UNIVERSITY OF TURKU

ABSTRACT

	Bachelor's thesis
x	Master's thesis
	Licentiate's thesis
	Doctor's thesis

Subject	Accounting and Finance	Date	16.6.2020
Author(s)	Olli Rihko	Student number	
		Number of pages	93 + Appendices
Title	The Inverted yield curve and hidden Markov models in predicting future bear markets		
Supervisor(s)	Prof. Luis Alvarez Esteban		

Abstract

Previous studies have shown very clearly that there is a connection between the inversion of the yield curve and future recessions measured as contractions in the GDP. The focus of this study was, however, the connection between financial markets and the inverted yield curve, which has not been studied as much as the production related dependences. Although it is difficult to draw a clear line between stock market crashes and the preceding inversions of the yield curve, the data suggests that the inverted yield curve is a sign of more volatile times ahead, which clearly heightens the risk of a major stock market downturn.

The empirical part of the study combined other factors and tools – hidden Markov models, credit spreads, S&P 500 returns, NY Fed recession probability model and the CAPE ratio – with the inverted yield curve in order to improve the predictability of future stock market crashes, which was then used to construct allocation strategies between the stock market and risk-free US Treasury rates.

The proposed strategies managed to beat the benchmark and the conducted transaction cost sensitivity analysis proves that the strategies are feasible at reasonable transaction cost levels. The greater annualized returns and lesser volatility of the strategies lead to significantly better risk-adjusted returns than what one would achieve using the simple buy and hold strategy. The obtained Sharpe ratios range from 0,531 to 0,824, whereas the Sharpe ratio of the benchmark total US stock market return is 0,499 during the (pseudo) out-of-sample forecasting period of 1986–2019. It should be noted that the analysis is conducted using past data and the true acid test of the strategies will be made during the current inversion period.

Although greater volatility leads to greater opportunities in the market, the effectiveness of the proposed strategies in this study is based on prudence and patience: the best option for the average investor could very well be to weather the storm in safer assets and then return to the stock market once the volatility of the market is lower again.

Key words	inverted yield curve, hidden Markov models, credit spreads, allocation strategies
Further information	

TURUN YLIOPISTO

TIIVISTELMÄ

	Kandidaatintutkielma
x	Pro gradu -tutkielma
	Lisensiaatintutkielma
	Väitöskirja

Oppiaine	Laskentatoimi ja rahoitus	Päivämäärä	16.6.2020
Tekijä(t)	Olli Rihko	Matrikkelinumero	
		Sivumäärä	93 + Liitteet
Otsikko	Kääntynyt korkokäyrä ja piilotetut Markov-mallit osakemarkkinoiden käännöksen ennustamisessa		
Ohjaaja(t)	Prof. Luis Alvarez Esteban		

Tiivistelmä

Aiempi tutkimus osoittaa selkeästi yhteyden kääntyneen korkokäyrän ja sitä seuraavan bruttokansantuotteen laskulla mitatun reaalitalouden laman välillä, mutta kyseisen korkokäyrän vaikutuksia rahoitus- ja ennen kaikkea osakemarkkinoilla ei ole akateemisesti tutkittu yhtä paljon kuin kansantaloustieteellisiä vaikutuksia. Vaikka selkeitä syy-seuraussuhteita osakemarkkinaromahdusten ja niitä edeltävien korkokäyrien kääntymisten välille on vaikea vetää, läpikäyty aineisto osoittaa kääntyvän korkokäyrän olevan merkki lähitulevaisuudessa kasvavasta volatiliteetista, joka kasvattaa olennaisesti vakavan osakemarkkinaromahduksen riskiä.

Tutkimuksen empiirinen osuus yhdistää kääntyvään korkokäyrään muita faktoreita ja työkaluja – kuten esimerkiksi piilotetut Markov-mallit, korkospreadit ja S&P 500 -tuotot –, joiden avulla pelkän korkokäyrän ennustevoimaa pyritään parantamaan sekä siten muodostamaan toimivia allokaatiostrategioita osakemarkkinoiden sekä riskittömän tuoton välillä.

Esitetyt strategiat onnistuivat voittamaan vertailuindeksinsä ja transaktiokustannuksista tehty herkkyysanalyysi osoittaa, että muodostetut strategiat ovat toteuttamiskelpoisia realistisilla transaktiokustannuksilla. Strategioiden vertailuindeksiä parempien keskimääräisten vuosituottojen sekä pienemmän volatiliteetin ansiosta niiden riskikorjattu tuotto on merkittävästi parempi kuin pelkästään vertailuindeksiin sijoittamalla. Tutkimuksessa esitettyjen strategioiden Sharpen luvut ovat välillä 0,531–0,824, kun taas vertailuindeksinä toimivan koko Yhdysvaltain osakemarkkinan kokonaistuoton Sharpen luku on 0,499 aikavälillä 1986–2019, joka toimii tutkimuksen (pseudo-)ennustejaksona. On kuitenkin huomattava, että strategioiden todellinen tulikoe tapahtuu vuonna 2019 tapahtuneen korkokäyrän kääntymisen jälkeisessä lähitulevaisuudessa.

Vaikka suurempi volatiliteetti johtaa suurempiin mahdollisuuksiin markkinoilla, niin tässä tutkimuksessa esitettyjen strategioiden tehokkuus perustuu ennen kaikkea kärsivällisyyteen ja varovaisuuteen: tavallisen sijoittajan paras vaihtoehto markkinoiden myllerryksessä voisi useimmiten olla hakeutuminen turvallisempien omaisuuserien pariin sekä palaaminen osakemarkkinoille myrskyn laantuessa.

Asiasanat	kääntynyt korkokäyrä, piilotetut Markov-mallit, korkospreadit, allokaatiostrategiat
Muita tietoja	

UNIVERSITY
OF TURKU
Turku School of
Economics

THE INVERTED YIELD CURVE AND HIDDEN MARKOV MODELS IN PREDICTING FUTURE BEAR MARKETS

Master´s Thesis
in Accounting and Finance

Author:
Olli Rihko

Supervisor:
Prof. Luis Alvarez Esteban

16.6.2020
Turku

Table of contents

List of figures

List of tables

1 INTRODUCTION

1.1 Motivation

Over the course of the 20th and the 21st century we have seen many stock market crashes, which are depicted by major downfalls of the whole stock market losing vast amounts of value in a matter of weeks. The attempts, therefore, to time the market and hedge from the dramatic dips have been of great interest among the market participants, although the attempts to time the market are useless in achieving additional return according to finance theory (see e.g. Malkiel 2007). However, timing the market is – and will probably be – an integral part of market participation, because sensing bigger trends and predicting future moves in markets are usually a part of active trading strategies, financial discussions and general speculation.

One of the most important or the most used tools to predict future recessions in finance over the last 30 years has been the slope of the yield curve, which is also referred to as the term spread. To put it short, the yield curve is inverted, when long-term interest yields fall below short-term yields, and this is usually seen as a sign of economic slowdown and diminished inflation expectations. The most used spread in academic studies is the difference between the US three-month treasury bill rate and the US ten-year bond rate, which provides the most evidence in the past data (e.g. Estrella & Hardouvelis 1991, 556–558; Adrian et al. 2019, 725–726). An inverted yield curve in the spread in question has occurred totally 8 times since 1961 and on 7 occasions the inversion has been followed by a recession (Adrian et al. 2019, 725), which solidifies the predictive power the inverted yield curve has. The only exception to the rule in 1966 was actually called a mini-recession or even a recession by some standards (Friedman 1970, 16–17).

The inverted yield curve has been the topic of various studies during the last 30 years, especially the effect on real economic activity has been studied numerous times. In this sense, it is no new idea to connect the inverted yield curve and recessions. The studied connections have been, however, mostly real economy based (GDP), whereas this study is focused on the effects of the inversion in financial markets. Furthermore, this study attempts to add some flavour to the previous studies with the inclusion of hidden Markov models and credit spreads. The credit spreads between safer and riskier assets can be seen as measures of uncertainty and restlessness in the market, which is why they are a good proxy for the general market sentiment, which is vital during volatile times in the market.

Predicting future recessions or timing the market in general is, as known, notoriously difficult and finding predictive patterns or formulas has been attempted by many, including some of the brightest minds in finance and economics. It might be therefore difficult to conquer new frontiers and find truly new information on the topic, but the aim of this study is to find new views and connections between different phenomena nevertheless. And if new views are not to be found, this study will hopefully bring deeper understanding of the inverted yield curve, hidden Markov models and market sentiment.

1.2 Objectives

The main objective of this study is to enhance the current knowledge of the effects of the inverted yield curve on financial economic activity and to drive forward the predictions of the related stock market downturns. The main method to reach the object is to include signals given by hidden Markov models into the equation. Whether or not this will be fruitful, remains to be seen.

In essence, the study can be summed up to three main questions, which are at the very core of this study:

- How does the inverted yield curve predict future stock market crashes?
- What other components or tools can be used to improve predictability?
- Is it possible to construct successful investment strategies using the inverted yield curve and other related components?

The first objective is to build a sound foundation to the study by viewing past predictions and methods, and thus acquiring a profound picture of the field. After that, various components are considered in order to push the models forward. The components and tools considered to improve the accuracy of the predictions are coming from the fields of time series analysis and behavioral finance, which can be seen as an effort to incorporate various fields and concepts in a single study.

In a broader sense, the aim of this study is to combine the implications of the yield curve, measured as the spread between different maturities; and hidden market processes reflected in credit spreads and stock market prices, which are tested using hidden Markov models (hereafter HMMs). With these three parts, the objective of the study is to find patterns and evidence, which could shed light on the necessary conditions of a future crash. In addition to this, all three factors are to be measured in quantitative terms in order to find statistical evidence between the factors and future recessions and not just speculate on different factors.

The empirical objective of the study is to find a way to combine the inverted yield curve with other factors to construct strategies that stay out of the stock market during unfavorable conditions. The main challenge of the study will presumably be to find the delicate balance between adding more variables and keeping the model simple enough.

Each new explaining factor should add some real value to the model, otherwise the study will be of no use.

Different factors and components are tested using allocation strategies between stock market and risk-free returns. As the focus of the study is more on market downturns than GDP-measured recessions, the final verdict of the different models and strategies will be made by simulating market operations with (pseudo) out-of-sample forecasting.

1.3 Previous studies

As stated before, the inverted yield curve has been studied quite intensively, especially in the 80s and 90s. Notable studies include e.g. the studies of Harvey (1986), Stock & Watson (1989), Estrella & Hardouvelis (1991) and Estrella & Mishkin (1996). More recent takes on the subject include e.g. Chinn & Kucko (2015) and Gogas et al. (2015). The specific views on different studies might be different, but they all are somehow connected first and foremost to the inverted yield curve and its implications on future real economic activity and especially economic downturns and recessions.

There are also other studies trying to predict future economic activity, downturns and stock market crashes using also other parameters and components than the inverted yield curve. There are obviously a lot of studies, books and articles, which focus on timing the market and predicting major turning points in the stock market, but relevant research to this study include the recent works of e.g. Chen (2009), Phillips et al. (2015), Kirschenmann et al. (2016) and Feldman & Liu (2018). These studies focus on the pricing of assets as well as macroeconomic variables, but the common factor is their goal to predict and anticipate financial turning points.

A common way to anticipate financial turning points is to see the market as a multi-regime system with different states, e.g. positive (bull) and negative (bear) states, which have their own characteristics such as mean and standard deviation. It should be also noted, however, that there could be even more states (e.g. negative – neutral – positive – exuberant), which define the underlying regimes.

The different regimes are usually modeled with switching-regime models such as a Markov switching model or a hidden Markov model. Relevant studies utilizing these models and financial applications include e.g. Guidolin & Timmermann (2007), Nguyen (2018) and Zhang et al. (2019). Even though most of the studies focus more on daily data, these methods are applied to monthly data in this study.

Other relevant works and studies, which are depicting the major turning points in the stock market in a more general sense include e.g. Kindleberger (2011), Shiller (2015) and Sornette (2017). These works are used as background information to give a sense of the bigger picture. After all, economic downturns are not just changes in variables, but also psychologic phenomena, which are caused by various reasons and affected by the overall sentiment of the market.

Especially the connection between credit-market sentiment, the business cycle and financial markets is of great interest in this study. The link between credit booms and future poor macroeconomic performance as a phenomenon of behavioral finance is studied by e.g. López et al. (2017), Bordalo et al. (2018) and Greenwood & Hanson (2013). These studies do not, however, comment very much on stock market sentiment, whereas this study emphasizes the link between credit and stock market sentiment.

The idea of the destabilizing effect of stability is relatively old. For example, presented in 1982, Hyman P. Minsky's financial instability hypothesis (FIH) states that prolonged periods of prosperity lead to speculative finance instead of hedge finance that is typical for periods after economic turmoil: Good economic performance leads to less conservative banking practices and excessive leverage, which in turn leads to "euphoric economy" and ultimately to debt deflation and economic crises (Keen 2013, 223–224). This study aims to find the underlying factors that reflect the changes in market sentiment, which could indicate that the destabilizing process has started.

The general conditions of the credit market are signaled by credit spreads, which in short price the credit risk between safer and riskier assets. When the credit market is stable, the spreads are narrow, but when the market is more stressed, the spreads widen. Relevant definitions and studies include e.g. Choudry (2013, 120), Mishkin (2019, 334–335) and Bektic´ & Regele (2018).

The aim of this section was to give a quick glance at different studies and fields relevant to this thesis. The most important studies, findings and models are presented in more detail in their own designated chapter.

1.4 Structure of the thesis

The thesis is divided into theoretical and empirical parts, accompanied with this introduction and conclusive remarks in the end. The theoretical foundation will be laid first followed by the presentation of the methodology and used data, which leads to the empirical part of the thesis.

The second chapter presents the characteristics of the yield curve, its inversion and the history of the inverted yield curve. The relevant studies are presented and discussed within the context of this study. Also, the stock market after the inversion is compared to the usual circumstances of the market and the most volatile periods traced. Furthermore, the most relevant predictive model based on the slope of the yield curve is presented and discussed.

The third chapter presents the hidden Markov model, which is the main method used in the study. We will go through the basics, the algorithms used and the particularities of the model. The aim is to present the model clearly, but concisely. Also, the empirical application of the model is discussed.

The fourth chapter presents the data used in the thesis. Data and the data sources are displayed, discussed and worked on. The applicability and reliability of the data used will also be considered. Methodological choices are also justified by fitting the data, methodology and selected statistical models together.

The fifth chapter focuses on the empirical side of the study: it presents the analysis and construction of the allocation strategies. The chapter considers various factors in detail and contains the *beef* of the thesis as it involves the training of the models used. Presumably, the chapter will also be the most extensive of the thesis.

The sixth chapter presents the results and further discusses the findings. Also, further testing and validation of the results is done. The chapter contains a lot of tables and figures to make the results easily digestible.

The seventh chapter summarizes the study, connects the major findings to other studies and their findings and draws a conclusion of the results obtained. If applicable, the results could be used also in an anticipatory manner considering the present data to predict the phase of the business cycle and possible future downturn. In addition, the study will be reflected and suggestions for future research made. The chapter offers also some concluding remarks.

2 THE INVERTED YIELD CURVE

The yield curve provides important information on the prices and maturities of the interest rates, which are essential in bond markets and banking in general. The yield curve is simply the interest rates of the same instrument – e.g. government bonds, money market instruments or off-balance sheet instruments – provided by the same issuer on different maturities plotted on the same curve. If the time dimension is added to the mix, the yields can be portrayed as a surface, which shows simultaneously different yield curves at different times.

The yield curve used and referred to in this study is the US Treasury yield curve, which contains the interest rates on different US Treasury fixed maturities ranging from one month to 30 years. However, this study focuses particularly on the spread between a 3-month Treasury bill and a 10-year Treasury bond, which is used in most of the previous studies.

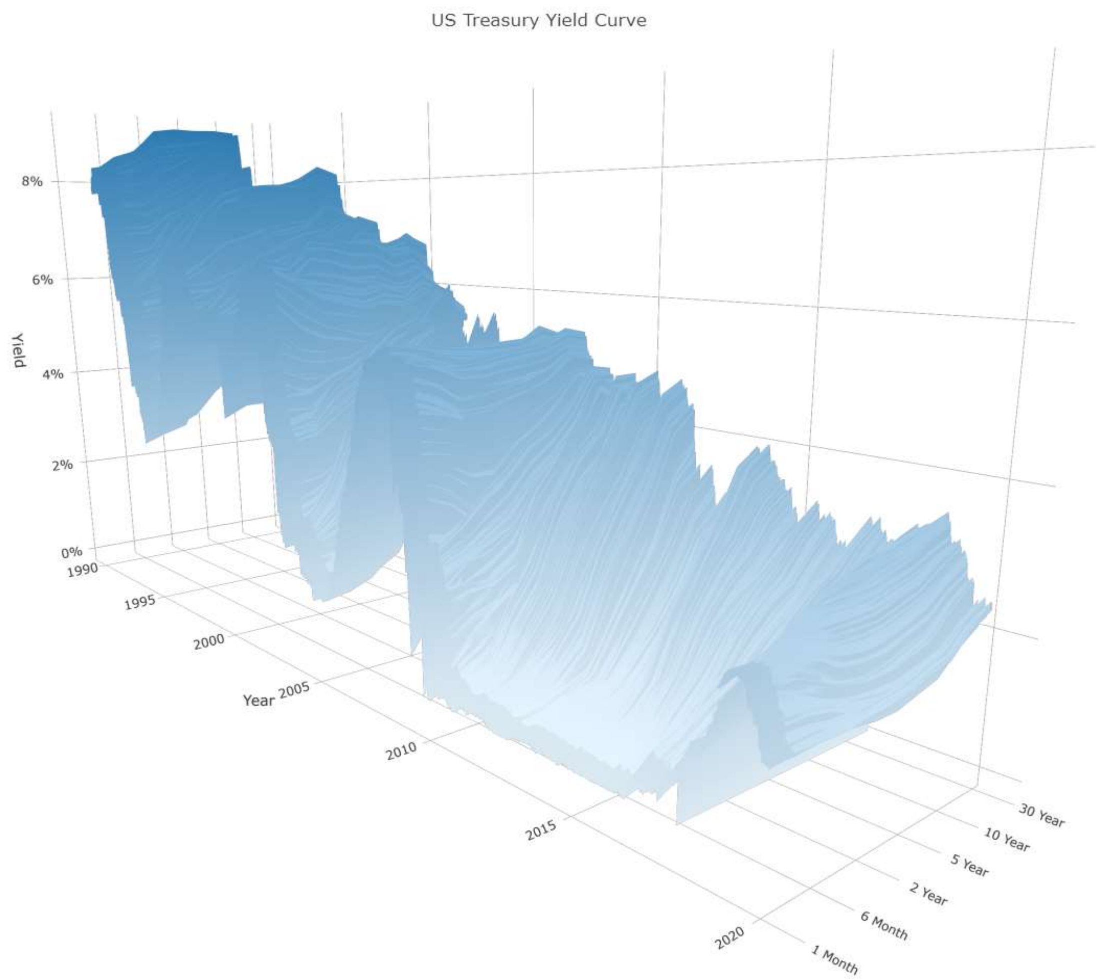

Figure 2.1 U.S. Treasury yields (1990-2020)

The yield curve can have different slopes: if short-term interest rates are higher than long-term, then the yield curve is inverted. In other words, the usually positive spread between longer and shorter interest rates turns negative. The inverted yield curve usually implicates lower expectations on future economic activity and inflation. (see e.g. Choudhry 2018, 132–136.; Veronesi 2010, 38–42.; Mishkin 2019, 175–185.)

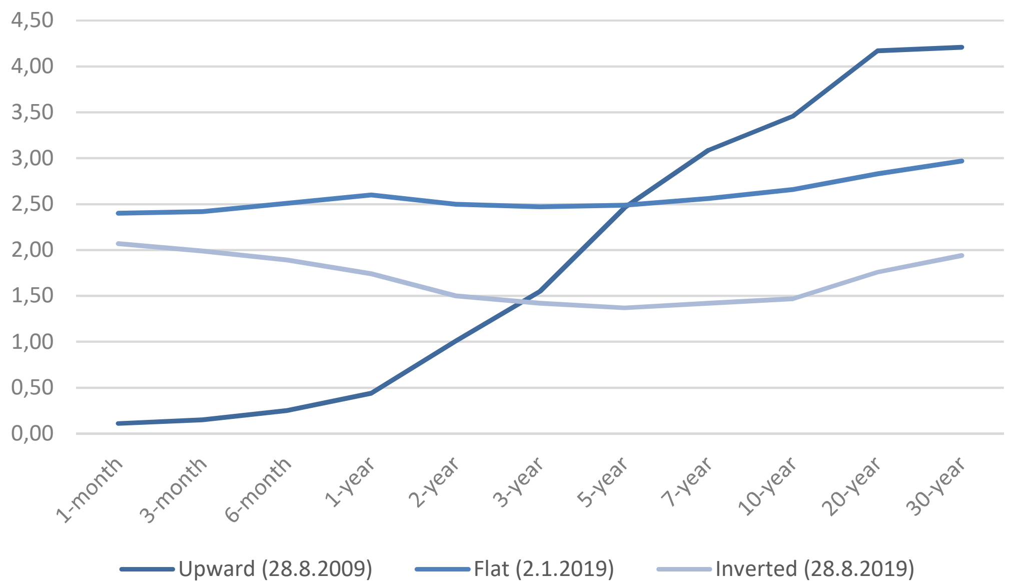

Figure 2.2 Different yield curves

The yield curve, its inversion and the relevant information it contains has been subject to intensive studying during the past and it is one of the most followed indicators in the financial industry. The evidence behind the phenomenon is strong and the most relevant studies regarding the inverted yield curve are presented next.

2.1 History

The first one, arguably, to study the inverted yield curve as a predictor of future economic slowdown was Campbell Harvey, who in his Ph.D. dissertation (1986) argued that the term structure could be used as predictor of US business cycles and thus future recessions. Sometimes acknowledged as “the godfather of the inverted yield curve” (e.g. Shen 2019), Campbell noticed the connection between negative spreads between different interest rate maturities and consumption growth. Although the spread is constructed as the difference between the one-year corporate yield and the yield on 90-day Treasury bills, the connections and the interdependence are clearly visible, as in later studies with different maturities and spreads. (Harvey 1986, 38–43.)

Other notable contributions to the topic in the late 80s and 90s include the studies of Stock & Watson (1989), who studied the effects of various factors as macroeconomic predictors. Stock & Watson found out that the inversion of the yield curve represented by the negative 10-year/1-year Treasury bond spread preceded an NBER defined cyclical peak by approximately one year on four occasions out of five. NBER refers to The National Bureau of Economic Research, which keeps official records of business cycle expansions and contractions in the United States (NBER 2019). Stock & Watson concluded that the bond term spread and the slope of the public debt curve are new potent variables to predict macroeconomic changes and could be added to the list of leading economic indicators. (Stock & Watson 1989, 383–385; 391.)

Probably the most popular model to predict future economic activity using the inverted yield curve was made in 1991, when Estrella and Hardouvelis developed a model, which gave exact probabilities for a recession in the upcoming 12 months based on the spread between 3-month T-bills and 10-year Treasury bonds. Additionally, the article discusses the driving forces behind the predictive qualities: current and expected future monetary policies probably are the main reason for the interdependence of recessions and yield curve inversions, but there could be other information in the slope of the yield curve as well. (Estrella & Hardouvelis 1991, 557–572.)

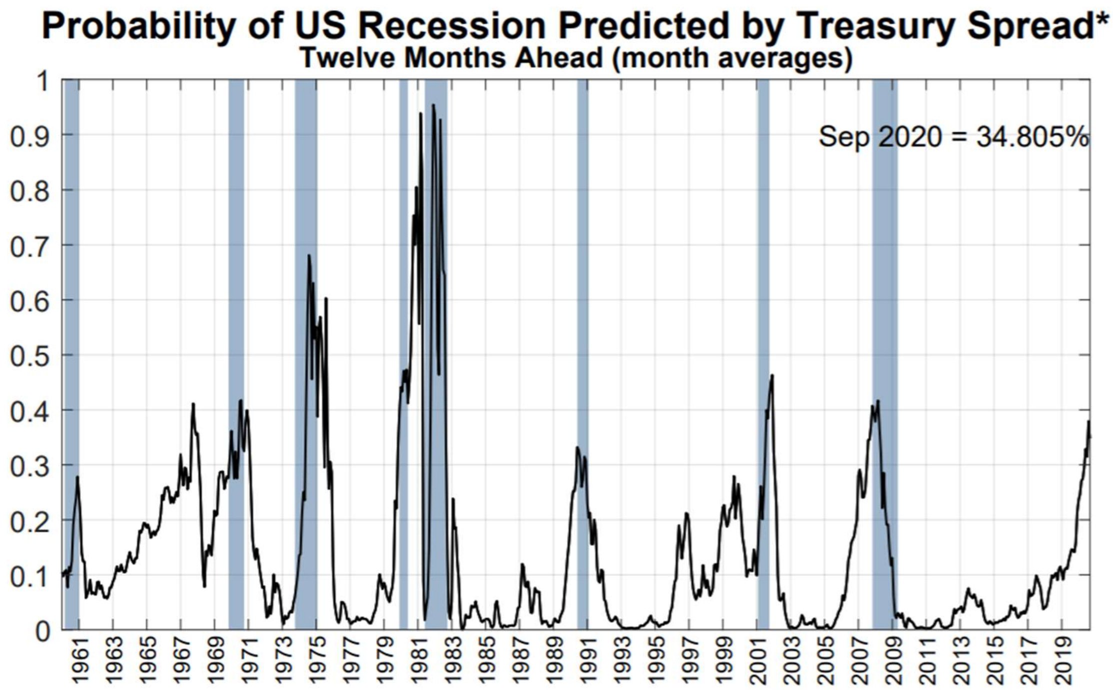

Figure 2.3 Probability of US recession (Federal Reserve Bank of New York 2019a)

The model developed by Estrella & Hardouvelis is still being used by the Federal Reserve Bank of New York, which updates the data and the probabilities of the model monthly (Federal Reserve Bank of New York 2019a). The model in question will be discussed in detail in Section 2.3.

The yield curve does have some predictive power on the GDP growth in general, but binary recession indicators, where a recession is coded as a one and other instances as zeros, are found to be more precise application of the inverted yield curve (Ang et al. 2006, 360–364). Furthermore, the fairly simple yield curve model for predicting recessions outperforms professional forecasters on many occasions. This imbalance can be seen as a puzzle, which has no clear explanation. One possible reason could be the economists' refusal to apply the past performance to the current situation, because of different times and situations. (Rudebusch & Williams 2009, 501–502.)

Notable more recent studies include e.g. Chinn & Kucko (2015), who studied the predictive power of the yield curve across different countries and timeframes. One key observation was that the yield curve possesses predictive power in more volatile times, but during steadier times there is no clear evidence of forecasting power for future GDP growth. However, the binary recession indicator estimations generate notable probabilities in the whole sample from 1970 to 2013. Also, The U.S. provides some kind of an exception: the statistical predictive power was not as significant in other countries as in the U.S. in six-month and one-year forecasts. (Chinn & Kucko 2015, 145–149; 151–152.)

During the recent years new techniques and methods have been used in forecasting future output from the yield curve. Gogas et al. applied a machine learning framework to try to improve the forecasting ability of the yield curve. According to their results, new sophisticated techniques can improve the overall forecasting accuracy. (Gogas et al. 2015, 641–644.)

2.2 Interpreting the evidence

The statistical connection between the inverted yield curve and future recessions is evident, but the mechanism and causality behind the phenomenon is contested. Moreover, the implications and importance of the inversion is not agreed on. Is the inversion relevant to future economic activity *per se* or is it just a reflection of the larger economy and monetary policy?

Although monetary policy is the most important factor on the slope of the yield curve, it is not the only determinant. Other relevant determinants, or relevant information contained by the slope of the yield curve, are expectations on real economic activity, expectations on inflation and expectations on market price movements in general. Furthermore, the term structure of interest rates can be used as a tool in monetary policy, but it shouldn't be any kind of a target. (Estrella & Mishkin 1997, 1376–1377; 1397–1399.)

Whether or not the role of the inverted yield curve itself is relevant, the evidence behind it as an indicator of future volatility is quite evident. If we look at the two-and-a-half-year-period after the inversion signal (the negative mean of the 3-month to 10-year spreads of the month in question), it can be seen in standard deviations that the stock market (as represented by the total market returns provided by Kenneth R. French (2020a)) during the period in question is way more volatile than outside this period:

Table 2.1 Descriptive statistics for continuously compounded monthly stock market returns

	Months 1-30 after inversion	***Other periods***	***All data 1959-2019***
Mean	0,00130	0,01132	0,00810
Standard Error	0,00341	0,00173	0,00161
Median	0,00727	0,01380	0,01257
Mode	-0,02153	-0,01765	0,01784
Standard Deviation	0,05225	0,03851	0,04361
Kurtosis	3,40718	8,44904	5,75958
Skewness	-0,34507	-1,03904	-0,77273
Range	0,34180	0,37794	0,41037
Minimum	-0,18814	-0,25670	-0,25670
Maximum	0,15366	0,12124	0,15366
First quartile	-0,02974	-0,01056	-0,01654
Third quartile	0,03912	0,03527	0,03590
Confidence Level(95%)	0,00672	0,00339	0,00316
n	235	497	732

The difference between the "danger zone" (months 1–30 after the inversion) and regular state of the market is even more evident if looked at the worst monthly stock market returns (1959–2019) and their occurrences during business cycles:

Table 2.2 Worst monthly simple market returns and the inverted yield curve (Months 1-30 after inversion)

Decline/month	*After inversion*	*Total*	*% of cases*
More than -10%	8	11	72,73 %
More than -9%	11	15	73,33 %
More than -8%	16	23	69,57 %
More than -7%	23	34	67,65 %
More than -6%	30	46	65,22 %
More than -5%	38	66	57,58 %
Total periods	235	732	32,10 %

As we can see, the probability of a particularly bad month is way more prominent after the inversion of the yield curve compared to more tranquil times. It should be, however, emphasized that the best monthly stock market returns are also significantly more probable after the inversion:

Table 2.3 Best monthly simple market returns and the inverted yield curve (Months 1-30 after inversion)

Increase/month	*After inversion*	*Total*	*% of cases*
More than 10%	6	12	50,00 %
More than 9%	8	16	50,00 %
More than 8%	10	24	41,67 %
More than 7%	17	42	40,48 %
More than 6%	26	68	38,24 %
More than 5%	41	107	38,32 %
Total periods	235	732	32,10 %

It is quite clear that the inversion of the yield curve leads to more volatile times in the stock market presenting both opportunities and threats to investors. The mechanism behind the inverted yield curve and economic slowdown might not be entirely clear, but the connection between the inversion and increased volatility in financial markets is evident.

Even though the major part of inverted yield curve relative analysis is based on the relationship between the curve and GDP growth, there are some studies which are linking the phenomenon to the stock market. The relationship between the inverted yield curve and bull/bear markets have been studied among other factors explaining the possible indicators for turning points between these different regimes (see e.g. Resnick & Shoesmith 2002 and Nyberg 2013). It should be noted, however, that the focus of the studies on the inverted yield curve has been more on predicting the future GDP (economics) than the stock market (finance).

2.3 The NY Fed inverted yield curve model

The benchmark of all yield curve based economic activity models and the model used extensively in this study is the model used by the Federal Reserve Bank of New York. As stated before, the model was developed by Estrella and Hardouvelis in 1991 as an attempt to quantify the effects of the slope of the yield curve as represented by the spread between U.S. Treasury 3-month and 10-year yields turning negative. The spread is calculated as simply the difference between the long-term, R^L, and short-term, R^S rates (Estrella & Hardouvelis 1991, 558):

$$SPREAD_t = R_t^L - R_t^S \tag{2.1}$$

Estrella and Hardouvelis use the yield curve as a predictor of a binary variable X, which simply indicates the presence ($X_t = 1$) or the absence ($X_t = 0$) of a recession and the probability of the recession is estimated from the slope of the yield curve four quarters earlier ($SPREAD_{t-4}$). The equation provides the probability of a National Bureau of Economic Research (NBER) defined recession in the current quarter t:

$$\Pr[X_t = 1 \mid SPREAD_{t-4}] = F(\alpha + \beta SPREAD_{t-4}) \tag{2.2}$$

where *Pr* denotes probability, *F* is the cumulative normal distribution and X_t equals unity during those quarters considered as official recessions by NBER. Furthermore, the model is a usual probit model and the unknown parameters of α and β are estimated by maximizing the log-likelihood of the function:

$$\log L = \sum_{X_t=1} \log F\,(\alpha + \beta \mathrm{SPREAD}_{t-4}) + \sum_{X_t=0} \log F\,(1 - \alpha - \beta \mathrm{SPREAD}_{t-4}) \tag{2.3}$$

Maximizing the function (2.3) with respect to unknown parameters α and β over the quarterly sample data from 1956:1 to 1988:4 leads to:

$$\Pr[X_t = 1 \mid SPREAD_{t-4}] = F(-0{,}56 - 0{,}78 SPREAD_{t-4}) \tag{2.4}$$

which gives a relatively simple formula with a pseudo-R^2 of 0,297 to estimate the probability of a recession given the yield curve slope a year earlier. (Estrella & Hardouvelis 1991, 562–565.)

The later articles by Estrella and other contributors have refined and sophisticated the model, but the main principle is still the same. Nowadays the model doesn’t use quarterly, but monthly data and the way to present the model is forward-looking rather than referring to past time, but functions remain basically the same. (see e.g. Estrella & Mishkin 1997; Estrella & Trubin 2006.)

The updated recession probability model presented by Estrella and Trubin (2006, 3.) is based on monthly observations of the slope of the yield curve and calculated in the same manner as the earlier version:

$$Recessm_{t+12} = F(\alpha + \beta sprd_t) \tag{2.5}$$

where the probability $Recessm_{t+12}$ is based on the same parameters as in the earlier version with the only difference being the notation ($Pr[X_t = 1 \mid SPREAD_{t-4}]$ vs. $Recessm_{t+1}$ and $SPREAD_{t-4}$ vs. $sprd_t$) and the time window perspective. As in the previous version, F is the cumulative normal distribution function.

One question to address with the evolution of the model is the estimation of the unknown parameters of α and β. As the parameters are estimated by maximizing the function (2.3), the parameters get different values with different time frames, naturally, because of different data sets. The values for different time frames are presented below (Table 2.4). It should be noted, that the values of α and β remain relatively stable and the point, where the yield curve is completely flat ($SPREAD_{t-4} = sprd_t = 0{,}00$), gives rather similar probabilities for a recession in 12 months.

Table 2.4 Estimated parameters of the NY Fed Recession probability model

	α	β	Pr [spread = 0]
1959-1988*	-0,56	-0,78	28,77 %
1959-2009**	-0,5333	-0,6333	29,69 %

* Estrella & Hardouvelis 1991, 562 ** Federal Reserve Bank of New York 2019a

Looking at the probabilities of a recession in 12 months when the spread between the long and short-term yields is zero (Table 2.4), one may ask why the probability of a recession is so low, if they have almost always lead to a recession? Even for most of the recessions (Figure 2.3), the probability reaches only around 0,4 at its highest point during the business cycle. Whether or not these figures are truly reflecting the odds of the US economy heading towards a recession, could be up for debate. On the other hand, recessions are relatively rare and complex events, which makes it almost impossible to issue exact probabilities.

3 HIDDEN MARKOV MODELS

Hidden Markov models (HMMs) are in short models, which try to uncover the hidden layers and states behind the level of observations. In other words, HMMs try to model the processes, which we really can't see or observe from the observable reality, but which are the driving forces behind the observations.

Hidden Markov models are used extensively in various fields of science, ranging from speech recognition (see e.g. Rabiner 1989) to weather forecasting (see e.g. Joshi et al. 2017) and biomedicine (see e.g. Anh et al. 2020). These models are also used heavily in finance research and recent examples include e.g. Nguyen (2018) and Zhang et al. (2019).

3.1 Preliminaries

Hidden Markov models are characterized first and foremost by two qualities: mixture models and Markov chains. HMMs allow the probability distribution of each observation to depend on the unobserved state (or regime) of a Markov chain, which allows them to accommodate both overdispersion and serial dependence. Furthermore, HMMs are rather simple, mathematically tractable and the likelihood is relatively straightforward to compute. (Zucchini & MacDonald 2009, 3–5.)

Mixture models are in general designed to accommodate unobserved heterogeneity in population: the population may consist of unobserved groups each having their own distinct distribution for the observed variable. An independent mixture distribution consists of a finite number of component distributions and a mixing distribution, which selects the state between the component distributions, which can be both discrete and continuous. For example, if there are m components, the probabilities assigned to different components are denoted by $\delta_1, \ldots, \delta_m$, probability or density functions are denoted by $p_1, \ldots, p_m$ and X denotes the random variable which has the mixture distribution, then it is easy to show that the probability or density function of X is given by:

$$p(x) = \sum_{i=1}^{m} \delta_i p_i(x) \tag{3.1}$$

Furthermore, for the discrete case the expectation of the mixture model can be given in terms of the component distributions. Letting Y_i denote the random variable with probability function p_i, it can be stated that

$$E(X) = \sum_{i=1}^{m} \Pr(C = i)E(X \mid C = i) = \sum_{i=1}^{m} \delta_i E(Y_i) \tag{3.2}$$

Even though the demonstration is made using discrete distributions, the same result holds also for a mixture of continuous distributions. (Zucchini & MacDonald 2009, 6–8.)

The second building-block of an HMM is Markov chains, which are restricted to discrete-time first order Markov chains in this study. A sequence of random variables $\{C_t : t \in \mathbb{N}\}$ is said to be a Markov chain, if it satisfies the Markov property

$$\Pr(C_{t+1} | C_t, \dots, C_1) = \Pr(C_{t+1} | C_t) \tag{3.3}$$

In other words, the history of the process is conditioned only on the most recent value of C_t since it reflects all the information of the past sequence. The Markov property can be seen the "first relaxation" of the assumption of independence, but the random variables $\{C_t\}$ are dependent in a specific way: the past and the future are dependent only through the present, which is also mathematically convenient. This feature is also referred to as the *memorylessness* of Markov chains. (Zucchini & MacDonald 2009, 16; Tsay & Chen 2019, 112.)

Important part of Markov chains are transition probabilities, in other words the conditional probabilities:

$$\Pr(C_{s+t} = j | C_s = i) \tag{3.4}$$

If these probabilities are independent of s, the Markov chain is (time-)homogenous. As stated before, only first order Markov chains are used in this study, hence only the matrix of one-step transition probabilities is used. If m denotes the number of states of the Markov chain, the matrix $\mathbf{\Gamma}$ is a square matrix of transition probabilities between the states with row sums of 1:

$$\mathbf{\Gamma} = \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1m} \\ \vdots & \ddots & \vdots \\ \gamma_{m1} & \cdots & \gamma_{mm} \end{pmatrix} \tag{3.5}$$

The matrix $\mathbf{\Gamma}$ is referred to as the transition probability matrix, which states the probabilities of staying in the same state or moving to another. Furthermore, a Markov chain with transition probability matrix $\mathbf{\Gamma}$ has a stationary distribution $\boldsymbol{\delta}$ (a row vector of nonnegative elements), if $\boldsymbol{\delta}\mathbf{\Gamma} = \boldsymbol{\delta}$ and $\boldsymbol{\delta}\mathbf{1}' = 1$. The key result is that a Markov chain has a stationary distribution, if it is irreducible and aperiodic. (Zucchini & MacDonald 2009, 17–18; Tsay & Chen 2019, 111–113.)

3.2 The model

A hidden Markov model is, simply put, "a doubly stochastic process with an underlying stochastic process that is not directly observable (it is "hidden") but can be observed only through another stochastic process that produces the sequence of observations" (Cappé et al. 2005, 42). An HMM is a special kind of dependent mixture model $\{X_t : t \in \mathbb{N}\}$, which consists of two parts. With $\mathbf{X}^{(t)}$ and $\mathbf{C}^{(t)}$ denoting the histories of the series from time 1 to time t, it is possible to summarize the simplest model of this kind by:

$$\Pr(C_t | C^{(t-1)}) = \Pr(C_{t+1} | C_t), t = 2,3, \ldots \qquad (3.6)$$

$$\Pr(X_t | X^{(t-1)}, C^t) = \Pr(X_t | C_t), t \in \mathbb{N} \qquad (3.7)$$

The first part (4.6), the unobserved parameter process $\{C_t : t \in \mathbb{N}\}$ satisfying the Markov property is not directly observable, and the second part (4.7), the state-dependent process $\{X_t : t \in \mathbb{N}\}$ is observable and depends on the parameter process. When C_t is known, the distribution of X_t is dependent only of the current state C_t and not on the previous states or observations. (Zucchini & MacDonald 2009, 30.)

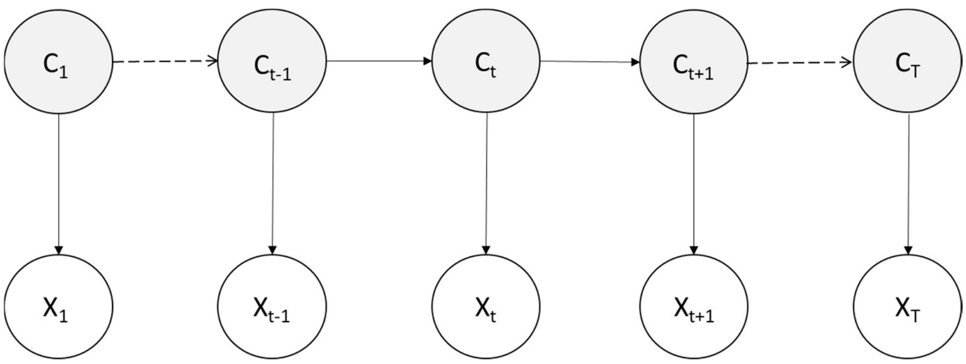

Figure 3.1 Graph of a basic HMM

There are also other ways to present hidden Markov Models graphically. One of the best examples is provided by Zucchini & MacDonald (2009, 31.), which is depicted in Figure 3.2. The model consists of state-dependent distributions p_1 and p_2, stationary distribution $\boldsymbol{\delta}$ and the probability transition matrix $\boldsymbol{\Gamma} = \begin{pmatrix} 0{,}9 & 0{,}1 \\ 0{,}3 & 0{,}7 \end{pmatrix}$. It should be noted, however, that as opposed to an independent mixture, here the distribution of C_t does depend on C_{t-1}. Nevertheless, there is a different distribution, discrete or continuous, for each state, just as in independent mixtures. It should be also noted, that stationarity of the chain $\{C_t\}$ is not needed in applications, only homogeneity is needed (Zucchini & MacDonald 2009, 75–76).

For other good, clarifying examples with graphs of HMMs please see e.g. Jurafsky & Martin (2019, 2 –3) or especially Rabiner (1989, 260), which is referred to in this study and plenty of other academic papers.

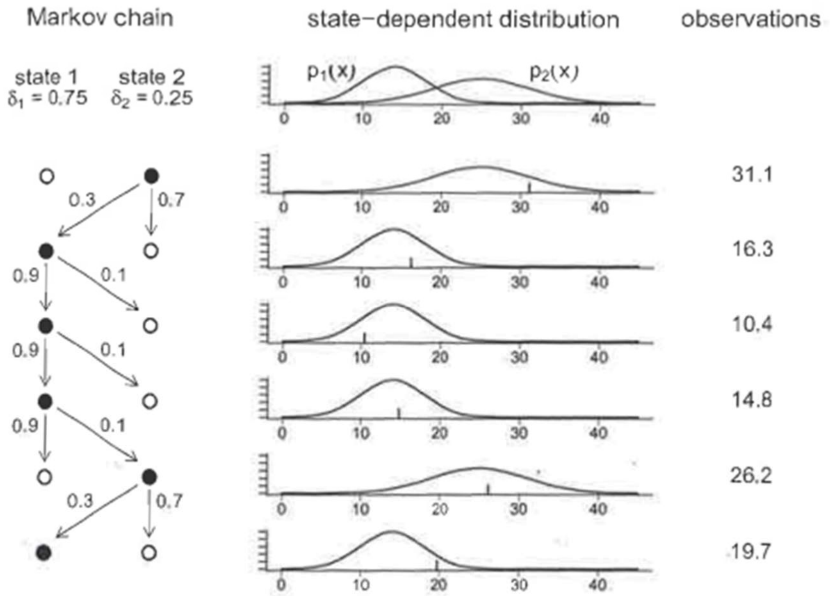

Figure 3.2 The observation-generating process in a two-state HMM

The basic elements of an HMM have been already discussed to some extent in this paper, but before moving forward, it is important to define them in a bit more formal and precise way. This is demonstrated by the classification and the procedure suggested by Rabiner (1989, 260–261). An HMM is specified by the following elements [the symbols used in this study in parenthesis]:

1. N, the number of states in the model (individual states are denoted as S = {S_1, S_2, …, S_N} and the state at time t is denoted by q_t) [individual states $C = \{C_1, C_2, \dots, C_m\}$ and the state at time t by c_t]
2. M, the number of distinctive observation symbols in the model e.g. vocabulary in speech recognition (individual symbols denoted as V = {v_1, v_2, …, v_M}) [NB! As this study uses continuous distributions, these symbols are not really used.]
3. The state transition probability distribution A = {a_{ij}} [equivalent to the transition probability matrix **Γ** presented earlier]
4. The observation symbol probability distribution in state j, B = {$b_j(k)$}, where $b_j(\mathrm{k}) = P[v_k \text{ at } t \mid q_t = S_j],\ 1 \leq j \leq N;\ \ 1 \leq k \leq M$

[NB! Once again, because of continuous distributions, the observation symbol probabilities in this study are state-dependent gaussian distributions $p_1, p_2, \dots, p_m$]

5. The initial state distribution $\pi = \{\pi_i\}$, where

$$\pi_i = P[q_1 = S_i], \quad 1 \leq i \leq N$$

If appropriate values of N, M, A, B and π are given, the HMM can generate an observation sequence of $O = O_1, O_2, \dots, O_T$ $[X = X_1, X_2, \dots, X_T]$ with the following procedure:

1. Choose an initial state $q_1 = S_i$ according to the initial state distribution π
2. Set $t = 1$
3. Choose $O_t = v_k$ according to the symbol probability distribution in state S_i e.g. b_i(k)
4. Transit to a new state $q_{t+1} = S_j$ according to the state transition probability distribution for state S_i, e.g. a_{ij}
5. Set t = t+1; return to step 3 if t < T, otherwise terminate the procedure

Even though the procedure explained by Rabiner focuses especially on HMMs in terms of speech recognition and discrete distribution, the basic principles are the same in other fields of study and continuous distributions as well, although there are some differences. Speech recognition uses *left-to-right HMMs* that start in a particular initial state, travels through many intermediate states and terminates in a final state. These HMMs are not ergodic and they produce a sequence of output, which usually has a random length. (Cappé et al. 2005, 33–34.)

Also, *left-to-right HMMs* are characterized by a large number of states, whereas *ergodic HMMs* have, in contrast, small number of states, if the state space is finite. Ergodic HMMs, which are used extensively in e.g. economics and finance, are usually at least irreducible and can have a unique stationary distribution, which allows for periodicity. Even though these models have differences, they are at their core very similar. The computational side of both *left-to-right* and *ergodic* HMMs have especially very much in common, which is beneficial in using algorithms related to these models. For example, the expectation maximization algorithm can be implemented similarly whatever the Markov chain structure. (Cappé et al. 2005, 33–34.)

3.3 Filtering and smoothing

In order to work empirically with hidden Markov models and the probabilities they produce, it is important to construct the appropriate measures for marginal probabilities and marginal distributions of each state in the model. Marginal probabilities are given by *filtering* of the state of the model, which is based on forward probabilities, also known as $\boldsymbol{\alpha_t}$. The marginal distributions are obtained by using both forward and backward probabilities, which are denoted by $\boldsymbol{\beta_t}$. This process is called *smoothing*. (Tsay & Chen 2019, 351–353; Zucchini & MacDonald 2009, 38–39.)

The forward probabilities can be defined for t = 1, 2, …, T as the (row) vector of $\boldsymbol{\alpha_t}$ as follows:

$$\alpha_t = \boldsymbol{\delta P}(x_1)\boldsymbol{\Gamma P}(x_2) \dots \boldsymbol{\Gamma P}(x_t) = \boldsymbol{\delta P}(x_1)\prod_{s=2}^{t}\boldsymbol{\Gamma P}(x_s), \tag{3.8}$$

where $\boldsymbol{\delta}$ denotes the initial distribution of the Markov chain, $\boldsymbol{\Gamma}$ the transition probability matrix and $\mathbf{P}$ is the matrix of state-dependent distributions:

$$P(x) = \begin{pmatrix} p_1(x) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & p_m(x) \end{pmatrix} \tag{3.9}$$

The backward probabilities can also be defined for t = 1, 2, …, T as the vector of $\boldsymbol{\beta_t}$ as follows:

$$\beta_t' = \boldsymbol{\Gamma P}(x_{t+1})\boldsymbol{\Gamma P}(x_{t+2}) \dots \boldsymbol{\Gamma P}(x_T)\mathbf{1}' = \left(\prod_{s=t+1}^{T}\boldsymbol{\Gamma P}(x_s)\right)\mathbf{1}', \tag{3.10}$$

where $\boldsymbol{\Gamma}$ and $\mathbf{P}$ have the same meanings as in equation 3.8. The backward probability function carries the convention that an empty product is the identity matrix, which is why $t = T$ yields $\beta_T = 1$.

The forward and backward probabilities have useful applications and it is no coincidence that they are named as probabilities. $\boldsymbol{\alpha_t}$ is in fact the vector of joint probabilities of the sequence **X**, in other words $\alpha_t(j)$, the *j* th component of $\boldsymbol{\alpha_t}$ is the joint probability

$$\Pr(X_1 = x_1, X_2 = x_2, \dots, X_t = x_t, C_t = j) \tag{3.11}$$

Similarly, $\boldsymbol{\beta_t}$ is the vector of conditional probabilities of the sequence **X**, in other words $\beta_t(j)$, the *j* th component of $\boldsymbol{\beta_t}$ is the conditional probability

$$\Pr(X_{t+1} = x_{t+1}, X_{t+2} = x_{t+2}, \dots, X_T = x_T \mid C_t = j) \tag{3.12}$$

It will then follow that, for t = 1, 2, …, T

$$\alpha_t(j)\beta_t(j) = \Pr(X^{(T)} = x^{(T)}, C_t = j) \tag{3.13}$$

Furthermore, these qualities can be used in filtering and smoothing. Given the observations $x_1, \dots, x_T$, the following set of statements can be made about the present and the past (respectively), letting L_T denote the likelihood:

$$L_T \Pr(C_t = i \mid X^{(T)} = x^{(T)}) = \begin{cases} \alpha_T(i) & \text{for } t = T \quad \text{filtering} \\ \alpha_t(i)\beta_t(i) & \text{for } 1 \leq t \leq T \quad \text{smoothing} \end{cases} \tag{3.14}$$

The filtering and smoothing parts are indeed state probabilities for present and past states. Hence, the conditional distribution of C_t given the observations can be obtained for i = 1, 2, …, m as:

$$\Pr(C_t = i \mid X^{(T)} = x^{(T)}) = \frac{\Pr(C_t = i \mid X^{(T)} = x^{(T)})}{\Pr(X^{(T)} = x^{(T)})} = \frac{\alpha_t(i)\beta_t(i)}{L_T} \tag{3.15}$$

Since $\beta_T(i) = 1$ for all *i,* both filtering and smoothing can be combined if looked only at the present state of the process. (Zucchini & MacDonald 2009, 32; 59–61; 86.)

3.4 Baum-Welch algorithm

Rabiner (1989, 261) presents the three basic problems for HMMs, which present a good framework for assuring the usefulness in real-world applications:

1. Given the observation sequence X and the model $\lambda = (A, B, \pi)$, how do we efficiently compute $P(X|\lambda)$, the probability of the observation sequence, given the model?
2. Given the observation sequence X and the model λ, how do we choose a corresponding state sequence C which is optimal in some meaningful sense (e.g. best "explains" the observations)?
3. How do we optimize the model parameters $\lambda = (A, B, \pi)$ to maximize $P(X|\lambda)$?

Problem 1 can be solved basically with just the forward probability ($\boldsymbol{\alpha_t}$) calculation already presented in Section 3.3 and the second problem can be solved various ways and is discussed in Section 3.5. The third problem, which is also described as the most difficult of the three, cannot be solved analytically, but with an expectation-maximization (EM) algorithm, which is discussed in detail in this section. (Rabiner 1989, 261–264.)

The EM algorithm is a way of performing maximum likelihood estimation when some of the data are missing and the algorithm exploits the fact the log-likelihood for complete data may be straightforward to calculate even though the likelihood for the observation data is not. This iterative process tries to find the log-likelihood for the parameters of interest $\boldsymbol{\theta}$ using both observed and missing data, also known as the complete-data log-likelihood (CDLL). (Zucchini & MacDonald 2009, 63–64.)

The iterative process is based on two different steps and the algorithm can be presented informally after choosing the starting values for the parameters $\boldsymbol{\theta}$ as follows:

- **E step:** Compute the conditional expectations of the missing data given the current estimate of $\boldsymbol{\theta}$ and the observations
- **M step:** Perform ML estimation of $\boldsymbol{\theta}$ just as if there is no missing data at all, in other words maximize the CDLL with respect to $\boldsymbol{\theta}$

The two steps are repeated until the algorithm has converged according to the chosen criterion, e.g. until the resulting change in $\boldsymbol{\theta}$ or CDLL is less than some threshold. The resulting value of $\boldsymbol{\theta}$ is then a stationary point of the likelihood of the observed data. There are some cases, however, when the stationary point found by the iterative process is not the global maxima, but the local maxima or a saddle point. (Little & Rubin 2002, 166–167; Zucchini & MacDonald 2009, 63–64.)

In the case of HMMs the EM algorithm is known as the Baum-Welch algorithm because of the groundbreaking work of Leonard Baum and his co-authors (see e.g. Baum & Petrie 1966; Baum et al. 1970) generalizing the forward-backward smoothing approach also to models, where the state space is not discrete anymore (Cappé et al. 2005, 52).

Originally developed for the needs of codebreaking, the algorithm provides the framework for estimating the parameters of an HMM. Let the states $c_1, c_2, \dots, c_T$ followed by the Markov chain be defined by the zero-one random variables as follows:

$$u_j(t) = 1 \quad \text{if, and only if,} \quad c_t = j, \ (t = 1, 2, \dots, T) \quad \text{and} \tag{3.16}$$

$$v_{jk}(t) = 1 \quad \text{if, and only if,} \quad c_{t-1} = j \ and \ c_t = k, \ (t = 2, 3, \dots, T) \tag{3.17}$$

With this notation, the complete-data log-likelihood (CDLL) of an HMM, in other words the log-likelihood of the observations $x, x_2, \dots, x_T$ plus the missing data $c_1, c_2, \dots, c_T$ is given by:

$$\log\big(Pr\big(x^{(T)}, c^{(T)}\big)\big)$$
$$= \sum_{j=1}^{m} u_j(1)\, log\delta_j + \sum_{j=1}^{m}\sum_{k=1}^{m}\left(\sum_{t=2}^{T} v_{jk}(t)\right)\log\gamma_{jk}$$
$$+ \sum_{j=1}^{m}\sum_{t=1}^{T} u_j(t)\log p_j(x_t) = term\ 1 + term\ 2 + term\ 3, \tag{3.18}$$

where $\boldsymbol{\delta}$ is the *initial* distribution of the Markov chain, the distribution of C_1, not necessarily the stationary distribution as stated before (the forward probabilities as computed for an HMM does not assume stationarity of $\{C_t\}$ and the backward probabilities are not affected by the (non-)stationarity of $\{C_t\}$). The EM (Baum-Welch) algorithm proceeds as follows:

- **E step:** Replace all the quantities $v_{jk}(u)$ and $u_j(t)$ by their conditional expectations given by the observations $x^{(T)}$ and current parameter estimates:

$$\hat{u}_j(t) = \Pr(C_t = j \mid x^{(T)}) = \frac{\alpha_t(j)\beta_t(j)}{L_T} \qquad \text{and} \tag{3.19}$$

$$\hat{v}_{jk}(t) = \Pr(C_{t-1} = j, C_t = k \mid x^{(T)}) = \frac{\alpha_{t-1}(j)\gamma_{jk}p_k(x_t)\beta_t(k)}{L_T} \tag{3.20}$$

- **M step:** After replacing $u_j(t)$ and $v_{jk}(t)$ with the estimates $\hat{u}_j(t)$ and $\hat{v}_{jk}(t)$, maximize the CDLL (equation 3.18) with respect to the parameters $\boldsymbol{\delta}$, $\boldsymbol{\Gamma}$ and the parameters of the state-dependent distributions. Equation 3.18 splits very conveniently into three parts, which can be maximized separately, since term 1 depends only on the initial distribution $\boldsymbol{\delta}$, term 2 on the transition probability matrix $\boldsymbol{\Gamma}$ and term 3 on the state-dependent parameters:

    1. $\sum_{j=1}^{m} u_j(1)\, log\delta_j$ with respect to $\boldsymbol{\delta}$
    2. $\sum_{j=1}^{m}\sum_{k=1}^{m}\big(\sum_{t=2}^{T} v_{jk}(t)\big)\log\gamma_{jk}$ with respect to $\boldsymbol{\Gamma}$
    3. $\sum_{j=1}^{m}\sum_{t=1}^{T} u_j(t)\log p_j(x_t)$ with respect to the state-dependent parameters of the j th state

The maximization of the third term depends very much on the assumed underlying state-dependent distributions. Even though the distributions can be a mix of various different distributions, normal distributions are used in this study, because of their qualities and computability. For a normal-HMM the state-dependent density is of the form

,

$$p_j(x) = \left(2\pi\sigma_j^2\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma_j^2}\left(x-\mu_j\right)^2\right) \tag{3.21}$$

and the maximizing values of the state-dependent parameters μ_j and σ_j^2 are:

$$\hat{\mu}_j = \sum_{t=1}^{T} \hat{u}_j(t)x_t \,/ \sum_{t=1}^{T} \hat{u}_j(t) \qquad \text{and} \tag{3.22}$$

$$\hat{\sigma}_j^2 = \sum_{t=1}^{T} \hat{u}_j(t)(x_t - \hat{u}_j)^2 \,/ \sum_{t=1}^{T} \hat{u}_j(t) \tag{3.23}$$

The Baum-Welch algorithm gives the tools to estimate the parameters of a hidden Markov model, given that the Markov chain of an HMM is homogenous but not necessarily stationary. For more complex observations, the Newton-Raphson method can be used to update the parameters, whereas for simpler cases the maximum can be obtained analytically. (Zucchini & MacDonald 2009, 59–67; Tsay & Chen 2019, 356–358.)

The EM algorithm answers the problem 3 of optimizing the model parameters as presented by Rabiner (1989, 261). We now move on to the decoding problem (Problem 2) and present the Viterbi algorithm.

3.5 Viterbi algorithm

It is vital in many HMM appliances to find out the most likely state at a certain time or to find out the most probable sequence of the unobservable Markov chain C given the observations X. There are different methods to discover the most likely states or sequences, but these methods are all in general called decoding.

The method of finding the most likely state at a certain time is called *local decoding*, which is based on the conditional distribution of C_t given the observations. For each time $t \in \{1, \dots, T\}$ it is therefore possible to determine the distribution of the state C_t given the observations $x^{(T)}$, which for *m* states is a discrete probability distribution with support $\{1, \dots, m\}$. For each $t \in \{1, \dots, T\}$ the most probable state i_t^*, given the observations, is:

$$i_t^* = \underset{i=1,\dots,m}{\operatorname{argmax}} \Pr(C_t = i \mid X^{(T)} = x^{(T)}) \tag{3.24}$$

This approach provides the most likely state for each *t* by maximizing the conditional probability $\Pr(C_t = i \mid X^{(T)} = x^{(T)})$. (Zucchini & MacDonald 2009, 82–83.)

Global decoding, on the other hand, refers to methods which try to find the most likely sequence of states, given the observations and the model. Whereas in local decoding the objective is to maximize the conditional probability over *i* for each *t*, in global decoding one seeks the sequence of $c_1, c_2, \dots, c_T$ which maximizes the conditional probability:

$$\Pr\left(C^{(T)} = c^{(T)} \,\middle|\, X^{(T)} = x^{(T)}\right) \tag{3.25}$$

Or more conveniently the joint probability:

$$\Pr\left(C^{(T)}, X^{(T)}\right) = \delta_{c_1} \prod_{t=2}^{T} \gamma_{c_{t-1},c_t} \prod_{t=1}^{T} p_{c_t}(x_t) \tag{3.26}$$

Maximizing over all possible state sequences $c_1, c_2, \dots, c_T$ would require m^T function evaluations, which clearly isn't feasible for longer series. There is, however, a dynamic programming algorithm, which makes it more efficient to determine the most likely sequence of states. In the HMM literature, this algorithm which makes it possible to efficiently compute the *a posteriori* most likely sequence of states is known as the *Viterbi algorithm* after Viterbi (1967). (Cappé et al. 2005, 125-127; Zucchini & MacDonald 2009, 83–86.)

In order to construct the Viterbi algorithm, we first define

$$\xi_{1i} = \Pr(C_1 = i, X_1 = x_1) = \delta_i p_i(x_i) \tag{3.27}$$

And for t = 2, 3, …, T

$$\xi_1 i = \max_{c_1, c_2, \dots, c_{t-1}} Pr(C^{(t-1)} = c^{(t-1)}, C_t = i, X^{(T)} = x^{(T)}) \tag{3.28}$$

Furthermore, it can be shown that the probabilities ξ_{tj} satisfy the following recursion for $t = 2, 3, \dots, T$ and $i = 1, 2, \dots, m$:

$$\xi_{tj} = \left(\max_i\left(\xi_{t-1,i}\gamma_{ij}\right)\right) p_j(x_t), \tag{3.29}$$

which provides an efficient way to compute the $T \times m$ matrix of values ξ_{tj}, as the computational effort is linear in *T*. The required maximization of the sequence of states $i_1, i_2, \dots, i_T$ can therefore be determined recursively from

$$i_T = \underset{i=1,\dots,m}{\operatorname{argmax}}\, \xi_{Ti} \tag{3.30}$$

and for $t = T - 1, T - 2, \dots, 1$, from

$$i_t = \underset{i=1,\dots,m}{\operatorname{argmax}} \left(\xi_{ti} \gamma_{i,i_{t+1}}\right) \tag{3.31}$$

It should be also noted that the quantity to be maximized in global decoding is simply a product of probabilities, it is possible to maximize its logarithm, which prevents numerical underflow. Furthermore, the Viterbi algorithm is applicable to both stationary and non-stationary underlying Markov chains. (Cappé et al. 2005, 125-127; Tsay & Chen 2019, 355–356; Zucchini & MacDonald 2009, 83–86.)

To summarize, global decoding aims to find out the most likely path and it is most recommended for applications in which the path itself is important and a part of the analysis (Zucchini & MacDonald 2009, 86) whereas local decoding provides the conditional probabilities for each state at each time. Both approaches are used in this study to decode hidden Markov models.

3.6 Empirical applications

As stated before, hidden Markov models have many empirical applications in various fields of science. Their prominence in finance has grown into new levels during the 2000s and there are many applications for different topics (see e.g. Mamon & Elliott 2007) with a special focus on the predicting financial time series, especially the stock market (see e.g. Zhang et al. 2019). This study relies on the same principles as the HMMs used in this study are applied to various financial time series.

The HMMs will be built on monthly data and they are updated monthly as if the data would have become available in real time. In other words, the starting point *(t =1)* will be the same, but the last period *(t = T)* will be updated each month as new figures become available. Figure 3.3 presents an example on the S&P 500 log-return series:

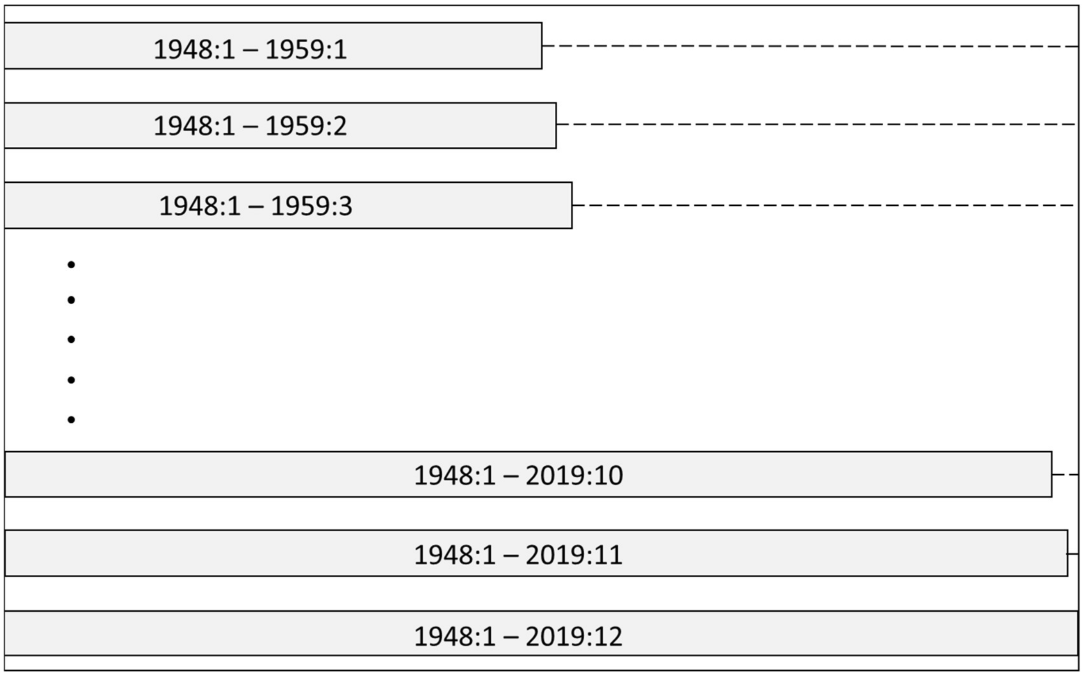

Figure 3.3 S&P 500 HMM parameter estimation windows

The latest estimations will be used as the basis for predicting the regimes in *T+1* and thus the direction of the time series in general. The HMM framework presented in this chapter will be applied to stock market returns, credit spread differences and other possible financial time series. The model used can be described as a gaussian mixture hidden Markov model (sometimes stated GMM-HMM), which is fitted to the data using the Baum-Welch algorithm (Tsay & Chen 2019, 358).

The data used in the study will be presented in detail in the following chapter. Furthermore, the models with a different number of states are also evaluated and chosen according to different information criteria for each time series.

The empirical hidden Markov models used in the study are based on the implementation of Tsay & Chen (2019, 358–365) and the R package HiddenMarkov (Harte 2017). The modifications of the code are those of the author and the codes are presented in Appendices 4 and 5. Empirical analysis is conducted using RStudio and MS Excel.

4 DATA

The data of the study consists of multiple time series, which are based on monthly observations of the available data. Since the study focuses on U.S. markets, the data will be easily available. The available time series begin from 1959 and the main emphasis will be on the data available before 1986, which will be used to build different trading strategies, in particular the period between 1972 and 1985 will be used as the testing/training period for different models and strategies.

The data from 1986 to 2019 is used in a (pseudo) out-of-sample forecast focusing on the latest recessions and market crashes of 1990, 2000 and 2008. This is done to test the strategies, which are constructed using the data before 1986.

4.1 The yield curve

The yield curve that is used in the study is based on the U.S. Department of the Treasury yield curve rates. However, the data in the study is provided by the Federal Reserve Bank of New York, which keeps a monthly updated database of the term spread between the 3-month and 10-year treasury yields on its website (Federal Reserve Bank of New York 2019c). The time series is constructed of monthly averages in order to get rid of daily changes, which can be substantial and therefore risk the persistence of the inverted yield curve signal (Federal Reserve Bank of New York 2019b). As the data is coming straight from the source, the reliability of the data shouldn't be an issue.

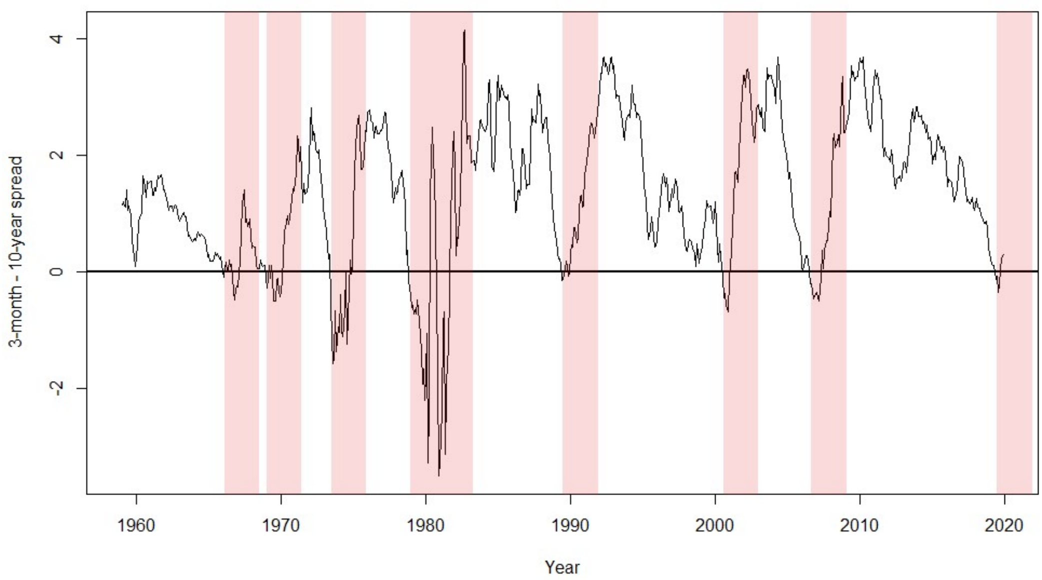

Figure 4.1 The 3-month – 10-year U.S Treasury yield with highlighted "danger zones" (Months 1-30 after inversion)

The danger zones refer to the periods after the inversion and they are as a concept different than the NBER recessions highlighted in Figure 2.3. The NBER-denominated recessions are GDP-based measures and specified after-the-fact, i.e. there is lag in identifying the recessions, whereas the highlighted danger zones in Figure 4.1 are just periods of higher volatility and lower returns in the S&P 500 as described in Section 2.2.

The yield spread series used in the study is from the beginning of 1959 to the end of 2019 consisting in total of 732 data points. The level of the spread itself is used as an indicator of market turbulence to some extent, but the signal of the yield curve inverting (term spread turning negative) is used more profoundly.

4.2 The stock market

The stock market in this study is represented by the Standard & Poor’s 500 Price Return Index, which covers the leading 500 U.S. publicly traded companies and doesn’t take the dividends of the companies into account. The price index is seen to be more reflective of the market and market pricing itself, why it’s considered to be a more suitable option in forecasting.

The S&P 500 was founded in 1927, but the timeframe for the data used in this study starts from 1948, because we consider the stock market to be so dynamic and ever-changing that the oldest data points are not as relevant as the newer ones. As the yield curve spread time series starts from 1959, the 132 data points (1948–1958) before it was considered necessary for the model to have reliable indications right from the start. The S&P 500 data is from the Eikon database, which is maintained by Refinitiv and is one of the leading information systems of the financial industry.

The stock market time series is based on monthly closing prices and the return series is constructed as a time series of continuously compounded returns (log-returns). The series consists of 864 data points and it is shown in Figure 4.2a. The histogram of the series is shown in Figure 4.2b and also the general price level of the index during the forecasting period (1986-2019) is shown in Figure 4.2c. The table of descriptive statistics can be found in the next section in Table 4.1.

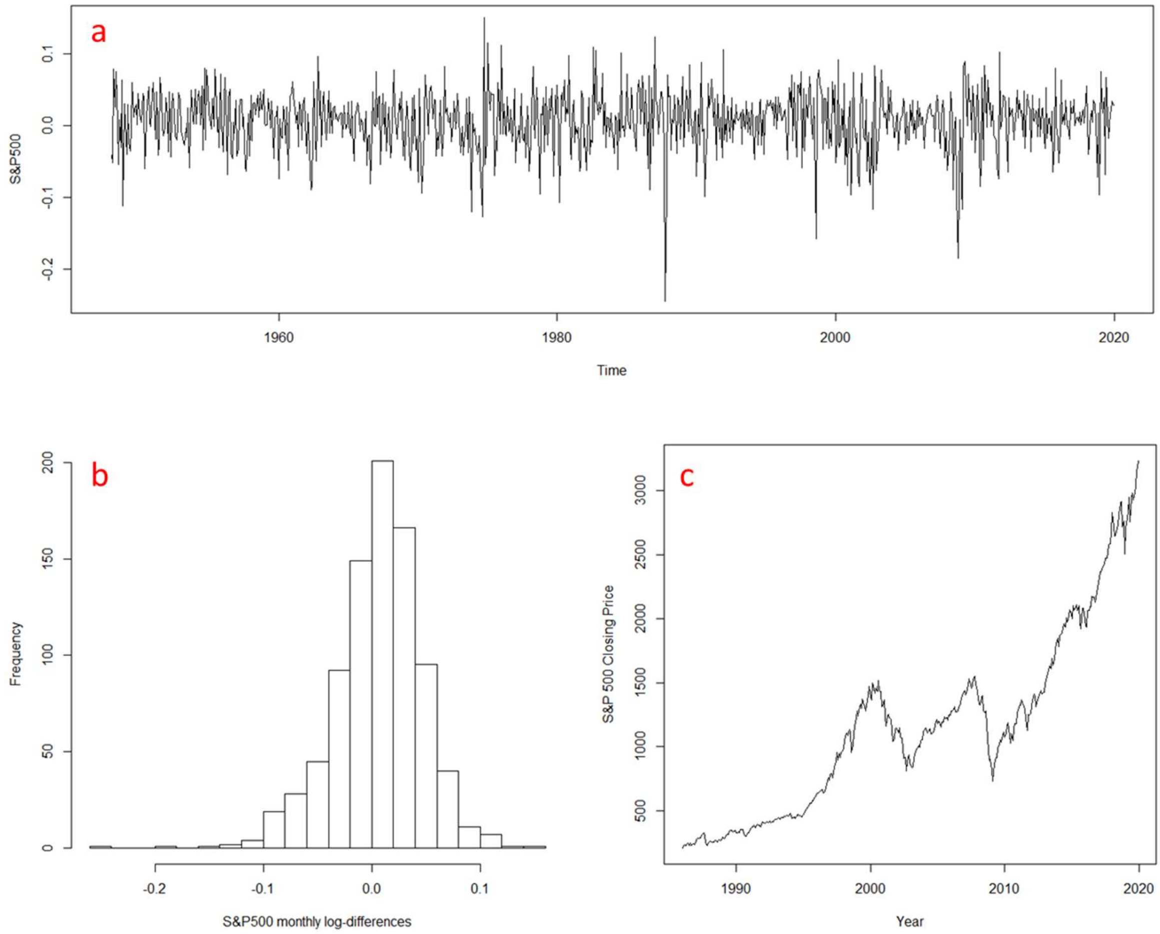

Figure 4.2 a) The log-returns of S&P 500 (1948-2019), b) Histogram of log-returns (1948-2019), c) Closing prices of S&P 500 (1986-2019)

The stock market series does not have a normal distribution, which is also clearly visible in Figure 6b. With skewness of -0,665 and kurtosis of 5,382, the distribution of returns is clearly leptokurtic (fat-tailed) and left-skewed (left-tailed). These measures, unsurprisingly, are in line with previous studies of the stock market (see e.g. Campbell et al. 1998, 19–20; Tsay 2005, 10–11).

Interpreting Figure 4.2a, it can be seen that the series is not stationary, because the variance (or volatility) is not constant, since it has different periods of high (e.g. 1930s) and low (e.g. 1990s) volatility, which is characteristic of stock markets and is also called volatility clustering (see e.g. Mandelbrot 1963; Ding et al. 1993, 84–85).

4.3 The credit spread

All the macroeconomic and credit market data used in the study is from the Federal Reserve Economic Database (FRED) maintained by the Federal Reserve Bank of St. Louis. The database and the time series, such as GDP, consumer price index and household wealth are considered to be reliable indicators of the factors they measure. FRED provides the monthly data for Moody's Seasoned Baa Corporate Bond Yield starting from 1919 (FRED 2019b) and the monthly data for US 10-year Treasury Constant Maturity Rate starting from 1962 (FRED 2019a), which are used to calculate the spread between the series, which is used as a proxy of the credit spreads in the US. The credit spread in question is also available from FRED on a daily basis starting from 1986, but the data used in this study is based on the monthly series.

The monthly series of credit spreads is based on the spread between the yields and transformed into a series of relative log-differences in order to make them more computable in an HMM setting. In other words, the series is transformed to a continuously-compounded-returns-like series, which has a more symmetric distribution even though not normal.

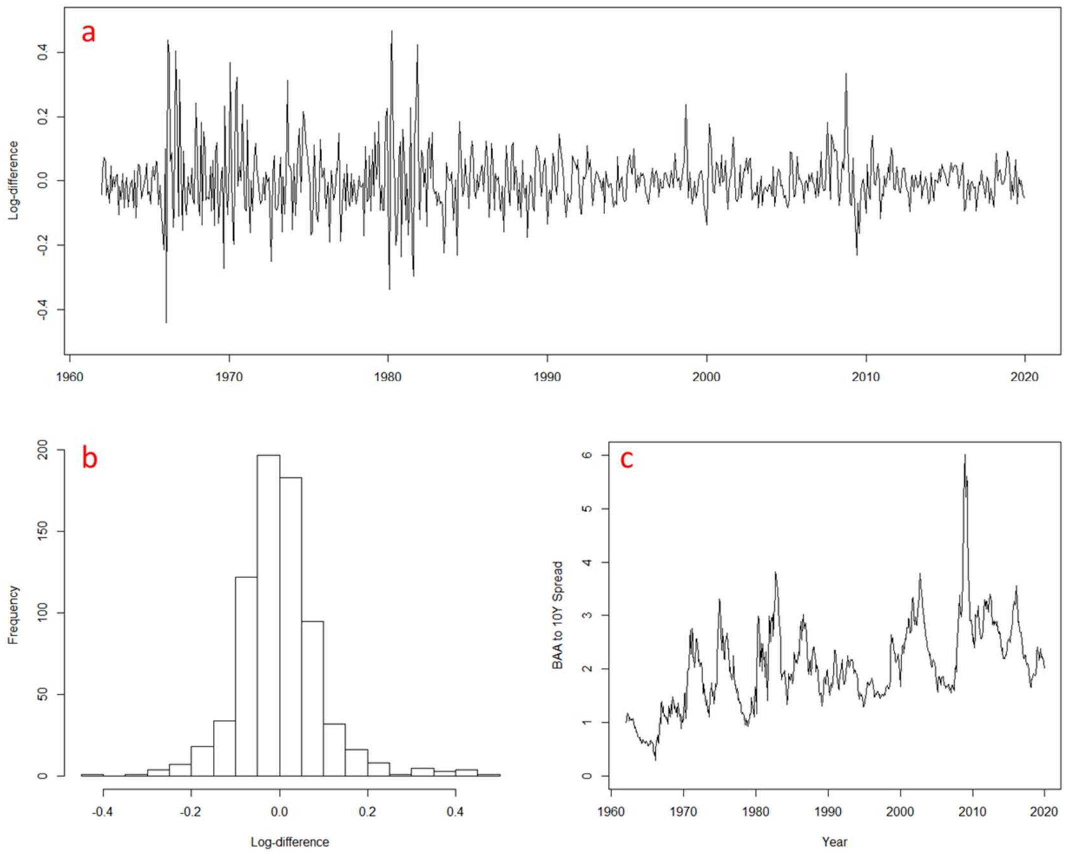

Figure 4.3 a) The log-differences of the credit spread (1962-2019), b) Histogram of log-differences (1962-2019), c) The BAA to 10Y Credit spread (1962-2019)

The credit spread series consists of 696 data points and does not have a normal distribution, which is also clearly visible in Figure 4.3b. With skewness of 0,805 and kurtosis of 7,706, the distribution of returns is clearly leptokurtic (fat-tailed) and right-skewed (right-tailed). The credit spread is usually relatively stable (Figure 4.3b) but can at times have large spikes and volatility (Figures 4.3a and 4.3c). Therefore, the figures would clearly indicate that the series is not stationary.

Table 4.1 presents descriptive statistics of both the time series used in HMM analysis, in other words the log-differences of the S&P 500 and the BAA to 10Y Spread.

Table 4.1 Descriptive statistics (BAA to 10Y spread & S&P 500)

S&P 500 1948-2019		*BAA to 10Y Spread 1962-2019*	
Mean	0,00620	Mean	0,00095
Standard Error	0,00141	Standard Error	0,00355
Median	0,00930	Median	-0,00264
Standard Deviation	0,04141	Standard Deviation	0,09371
Sample Variance	0,00171	Sample Variance	0,00878
Kurtosis	5,38243	Kurtosis	7,70560
Skewness	-0,66503	Skewness	0,80517
Range	0,39647	Range	0,90956
Minimum	-0,24543	Minimum	-0,44135
Maximum	0,15104	Maximum	0,46822
n	864	n	696

4.4 Additional time series

The Chicago Board Options Exchange (CBOE) Volatility Index (VIX) is based on the implied volatilities of the S&P 500 index options and is known as a reflective measure of the underlying state of the market and especially the general sentiment of the market. The Volatility Index, originally based on the trading of S&P 100 (OEX) options and trading of S&P 500 (SPX) options since 2003, data is available directly from the Chicago Board Options Exchange starting from 1986 for the original VXO index and 1990 for the VIX index (CBOE 2020). It is, therefore, only available for the analysis of the last few market crashes, which is naturally a bit problematic for the analysis, which is why the VIX is not used as intensively as the other measures.

The volatility index is seen as the most accurate measure of future volatility of the market and sometimes referred to as the fear factor or the fear gauge (see e.g. Whaley 2000). The volatility index provides important information on market sentiment, which could be used also in this study. The data used in the study is the highest points of the

VIX each month. The VXO and VIX series consists of 408 and 360 data points, respectively, and are shown in Figure 4.4.

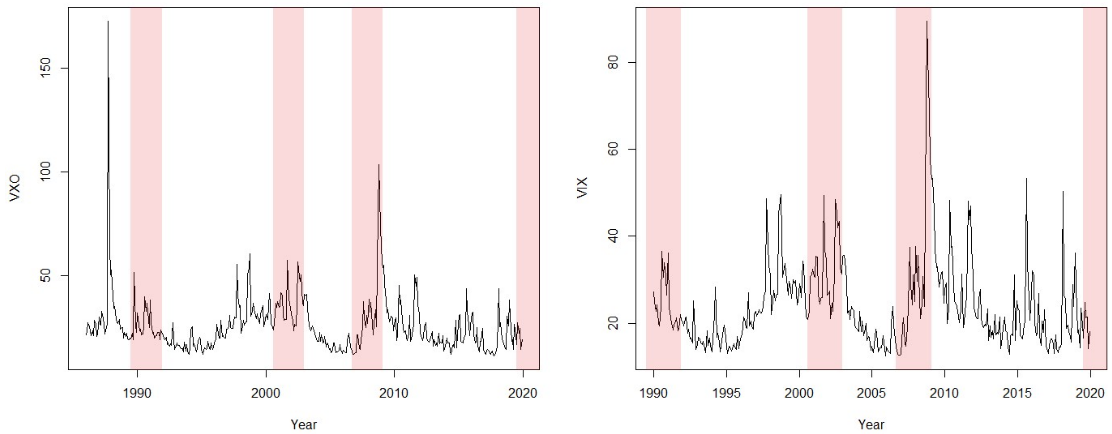

Figure 4.4 VXO (1986-2019) and VIX (1990-2019) with "danger zones" in red

Also, the cyclically adjusted total return price earnings ratio (TRCAPE) presented by Robert Shiller is considered in the study as a proxy for the relative valuation level of the US stock market. The monthly data (1871-2019) is available from Robert Shiller's website (Shiller 2020) and is considered to be reliable to use as such, even though some confirmatory calculations are made. The CAPE is "the real (inflation-corrected) S&P Composite Index divided by the ten-year moving average of real earnings on the index" (Shiller 2015, 6).

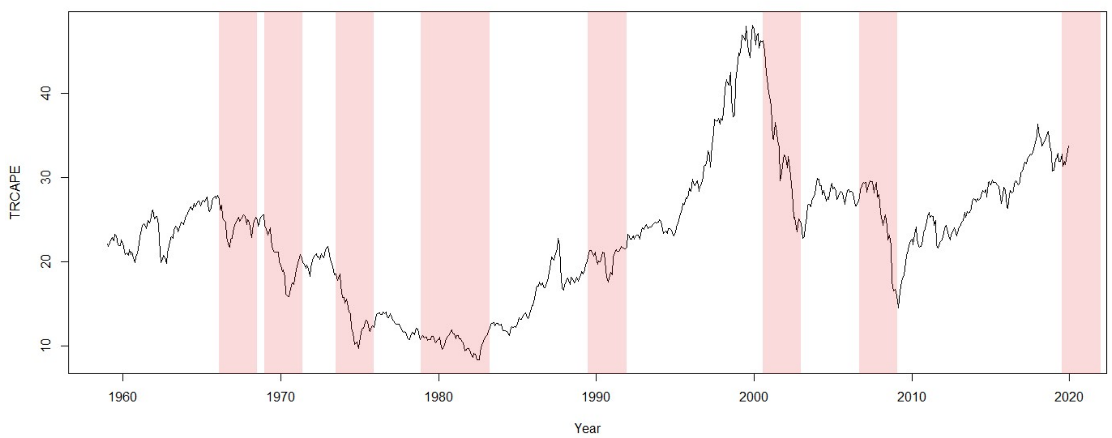

Figure 4.5 TRCAPE (1959-2019) with "danger zones" in red

TRCAPE is a newer relative valuation measure of the legendary CAPE developed by Robert Shiller and John Campbell, which is sometimes also referred to as the Shiller P/E. Whereas the traditional measure does not make any difference between dividends and share repurchases, which could have affected the CAPE, the TRCAPE reinvests dividends into the price index, which appropriately scales the earnings per share between various times. (Shiller 2020)

Various volume and turnover data were also considered, but there were no reliable measures available. For example, NYSE trade and turnover data was inconsistent with missing years in between (World Federation of Exchanges 2020) and S&P 500 volume as itself was considered too unreliable because of changes in investor preferences e.g. derivative products and the rise of ETFs have altered the role of the index itself compared to what it has been in the past. Volume and turnover-related measures are thus left for future research.

4.5 Data for empirical testing

The data, or the return series to be more precise, used in this study are the ones provided by Kenneth French (French 2020a). The empirical allocation strategies done in chapters 5 and 6 use this data as proxies for the stock market total return and the risk-free rate. The total stock market return (RM) includes dividends, which makes the strategies more realistic. The risk-free rate (RF) in turn is based on the 1-month US Treasury bill rate. (French 2020b.)

These data series have been chosen because of their longevity, simplicity and reliability (for methodology, see French 2020b). The market return and risk-free rate series are considered the best approximations of total returns to an average investor. These series are used in order to make the empirical testing of different allocation strategies as realistic as possible, even though it is almost impossible to capture all the true returns and costs over a very long period of time.

5 EMPIRICAL TESTING

What to do when the yield curve inverts? How can you try to time the market downturn and the volatile times ahead? There are several ways to tackle the issue, which is demonstrated in this chapter. The empirical tests are conducted with market return, risk-free rate and plain cash. The methods are considered from simple rules to more complex ones with statistical models.

The allocation strategies are done as simple allocations between the stock market and the risk-free rate, in other words the portfolios always consist of only one of the two assets (100% in stocks or bonds) in order to avoid excessive trading costs. The trading cost used in the empirical part is 0,5% per trade and taxes or management fees are not taken into account.

The different strategies' performances are compared to each other as the ability to grow the capital of $1 from the beginning of 1972 until the end of 1985. The pre-1972 data is not included in the performance analysis, because some data is needed to evaluate and train algorithms; the more complex methods would have a disadvantage compared to the simpler ones if the period in question would have been included in the analysis. The performances of the proposed strategies are displayed at the end of the chapter in Table 5.18.

5.1 Naïve allocation rules using the inverted yield curve

As stated in Sections 2.2 and 4.1, the immediate period (months 1-30) after the inversion of the yield curve the stock market is more volatile with statistically smaller returns when compared to other periods. This fact is used in a simple manner and tested with different time periods off the stock market, which are then combined in a dynamic way. In other words, the time after the inversion is considered too risky and is tried to be avoided by investing to the risk-free rate for a certain period of time.

First, different starting times to get off of the stock market are considered: right after the monthly signal (Month number 1), six months later (7) and a year later (13). Then, different times to get back to the stock market are considered: 6 months after the inversion (7), a year (13), 18 months after (19), two years (25) and the full 30 months after the inversion. Different strategy combinations are done within these borders: a total of 12 different simple strategies are introduced in Table 5.1.

Table 5.1 The different naive strategies and their months off the stock market

		Finish				
		6	12	18	24	30
Start	1	1 - 6	1 - 12	1 - 18	1 - 24	1 - 30
	7	-	7 -12	7 - 18	7 - 24	7 - 30
	13	-	-	13 - 18	13 - 24	13 - 30

The periods of staying out of the stock market ranges from six months to two and half years. The interval of 6 months is considered a good compromise between overfitting and being too approximate; 3 months seems too precise and not comparable between different cases, whereas 12 months seems not precise enough to really capture the effects of the phenomenon.

These naïve allocation strategies are tested against each other and the stock market for every incident of the inverted yield curve with except of the two inversions in 1978 and 1980 with just 23 months apart which are handled as a single volatile period after the inversion (see also Figure 4.1). The best performing strategies (months 1-30 after the inversion) for every occasion of the inverted yield curve are presented in Table 5.2.

Table 5.2 Best performing strategies for inverted yield curves

Year of the inversion	Best simple strategy
1966	1-6
1968	1-18
1973	1-18
1978	7-18
1989	7-18
2000	1-30
2006	13-30

It should be noted that there were many simple strategies in every occasion outperforming the stock market, but Table 5.2 shows only the best performing strategies. The presented strategies are, however, constructed and measured after-the-fact, which is why they aren't feasible trading strategies as such.

In order to benefit from the simple strategies, we must build more realistic and dynamic allocation strategies. This is done in two simple ways: the first allocation strategy (1^{st} strategy) relies on the best performing simple strategy considering the whole data set until the end of the last inversion period, whereas the second strategy (2^{nd} strategy) relies on the most efficient strategy considering only the latest inversion period. These simple dynamic strategies presenting the periods of investing in the risk-free rate are shown in Table 5.3.

Table 5.3 Dynamic inverted yield curve strategies

Inversion of the yield curve	*1. The Curve (whole data set)*	*2. The Curve (latest inversion)*
1966:1	N/A	N/A
1968:12	1-6 (1969:1 - 1969:6)	1-6 (1969:1 - 1969:6)
1973:6	1-18 (1973:7 - 1974:12)	1-18 (1973:7 - 1974:12)
1978:11 (1980:10)	1-18 (1978:12 - 1980:5; 1980:11 - 1982:4)	1-18 (1978:12 - 1980:5; 1980:11 - 1982:4)
1989:6	1-18 (1989:7 - 1990:12)	7-18 (1990:1 - 1990:12)
2000:7	1-18 (2000:8 - 2002:1)	7-18 (2001:2 - 2002:1)
2006:8	1-18 (2006:9 - 2008:2)	1-30 (2006:9 - 2009:2)
2019:6	1-18 (2019:7 - 2020:12)	13-30 (2020:7 - 2021:12)

The returns from the proposed simple dynamic strategies in the testing period 1972–1985 are fairly good in comparison to the stock market. Whereas the stock market had decent annualized returns of 10,57%, the strategies (both have the same allocations until 1989) fared much better with annualized returns of 14,71%. Furthermore, to return to the original performance comparison, a dollar invested in the beginning of 1972 had very different values in the end of 1985: $4,08 for the benchmark stock market and $6,83 for the strategies. The performances of the strategies are shown in Table 5.18 at the end of the chapter.

It should be noted, however, that the late 70s and the early 80s had very high US Treasury bill returns which had no precedents and which haven't occurred afterwards. These unusually high returns caused by the strong inflation made the stock market less attractive, which made it easier for the strategies to stay out of the stock market.

5.2 Credit spreads and the inverted yield curve

As stated before, the spread between US Treasury and BAA corporate yields reflect the credit conditions in the USA, in other words the credit spreads are a general look on the creditworthiness of the companies on the lowest level of investment grade corporate bonds. The spread in question can be seen as a proxy of the overall state of the market: if the spreads are low, the economy is considered to be in a good and stable state, whereas high spreads indicate suspicion among market participants and worse confidence on the economy in general.

In this section the indications of the yield spread and its relative changes are used in combination with the yield curve to build allocation strategies which aim to stay out of the stock market during the stormy conditions. In more detail, the signal given by the 4-state HMM applied to the log-changes in the credit spread is combined with the inversion of the yield curve: when the yield curve inverts, the first signal of a negative market regime (state of strong widening of the spread) results in getting off of the stock market and investing in the risk-free rate. The different credit spread regimes, their expected means and standard deviations are presented in Table 5.4.

Table 5.4 Means and standard deviations of the states of the model (Data used 1962:1 – 1971:12)

	State 1	***State 2***	***State 3***	***State 4***
Mean	-0,03439	-0,02778	0,03717	0,08398
Sd	0,08721	0,04313	0,02497	0,20278

The different hidden states have their own characteristics that can be seen in their observations: whereas states 1–3 are relatively stable, state 4 with its significantly larger mean and standard deviation in observations represent a different time in the credit market. If these observations are thought to be generated by a hidden process which reflects the general sentiment of the market, state 4 can be seen as the unstable period of distrust, lack of confidence and – at worst – flight-to-liquidity which widens the spreads between safe and riskier assets very rapidly. The allocation strategies in this section are based on these assumptions: when the unstable state is probable, it is time to get off the stock market and invest in the risk-free rate.

As stated in Section 3.5, the HMMs used in this study are updated monthly as if done in real time. The credit spread HMM transition probabilities, parameters and both local and global decodes are evaluated at the end of each month to assess the possible signals immediately with the first data set being 1962:1 – 1968:1 and the last 1962:1 – 2019:12. The chosen HMMs seem to be able to model the hidden layers of the phenomenon quite well as the partial autocorrelations of the residuals are not significant. Also, the theoretical quantiles would seem to indicate that the residuals follow the Gaussian distribution more closely than the credit spread series.

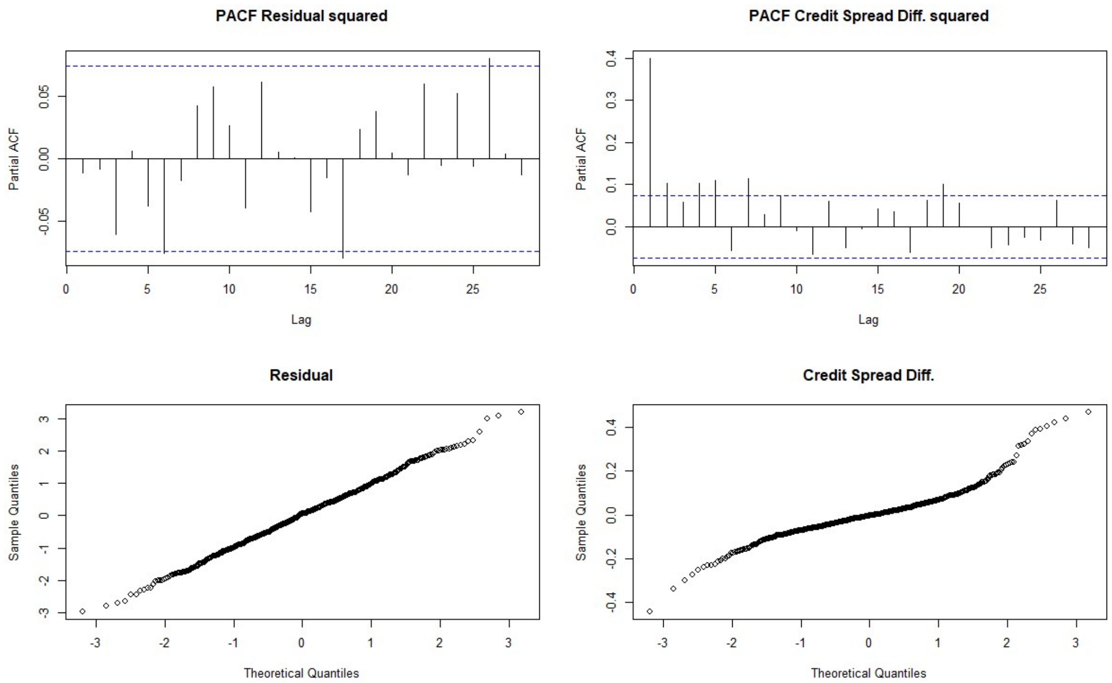

Figure 5.1 Credit Spread HMM residual analysis (Data 1962-2019)

The starting point, the initial distribution ***δ*, **the initial transition probability matrix **Γ,** and initial estimations for means and standard deviations will remain the same for every update of the model, only the last data point of the time series will change and thus new parameters and probabilities are calculated for the whole process. Because the means, standard deviations and state probabilities are not stable (they have different values at different times), the state with the highest mean (state 4 most of the time, sometimes state 3) is tried to be avoided and the standard deviation is not regarded as important as the mean.

Table 5.5 The initial distribution, initial estimated means, initial standard deviations and initial transition probabilities used in the HMM (Credit spread)

	State 1	State 2	State 3	State 4
δ	0,25	0,25	0,25	0,25
μ	-0,10	0,00	0,05	0,10
σ	0,10	0,05	0,05	0,10

	State 1	State 2	State 3	State 4
State 1	0,85	0,05	0,05	0,05
State 2	0,05	0,85	0,05	0,05
State 3	0,05	0,05	0,85	0,05
State 4	0,05	0,05	0,05	0,85

The inverted yield curve (months 1-30 after the inversion) and the signal provided by the credit spread are used to formulate simple allocation rules. We use both local and global decoding in assessing whether or not to turn to the risk-free rate:

- if the state with the highest mean is the most probable (Prob.>0,5) at time t, it is considered a signal to start investing in the risk-free rate in $t+1$ (local decoding)
- if the most likely sequence (using Viterbi algorithm) ends in the state with the highest mean at time t, we start to invest in the risk-free rate in $t+1$ (global decoding).

Table 5.6 Signals provided by the inverted yield curve and the credit spread HMM

Inversion of the yield curve	*Local decoding*	*Global decoding (Viterbi algorithm)*
1966:1	N/A	N/A
1968:12	1969:10	1969:10
1973:6	1973:9	1973:9
1978:11	1980:2; 1980:12	1980:2; 1981:6
1989:6	1989:6	1989:6
2000:7	2001:9	2001:9
2006:8	2007:8	2007:8
2019:6	?	?

The differences between the signals provided by the two methods are very similar with the only difference being the later part of the 1978 inversion. The Viterbi algorithm does not provide any real insight or more accurate signals in this case, which isn't really surprising considering its designated use in assessing the most likely sequence, which was not really used in this setting. Given these findings, we find it more credible to use local decoding in the forecasting period.

The credit spread based HMM approach is based on the dynamics of the credit spread and the hidden states of market sentiment behind it; the surges in the spread usually indicate lack of confidence within financial markets (see e.g. Greenwood & Hanson 2013; López et al. 2016) representing a worse market state for the average investor. Also, we view these changes in the bond market sentiment as relevant signals to the stock market because of the interconnectedness of the bond and stock markets. These worsening conditions to invest in riskier assets are tried to be avoided by investing in the risk-free rate for a certain period of time.

As in Section 5.1, different time periods are tested to evaluate the best time to get back to the stock market after the negative market state signal: 9, 12, 15 and 18 months after the signal are tested. Once again, different periods are compared to each other and the best simple strategies for each inversion are presented in Table 5.7.

Table 5.7 Best performing credit spread strategies for each inversion

Year of the inversion	*Best simple credit spread strategy*
1966	N/A
1968	9 months
1973	12 months
1978	18 months (Local)
1989	15 months
2000	12 months
2006	18 months

Once again, in order benefit from the simple credit spread strategies, we must build more realistic and dynamic allocation strategies. As in Section 5.1, this is done in two simple ways: the first allocation strategy (3rd strategy) relies on the best performing simple credit spread strategy considering the whole data set until the end of the last inversion period, whereas the second allocation strategy (4th strategy) relies on the most efficient strategy considering only the latest inversion period. These dynamic strategies based on credit spread signals (Appendix 3) with their respective periods of investing in the risk-free rate are shown in Table 5.8. As for the latest inversion in 2019:6, there were no signals in the data and the signal is assumed to surface in 2020 or 2021.

Table 5.8 Credit spread HMM based dynamic simple allocation strategies

Inversion of the yield curve	*3. The Spread (whole data set)*	*4. The Spread (latest inversion)*
1966:1	N/A	N/A
1968:12	N/A	N/A
1973:6	9 months (1973:9 - 1974:5)	9 months (1973:9 - 1974:5)
1978:11 (1980:10)	12 months (1980:2 - 1981:11)	12 months (1980:2 - 1981:11)
1989:6	12 months (1989:6 - 1990:5	18 months (1989:6 - 1990:11)
2000:7	15 months (2001:9 - 2002:11)	15 months (2001:9 - 2002:11)
2006:8	15 months (2007:8 - 2008:10)	12 months (2007:8 - 2008:7)
2019:6	15 months (?)	18 months (?)

The returns from the proposed strategies do beat the market in the testing period of 1972–1985 but fail to match the returns of the strategies proposed in Section 5.1. Whereas the benchmark stock market had annualized returns of 10,57%, the strategies (both have the same allocations until 1989) fared much better with annualized returns of 12,92%. To use the original performance comparison, a dollar invested in the beginning of 1972 had very different values in the end of 1985: $4,08 for the benchmark stock market and $5,48 for the strategies. The performances of the strategies are compared in Table 5.18 at the end of the chapter.

The allocation strategies presented in this section are clearly outperforming the stock market, especially if the lower volatility is taken into account. The relevant question is, however, are the returns based more on the inverted yield curve than the signals given by the changes in the credit spread? The question is rather hard to answer, but the credit spread HMM based strategies provide an efficient alternative to simple yield curve inversion strategies.

5.3 S&P 500 and the inverted yield curve

As stated before, there are many applications of time series analysis and forecasting on the stock market and also the S&P 500 in particular. Even though it is a common topic in modeling and forecasting, it is rare to combine these forecasts with the inverted yield curve. This section aims to bring together the modeling of the S&P 500 with an HMM and the inversion of the yield curve to construct useful allocation strategies.

Hidden Markov models are usually used to estimate different market regimes, which all have their own qualities. This study uses a gaussian mixture HMM, which tries to understand the hidden states of the stock market, which are usually described as growth (bull) market, neutral market and negative (bear) market.

Different number of states were tested in order to see which model captures best the market dynamics and after extensive testing the 4-state model seems to be the best to estimate the relevant hidden states. The proposed allocation strategies are based on the simple rule of trying to avoid the negative (bear market) states after the yield curve has inverted. Different states, their expected means and standard deviations for the 4-state model are presented in Table 5.9.

Table 5.9 a) Means and standard deviations of the states of the model b) Transition probabilities between states (Data used 1948:1 – 1971:12)

a	State 1	State 2	State 3	State 4
Mean	-0,06657	0,00152	0,02112	0,07272
Sd	0,01939	0,03653	0,02075	0,00708

b	State 1	State 2	State 3	State 4
State 1	0,60163	0	0	0,39875
State 2	0,01075	0,8979	0,07181	0,019532
State 3	0,00	0,22	0,78	0,00
State 4	0,00	0,57	0,43	0,00

Just as in the previous section, the different hidden states have their own characteristics that can be seen in their observations: State 1 represents the bear market, when the expected returns are negative; state 2 is the relatively neutral state; state 3 is characterized by moderate growth, whereas state 4 is the clear bull market with rapid growth.

If these observations are thought to be generated by a hidden process reflecting the general sentiment of the market, state 1 can be seen as the period of almost outright panic: the losses are high and there are virtually no chance of a profit because of the small standard deviation of the state. Also state 2 can be viewed as avoidable, but the allocation rules used in this section are based solely on negative means and the inverted yield curve. Because the means, standard deviations and state probabilities are not stable (they have different values at different times), we propose the following allocation rules:

- As all the states are relatively persistent (except for state 4), if the weighted average of expected means of the states is negative at time t and the yield curve has inverted (months 1-30 [signals 0-29] after inversion), we invest in the risk-free rate in $t+1$ (local decoding)
- As all the states are relatively persistent (except for state 4), if the most likely sequence (using Viterbi algorithm) ends in a state that has a negative mean at time t and the yield curve has inverted (months 1-30 [signals 0-29] after inversion), we invest in the risk-free rate in $t+1$ (global decoding)

As stated before in Sections 3.5 and 5.2, the HMMs used in this study are updated monthly as if done in real time. The S&P500 HMM transition probabilities, parameters and both local and global decodes are evaluated at the end of each month to assess the possible signals immediately with the first data set being 1948:1 – 1959:1 and the last 1948:1 – 2019:12. Just as with the credit spread HMM, the S&P HMMs seem to be able to model the hidden layers of the phenomenon quite well as the partial autocorrelations of the residuals are insignificant. Also, the theoretical quantiles would seem to indicate that the residuals are relatively Gaussian.

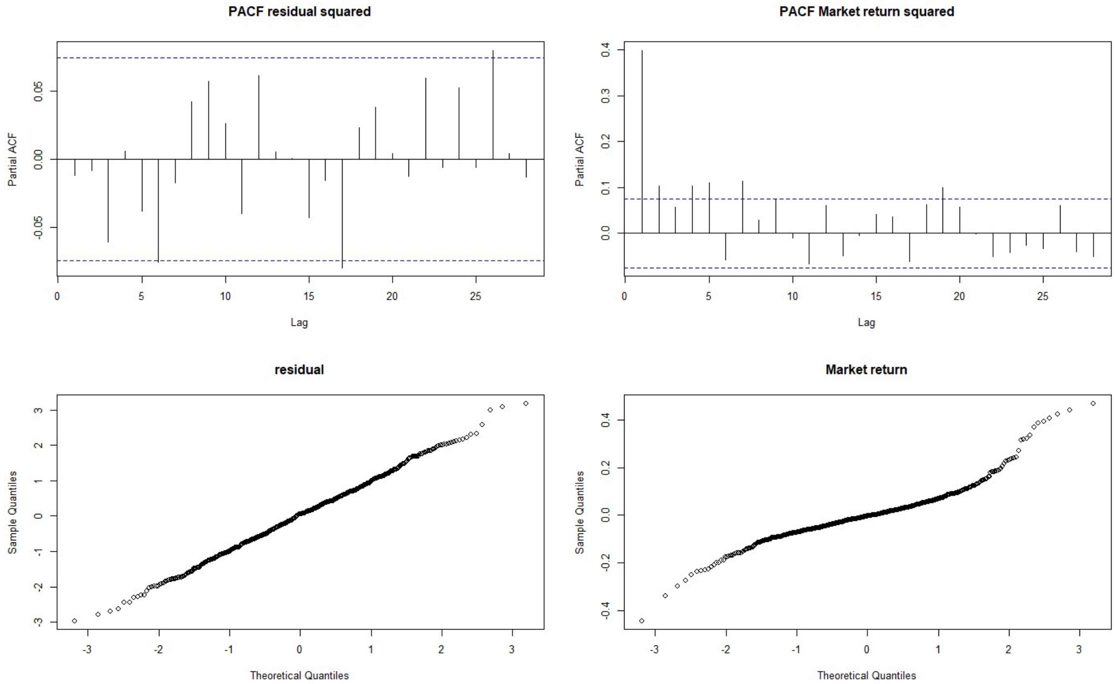

Figure 5.2 S&P HMM residual analysis (Data 1948-2019)

The starting point, the initial distribution $\boldsymbol{\delta}$**,** the initial transition probability matrix $\boldsymbol{\Gamma}$**,** and initial estimations for means and standard deviations will remain the same for every update of the model, only the last data point of the time series will change and thus new parameters and probabilities are calculated for the whole process.

Table 5.10 The initial distribution, initial estimated means, initial standard deviations and initial transition probabilities used in the HMM (S&P 500)

	State 1	State 2	State 3	State 4
δ	0,25	0,25	0,25	0,25
μ	-0,10	0,00	0,05	0,10
σ	0,10	0,05	0,05	0,10

	State 1	State 2	State 3	State 4
State 1	0,85	0,05	0,05	0,05
State 2	0,05	0,85	0,05	0,05
State 3	0,05	0,05	0,85	0,05
State 4	0,05	0,05	0,05	0,85

The allocation strategies proposed in this section relate to the two allocation rules presented above: the first strategy (5^{th} strategy) relies on local decoding and expected means, while the second strategy (6^{th} strategy) is based on global decoding and most likely sequences. The signals provided by the HMMs in 5^{th} and 6^{th} strategies are presented in Appendices 1 and 2, respectively, because of the vast number of signals provided by the both strategies.

Whereas the first four strategies using the inverted yield curve and the credit spread were updated basically at the end of each inverted yield curve period, the strategies constructed in this section are updated monthly according to signals given by the models. This is why it is quite obvious that these models trade a lot more, which could have a significant effect on the performance of the models, especially if higher transaction costs are assumed.

The returns from the proposed S&P 500 related strategies beat the market in the testing period of 1972–1985, but do not reach the returns of the earlier proposed strategies. Whereas the benchmark stock market had annualized returns of 10,57%, the 5th strategy fared slightly better with annualized returns of 10,81% and the 6th strategy also provided annualized returns of 10,86%. To use the original performance comparison, a dollar invested in the beginning of 1972 had the following values in the end of 1985: $4,08 for the benchmark stock market and $4,21 for the 5th strategy and $4,23 for the 6th strategy. The performances of the strategies are compared in Table 5.18 at the end of the chapter.

The 5th and 6th strategies itself and the returns of the strategies are fairly similar even though there are some differences. For example, the 5th strategy trades more which makes it more sensitive to changes in transaction costs which is discussed further in Chapter 6. Altogether the strategies constructed in this section does not seem as solid as the first four, but they do seem to create some value, which is why they are included in further analysis.

Once again, the relevant question is whether the HMM provides any true value at all or should the improved returns be credited only for the yield curve? To evaluate the question, we compare the returns of the strategies to the strategy of just staying out of the stock market for the period after the inversion (Months 1-30).

Table 5.11 The S&P 500 HMM vs. Avoiding danger zones

	5. The S&P 500 (weighted	*6. The S&P 500 (Viterbi)*	*Simple comparison (Risk-free rate 1-30)*
$1 (1.1.1972 - 31.12.1985)	$4,21	$4,48	$3,34
Annualized returns (%)	10,81 %	11,29 %	9,01 %

Most likely the models used in this section do carry some weight if compared to the simplest avoidance strategy of them all, but if we compare the 5th and 6th strategies to the first two, the yield curve is more substantial in avoiding volatile periods. However, if the two approaches are combined, there could be more room for improvement.

5.4 NY Fed prediction model and the inverted curve

One possible option to construct simple allocation strategies is to use the recession probability model of the Federal Reserve of New York, which is described in detail in section 2.3. If the yield curve itself – as it seems to be looking back at the earlier allocation strategies – has a lot predicting power in future bear markets, could it be that the recession probability model itself can be used in allocation strategies?

As stated in Section 2.3, the flat yield curve (Spread = 0,00) produces the probability of a recession of approximately 28-30% in 12 months. The model is designed to predict NBER defined recession, but this section evaluates, whether the recession probability model has the same indications for the stock market. As this study in general argues that the inversion itself is more important than the level of the negative spread, the following simple allocation rule is suggested:

- If the probability of a recession is more than 30% in t, we invest in the risk-free rate in t (the probability is known 12 months ahead so it is possible to invest at the same time)

It could be noted also that these allocation rules can be made with the yield curve itself; there is no true need for the NY Fed recession model since the allocation rules presented in this section are fundamentally just based on negative spreads of the yield curve.

Also, different options are studied regarding the lead time of the model. The NY Fed recession model excepts the lead time to be 12 months from the measurement, but it is expected in this study that the stock market reacts negatively sooner than the economy itself goes into recession. For that reason also shorter lead times are tested: 1, 3, 6, 9 and 12 months from the measurement are considered and their performances tested against each other. The best performing lead times for each inversion are presented in Table 5.12.

Table 5.12 Best performing lead times of the NY Fed recession model for each inversion

Year of the inversion	*Best lead time*
1966	-
1968	3 months
1973	3 months
1978	9 months
1989	9 months
2000	3 months
2006	9 months

The best performing lead times are used to construct dynamic allocation strategies trying to benefit from these observations. Two different strategies are proposed just as in previous sections: the first allocation strategy (7th strategy) relies on the recession probability using the best performing lead time considering the whole data set until the end of the last inversion period, whereas the second strategy (8th strategy) relies on the recession probability using the most efficient lead time considering only the latest inversion period.

Table 5.13 Dynamic lead times of the NY Fed recession probability model for each inversion

Inversion of the yield curve	*7. The NY Fed model (whole data set)*	*8. The NY Fed model (latest inversion)*
1966:1	N/A	N/A
1968:12	N/A	N/A
1973:6	6 months	3 months
1978:11 (1980:10)	3 months	3 months
1989:6	3 months	9 months
2000:7	3 months	9 months
2006:8	3 months	3 months
2019:6	3 months	9 months

The above proposed lead times are applied to the probabilities issued by the NY Fed recession probability model, in other words the regular lead time of 12 months is changed to the figures proposed in Table 5.13. This method gives us the periods in which to invest in the risk-free rate. The resulting periods out of the stock market are not presented here because of the multiple trades done in each inversion, but the periods can be easily recalculated with the NY Fed recession probability data.

The returns from the proposed recession probability model related strategies beat the market in the testing period of 1972–1985 but vary quite much with each other which could be an indicator of the randomness of the phenomenon. Whereas the benchmark stock market had annualized returns of 10,57%, the 8th strategy clearly outperformed the benchmark with annualized returns of 14,61%. The 7th strategy, on the other hand, also provided solid annualized returns of 12,00%, but failed to reach the profits of the 8th strategy.

To use the original performance comparison, a dollar invested in the beginning of 1972 had the following values in the end of 1985: $4,08 for the benchmark stock market and $4,89 for the 7th strategy and $6,75 for the 8th strategy. The performances of the strategies are compared in Table 5.18 at the end of the chapter.

Once again, the reliability of the strategies could be questioned because of the rather large differences between the strategies. Is the good performance of the strategies because of their ability to interpret causalities or is it just good luck? And furthermore, is it possible to time this kind of a complex phenomenon in such a simple way? Obviously, there are no certain answers, but these issues are addressed by presenting all returns for each simple strategy for each inversion in the pre-1986 testing period in Table 5.14, where all total returns (months -1 to 30, longer period in 1978–1983) below the stock market are highlighted in red.

Table 5.14 Total returns (Months -1 to 30) for all lead times for each inversion pre-1986

		Lead time					Bench-
		1	3	6	9	12	mark
Inversion	1966:1	0,87 %	3,36 %	10,31 %	7,68 %	15,32 %	23,90 %
	1968:12	12,74 %	30,45 %	23,89 %	15,34 %	-9,36 %	1,55 %
	1973:6	22,21 %	56,67 %	13,44 %	-1,96 %	-15,93 %	-7,61 %
	1978:11 (1980:10)	108,29 %	109,66 %	124,59 %	130,38 %	70,00 %	115,14 %

If we ignore the longest lead time, which implies that the stock market and the economy in general would react to the inversion of the yield curve at the same pace, the strategies using shorter lead times fare moderately well with 62,5% (10/16) of cases beating the benchmark, but the mixed results in general would indicate that the effect is not predictable, even though not totally random.

All in all, the NY Fed recession probability model and the yield spread itself do seem to have some predictive power over future bear markets, which could be used in allocation strategies. Thus, these strategies could be also combined with other indicators and strategies.

5.5 Other indicators and the inverted yield curve

The inverted yield curve has proven to be a reliable predictor of future NBER-denominated recessions and so far in this study also a noteworthy predictor of future bear markets. Further applications are considered in this section using other indicators combined with the yield curve, notably the cyclically adjusted total return price earnings ratio (TRCAPE). As the volatility index (VIX/VXO) does not have any history before 1986, it cannot be used in this chapter.

As stated before, the TRCAPE is a measure of relative valuation of the stock market dividing the total return price with cyclically adjusted earnings of the last 10 years taking into account both dividends and share buybacks. For this reason, the TRCAPE has less fluctuations than the ordinary P/E ratio and can therefore be considered more useful in a long time series setting. The TRCAPE is combined with the inverted yield curve in this section by considering boundaries to relative value changes within the inversion period.

In order to measure the relative changes in the TRCAPE we propose the following indicator to assess the current value in comparison to the highest value of the last 5 years (60 months):

$$TRC5Y_t = \frac{TRCAPE_t}{\max(TRCAPE_{t-60}, \dots, TRCAPE_{t-1})} \tag{5.1}$$

We aim to find reasonable relative value levels in comparison with the highest value of the last 5 years with the new indicator. In other words, we try to evaluate whether stock market valuations are already at reasonable levels or if there is more room for decline. It could be possible to do the same with the relative changes of the S&P 500, but the smoothed relative valuations can be considered to be more comparable across time (see e.g. Shiller 2015, 6; Siegel 2016).

We propose very simple strategies regarding the TRCAPE and the inverted yield curve. First, we look at the rate of $TRC5Y_t$ when the yield curve inverts (Month 0), then:

- If the rate is over 90%, we invest in the risk-free rate until the rate has gone down 20 percentage points or the inverted period is over (more than 30 months from the inversion)
- If the rate is over 80%, we invest in the risk-free rate until the rate has gone down 10 percentage points or the inverted period is over (more than 30 months from the inversion)
- If the rate is over 70%, we invest in the risk-free rate until the rate has gone down 5 percentage points or the inverted period is over (more than 30 months from the inversion)
- If the rate is below 70%, we continue to invest in the stock market

The strategy is based on elevated relative valuations and on prudence when trying to benefit from the exuberant valuations and market corrections: The proposed strategy fails to time the bear market perfectly, but it manages to hedge the portfolio against it to some extent. The periods of staying out of the stock market as identified by the strategy are presented below.

Table 5.15 The TRCAPE and the inverted yield curve strategy

Inversion of the yield curve	*9. The TRCAPE*
1966:1	1966:2 - 1966:10
1968:12	1969:1 - 1970:2
1973:6	1973:7 - 1973:12
1978:11 (1980:10)	1980:11 - 1981:11
1989:6	1989:7 - 1990:11
2000:7	2000:8 - 2001:4
2006:8	2006:9 - 2008:11
2019:6	2019:7 - ?

The returns from this very simple strategy are indeed interesting: The strategy beats the benchmark in the testing period of 1972–1985 by a clear margin even though not by a landslide. Whereas the benchmark stock market had annualized returns of 10,57%, the TRCAPE and inverted yield curve based strategy (9th strategy) fared better with annualized returns of 10,81%. In the terms of the original performance comparison, a dollar invested in the beginning of 1972 had the following values in the end of 1985: $4,08 for the benchmark stock market and $4,21 for the 9th strategy. As before, the performances of the strategies are shown in Table 5.18 at the end of the chapter.

The strategy introduced in this section is based on the assumption that a bear market always follows the inversion of the yield curve. If this was not to happen after the inversion, the proposed strategy would probably lose rather clearly to the stock market as the strategy would invest in the risk-free rate for two and a half years for nothing basically. In this sense the TRCAPE based strategy is riskier than the earlier ones because of the downside risk of missing a lot of potential returns caused by the long period out of the stock market.

On the other hand, all the strategies presented in this study are based on the assumption that history repeats itself and so far the evidence points at the direction that the inverted yield curve really does precede periods of worse returns in the stock market. In this sense the proposed TRCAPE based strategy is just as reasonable as the other strategies.

5.6 The comprehensive bear market predictor (*The BEAR*)

So far, this study has introduced nine different strategies using the inverted yield curve and other indicators to try to avoid the periods of volatility and losses in the stock market. Could it be possible to combine different indicators and signals to create a new, comprehensive bear market predictor to further boost the returns of the previous strategies? This section is dedicated to a symbiosis which tries to bring together all the observations and present the probability of a bear market in the next period as a one combined percentage rate.

In order to combine the probabilities of different strategies, each used indicator is scaled to a percentage rate between 0% and 100%. First off, the rescaled 2nd strategy is transformed by marking the best performing yield curve periods presented in Table 5.3 as 100%. To take into account the differences between inversions, a smoothed tail for each inversion is presented: after the 100% period, three months of 50% probability and three months of 25% probability are proposed. Furthermore, if the start of the period is not the inversion itself, the same smoothing is proposed also up front. Also, the probabilities are moved from t to $t-1$ to have the allocation strategies in the right period (e.g. if the strategy is 7–18, the 100% probabilities should be in months 6-17). The inversion of 2000 is presented as a clarifying example:

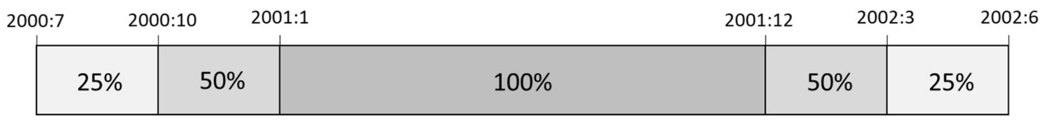

Figure 5.3 Scaling of the 2nd strategy

With these modifications *The Curve* (2nd strategy) is presented in a quantitative and comparable way and thus it can be a part of the comprehensive predictor.

Second, the credit spread signals are transformed to percentage rates in a similar manner as the inverted yield curve. The periods out of the stock market indicated by the HMM based signals are presented in Table 5.8. The periods in question are indicated by 100% probability and the end of the period is smoothed as before: after the 100% period, three months of 50% and 25% probability are proposed. The credit spread signal of 1973:9 is presented as an example:

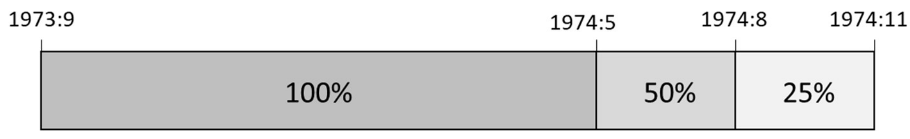

Figure 5.4 Scaling of the 4th strategy

The rescaled 4th strategy, *The Spread*, is the second element in constructing the comprehensive bear market predictor. As it is scaled just as the first element, both elements can be compared against each other and thus it is possible to optimize their weights when pursuing the ultimate strategy.

Third element of the symbiosis are the signals provided by the S&P 500 hidden Markov model. In order to scale the strategy as the first two, the weighted average of expected means is transformed to a scale of 1–100% using the means and standard deviations of the S&P 500 log-return series:

- The lower border, equal to 100% probability, is one and a half standard deviations below the mean of the series
- The higher border, equal to 0% probability, is the mean of the series

For the first 20 years of the data set (1948-1967), the means and standard deviations are calculated the whole time series so far, whereas starting from 1968 the means and standard deviations are calculated using the last 20 years of data (latest 240 data points). The left-side border is chosen to be 1,5 standard deviations of the log-return series, because of the leptokurtic distribution shown in Figure 4.1a. The initial thought was to use 2 standard deviations, but probabilities over 70% were so rare that it was deemed not to be representative of the true probabilities.

The scaled probabilities are also with the chosen borders most of the time 0,00%, but the signals reflect more their predictive value as shown in Figure 5.5, which allows the indicator to have more differentiated values. With these measures the 5th strategy, *The S&P 500*, is rescaled and it can be combined with the other indicators.

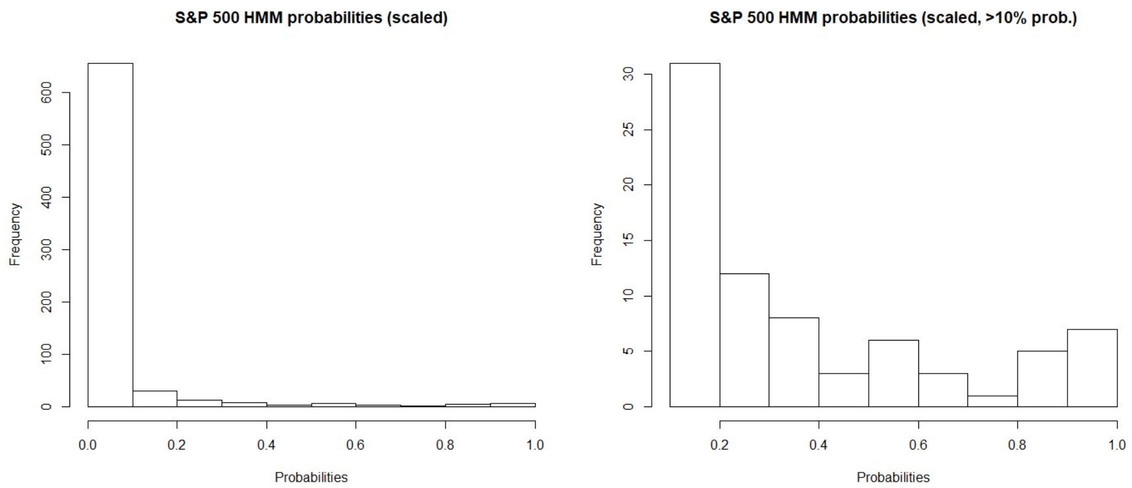

Figure 5.5 The scaled S&P 500 HMM probabilities

The fourth ingredient is the NY Fed recession probability model, which already gives percentage rates for a NBER recessions and which was applied to forecasting bear markets in Section 5.4. The indicator is simply rescaled by choosing 30% as the upper border (probability of 100%) and keeping the lower border of 0% the same. The lead times used are those of the 8th strategy, *The Fed*, which are based on the latest inversion only.

The TRCAPE based strategy is not used as such, but it is included in the comprehensive predictor as a possible backstop, if the relative valuation is low when compared to the highest value of the last 5 years. In order to do this, a simple addition to the measures presented above in this section is proposed. If $TRC5Y_t$ falls below 50% during the inversion period, the probability of a bear market is reduced by 30% for the rest of the inverted period. The end of the inverted period is marked with T_{inv} that equals the 30th month after the latest inversion of the yield curve.

The components of the comprehensive bear market predictor are now introduced, but the predictor still needs appropriate weights for all the factors. In order to have a truly comprehensive bear market predictor, each factor must have some weight (>0%), but none of them are allowed to be too dominant (>50%). Also, further weight limits are proposed to take advantage of previous studies and the earlier strategies:

- For *The Curve* and *The Spread*: $20\% \leq w \leq 50\%$
- For *The S&P* and *The Fed*: $5\% \leq w \leq 30\%$

These weight limits are based on the notion that previous inversions of the yield curve and the credit spread HMMs are more reliable indicators of a future bear market than the yield curve itself and the stock market HMMs.

Furthermore, the steps between tested weights are calibrated to 5% in order to avoid overfitting and the sum of all the factor weights must be 100% in all occasions. With these restrictions, there are a total of 170 different factor weight allocations. All of the allocations are tested within the training period by simply calculating the return series just as in the previous sections with one simple allocation rule: if the probability of a future bear market exceeds 50% at time t in the comprehensive bear market predictor, the strategy invests in the risk-free rate in $t + 1$. The return series are calculated using MS Excel and VBA.

The method of choosing the appropriate weight allocation is similar to the one used in previous sections and is purely based on the total returns of the respected return series: the weight allocation with the best return in the training period of 1972–1985 is chosen. It should be noted, however, that because the weight allocation is optimized within the training period, the true performance of any allocation is truly measured in the pseudo out-of-sample period. Also, it should be noted that all of the 170 weight allocations do beat the stock market within the training period, which implies that the used factors and proposed weight allocations could be on the right track.

There were many factor weight allocations which reached the same total return during the training period, thus a new rule must be introduced to separate the allocations. The factors were sorted by their perceived importance using the first nine strategies and previous studies (1. *The Curve*, 2. *The Spread*, 3. *The Fed*, 4. *The S&P*) and the highest values were chosen according to the pecking order. The seven equal performing allocations are presented in Table 5.16, where the 1st weight allocation is chosen because of its focus on *The Curve.* If there were two allocations with the same emphasis on *The Curve*, then the next comparison would be done using *The Spread* and so on.

Table 5.16 Best performing factor weight allocations for 1972-1985

	The Curve	*The Fed*	*The Spread*	*The SP*	*31.12.1985*
1	40 %	5 %	35 %	20 %	$8,41
2	35 %	5 %	35 %	25 %	$8,41
3	35 %	10 %	35 %	20 %	$8,41
4	30 %	5 %	40 %	25 %	$8,41
5	35 %	5 %	40 %	20 %	$8,41
6	30 %	15 %	40 %	15 %	$8,41
7	30 %	10 %	40 %	20 %	$8,41

As visible in Table 5.16, the chosen factor weight allocation is 40% – 5% – 35% – 20% for the respective factors. With the proposed factors and weights it is now possible to propose a relatively simple formula which will be used in the pseudo out-of-sample period:

$$\begin{aligned} The\ BEAR_t = {} & 0{,}4 * The\ Curve_t + 0{,}05 * The\ Fed_t + 0{,}35 * The\ Spread_t \\ & + 0{,}2 * The\ S\&P_t \\ & + (\,TRC5Y_t < 0{,}5\ \rightarrow\ -30\%\ \ until\ \ t = T_{inv}) \end{aligned} \tag{5.2}$$

As stated before in this section, the investing strategy related to the comprehensive bear market predictor is very simple:

- if the probability of a future bear market ($BEAR_t$) exceeds 50% in t, we invest in the risk-free rate in $t + 1$

The weights of the factors are evaluated using the training period so it is relatively probable that the weights must be re-evaluated during the pseudo out-of-sample testing, but during the training period the proposed 10th strategy, *The BEAR*, suggested to stay out of the stock market during the following periods:

Table 5.17 Periods out of the stock market (*The BEAR*)

Inversion of the yield curve	*10. The BEAR*
1973:6	1973:10 - 1974:9
1978:11 (1980:10)	1980:3 - 1980:8; 1980:11 - 1982:3

The returns from the proposed comprehensive strategy quite obviously beats the market in the testing period of 1972–1985 since the best performing strategy is chosen out of the 170 different options, in other words the strategy was optimized after-the-fact and therefore it is not comparable to the other proposed strategies. The 10th strategy outperforms the stock market and all the earlier strategies. The true test will be, however, the testing period of 1986–2019.

To use the original performance comparison, a dollar invested in the beginning of 1972 had the following values in the end of 1985: $4,08 for the benchmark stock market and $8,41 for the proposed comprehensive bear market predictor strategy.

As stated before, the true performance of the 10th strategy will be really evaluated during the pseudo out-of-sample period because of the process. The main idea behind the 10th strategy is that the dynamics between different yield curve inversions would have a lot of similarities, even though every situation is unique. In this sense the 10th is not different from the other strategies as they rely on the similarities and predictability between different occasions. In order to make the 10th strategy also a bit more dynamic, the weights will be re-evaluated after every inversion in Chapter 6.

We have now formulated all of the proposed ten different strategies and their performance will be next evaluated. There were already differences between the performances of the strategies in the training period and it is probable that the differences will only grow larger during the longer evaluation period. Table 5.18 shows the annual returns of the strategies and the benchmark sorted by their performance. The main hypothesis for the evaluation period (1986–2019) is that the ranking should remain approximately the same. If not, the efficiency of the strategies is up to debate and it could be argued that the effects of the inverted yield curve are purely random.

Table 5.18 Annualized returns of the strategies (1972–1985)

Ranking	*Strategy*	*Annualized excess returns (%) 1986-2019*	*Standard deviation 1986-2019*	*Sharpe ratio*	*Skewness*	*Kurtosis*
1.	*10. The BEAR*	*10,83 %*	*13,14 %*	0,824	-0,881	7,599
2.	*10.1 The BEAR/SHORT*	*12,07 %*	*14,79 %*	0,817	-0,588	5,790
3.	*2. The Curve (latest)*	*10,40 %*	*13,31 %*	0,781	-0,865	7,279
4.	*9. TRCAPE*	*9,34 %*	*13,46 %*	0,694	-0,858	7,084
5.	*3. The Spread (history)*	*9,32 %*	*13,72 %*	0,679	-0,876	6,770
6.	*5. The S&P 500 (weighted averages)*	*9,13 %*	*14,01 %*	0,652	-0,825	6,259
7.	*4. The Spread (latest)*	*8,51 %*	*14,05 %*	0,606	-1,026	7,150
8.	*1. The Curve (history)*	*8,29 %*	*13,89 %*	0,597	-1,012	7,317
9.	*6. The S&P 500 (Viterbi)*	*8,20 %*	*14,37 %*	0,571	-0,838	5,904
10.	*8. The NY Fed model (latest)*	*7,81 %*	*14,66 %*	0,533	-0,954	6,245
11.	*7. The NY Fed model (history)*	*7,83 %*	*14,74 %*	0,531	-0,911	6,117
12.	*Total market return (Benchmark)*	*7,56 %*	*15,15 %*	0,499	-0,907	5,698

6 RESULTS

This study has now presented various allocation strategies that try to profit from the inversion of the inverted yield curve. All the inversions of the yield curve during the evaluation period (1986–2019) are gone through one by one to see how the strategies react and perform. Furthermore, the comprehensive bear market predictor will be reweighted after every inversion period.

The evaluation period is divided into four parts: the first (1986–1992) is characterized by the early 1990s recession, the second (1993–2003) by the rise and fall of the Dot-com bubble and the third (2004–2009) by The Great Recession. The fourth part (2010–2019) is the era of expansive monetary policies worldwide, but it is not really evaluated since the inverted period is cut short by the end of the data set and most of the strategies just invest in the stock market the whole period.

The first three sections of the chapter belong to the first three parts of the evaluation period, respectively. In the last section, the strategies are evaluated thoroughly during the whole evaluation period and we try to imitate the realistic choices that would have probably been made during the testing period.

6.1 The early 1990s recession (1986–1992)

The first part of the evaluation period starts in the beginning of 1986, the inversion of the yield curve takes place in June 1989 and the end of the period is 42 months after the inversion in the end of 1992. The end of the 80s was an era of rapid growth and is characterized by strong total stock market returns, even though the era contained the Black Monday of 1987 and the oil crisis of 1990.

Almost all of the proposed strategies performed well and managed to beat the market, which also enjoyed phenomenal (13,93% p.a., whereas historical average using French 2020a data 1928–2019 is 9,74% p.a.) annualized returns. The best performing strategies of the training period also topped the list during the first part of the evaluation period, which could be seen as an implication that the captured effects during the training period should not be totally random. The value of the $1 invested in the beginning of 1986 until the end of 1992 and the annualized returns for the strategies and the benchmark are presented in Table 6.1.

Table 6.1 Strategy performance 1986–1992

Ranking	*Strategy*	*Annualized returns (%) 1986-1992*	*$1 (1.1.1986 - 31.12.1992)*
1.	*10. The BEAR*	16,14 %	$2,85
2.	*2. The Curve (latest)*	15,88 %	$2,81
3.	*9. TRCAPE*	15,29 %	$2,71
4.	*8. The NY Fed model (latest)*	14,99 %	$2,66
5.	*6. The S&P 500 (Viterbi)*	14,95 %	$2,65
6.	*1. The Curve (history)*	14,89 %	$2,64
7.	*4. The Spread (latest)*	14,89 %	$2,64
8.	*5. The S&P 500 (weighted averages)*	14,44 %	$2,57
9.	*Total market return (Benchmark)*	13,93 %	$2,49
10.	*7. The NY Fed model (history)*	13,37 %	$2,41
11.	*3. The Spread (history)*	12,90 %	$2,34

The periods out of the market for the proposed strategies are presented in Chapter 5 except for *The BEAR* because of the dynamic weight allocations after every inversion. With the weight allocations established in the training period that are presented in Equation 5.2, the model suggested to stay out of the stock market between 1989:10 and 1990:12. The model manages to avoid excessive trading as it signals only a single period and thus minimizes transaction costs. On the other hand, this is no surprise as the weights of the strategy components with fixed time spans (*The Curve* 40% and *The Spread* 35%) dominate the components with the possibility of multiple trades (*The Fed* 5% and *The S&P* 20%). The TRCAPE backstop is not used during the first part of the evaluation period as the lowest level of $TRC5Y_t$ is 76,98%, which is reached in 1990:10.

The top 3 strategies of the training period and their performance during 1989–1992 are presented visually in Figure 6.1. Whereas the 2nd and 10th strategies perform at the same level as in the training period, the 1st strategy fails to live up to the expectations, which can be seen also in Table 6.1. The underperforming strategy manages, however, to beat the market anyway.

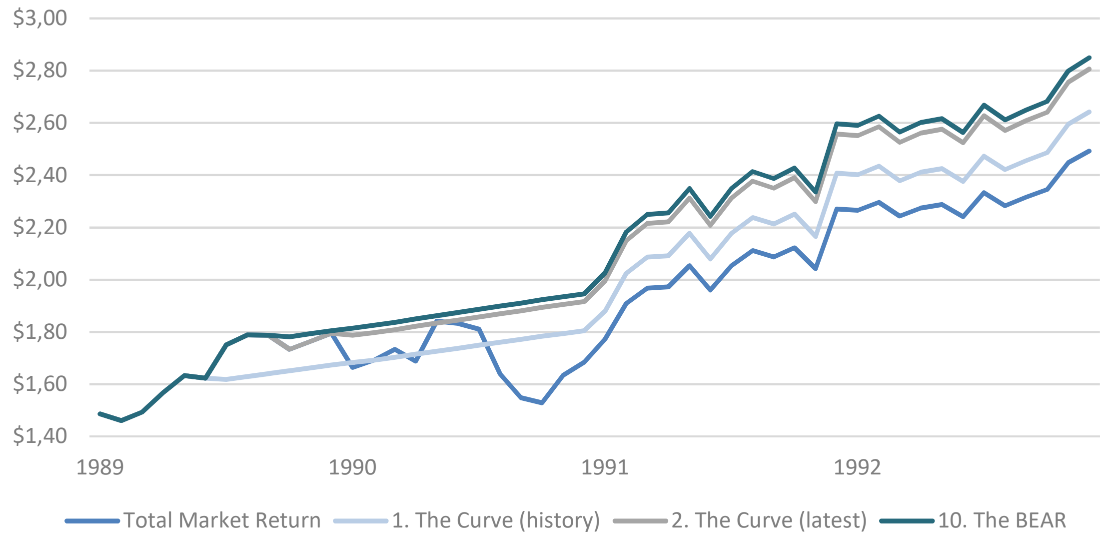

Figure 6.1 Assumed top 3 strategies' performance 1989–1992

As stated in Chapter 5, a total of 170 different weight allocations were tested for *The BEAR*. Because the strategy was optimized within the training period, the weight allocations are re-evaluated after the first part of the evaluation period according to the rules presented in Section 5.6. The best performing allocations for the period of 1986–1992 are shown in Table 6.2.

Table 6.2 *The BEAR* allocations 1986–1992

Ranking	***The Curve***	***The Fed***	***The Spread***	***The S&P***	***Annualized returns % (1986-1992)***
1.	50 %	5 %	35 %	10 %	16,14 %
2.	40 %	10 %	35 %	15 %	16,14 %
3.	40 %	20 %	30 %	10 %	16,14 %
4.	35 %	25 %	30 %	10 %	16,14 %
5.	30 %	15 %	35 %	20 %	16,14 %
6.	30 %	25 %	30 %	15 %	16,14 %
7.	25 %	30 %	30 %	15 %	16,14 %

The weight allocations of *The BEAR* are now re-evaluated according to the best performing weight allocations. The tiebreaking rules are also presented in Section 5.6 and thus the new formula for the comprehensive bear market predictor can be presented as:

$$\begin{aligned} The\ BEAR_t = {} & 0{,}5 * The\ Curve_t + 0{,}05 * The\ Fed_t + 0{,}35 * The\ Spread_t \\ & + 0{,}1 * The\ S\&P_t \\ & + (\ TRC5Y_t < 0{,}5\ \rightarrow\ -30\%\ \ until\ \ t = T_{inv}) \end{aligned} \tag{6.1}$$

These weights will now be used for the next evaluation period of 1993–2003. In other words, *The BEAR* is also re-evaluated for the upcoming period. It should be noted that other strategies are also dynamic and thus re-evaluated after each inversion. For the re-evaluation of the other strategies, please see Chapter 5.

6.2 The rise and fall of the Dot-com bubble (1993–2003)

The second part of the evaluation period starts in the beginning of 1993, the inversion of the yield curve takes place in July 2000 and the end of the period is 41 months after the inversion in the end of 2003. The end of the 90s was the era of the original "irrational exuberance" as stated by Alan Greenspan in 1996 and is characterized by overvalued tech stocks with Nasdaq leading the way. The Nasdaq Composite stock market index rose 400% between 1995 and 2000 only to fall again 78% by October 2002.

Most of the proposed strategies performed well during the period and managed to beat the market, which had annualized returns (10,7% p.a.) mostly in line with historical averages. The best performing strategies of the training period and the first evaluation period also topped the list during the second part of the evaluation period, which further solidifies the efficiency of the best performing strategies. *The BEAR* especially continued its strong performance against other strategies and the benchmark.

The value of the $1 invested in the beginning of 1986 until the end of 2003 and the annualized returns during the second part (1993–2003) for all of the strategies and the benchmark are presented in Table 6.3.

Table 6.3 Strategy performance 1993–2003

Ranking	*Strategy*	*Annualized returns (%) 1993-2003*	*$1 (1.1.1986 - 31.12.2003)*
1.	*10. The BEAR*	13,72 %	$11,72
2.	*1. The Curve (history)*	13,43 %	$10,57
3.	*2. The Curve (latest)*	12,54 %	$10,30
4.	*7. The NY Fed model (history)*	12,54 %	$8,83
5.	*9. TRCAPE*	12,42 %	$9,81
6.	*3. The Spread (history)*	11,91 %	$8,06
7.	*4. The Spread (latest)*	11,91 %	$9,11
8.	*5. The S&P 500 (weighted averages)*	11,56 %	$8,57
9.	*8. The NY Fed model (latest)*	11,46 %	$8,77
10.	*Total market return (Benchmark)*	10,69 %	$7,61
11.	*6. The S&P 500 (Viterbi)*	8,09 %	$6,24

The period out of the stock market proposed by *The BEAR* is fairly similar to the period during the first part of evaluation period. With the weight allocations chosen according to the best performance during 1986–1992 as presented in Equation 6.1, the model suggested to stay out of the stock market between 2001:2 and 2002:5. This time, however, *The BEAR* does not come up with a single period as it suggests also to stay out of the stock market in 2002:7, which is clearly suboptimal. Furthermore, *The BEAR* misses two miserable months (-10,21% in 2000:11 and -10,21% in 2002:9) at both ends, but it still outperforms all the other strategies and the benchmark. The TRCAPE backstop is not really used during the second part of the evaluation period even though $TRC5Y_t$ reaches 49,95% in 2002:9 as *The BEAR* is already below 50%.

The top 3 strategies of the training period and their performance during 2000–2003 are presented visually in Figure 6.2. The 1st and 10th strategies are the clear winners here, whereas the 2nd strategy loses some ground despite performing quite well.

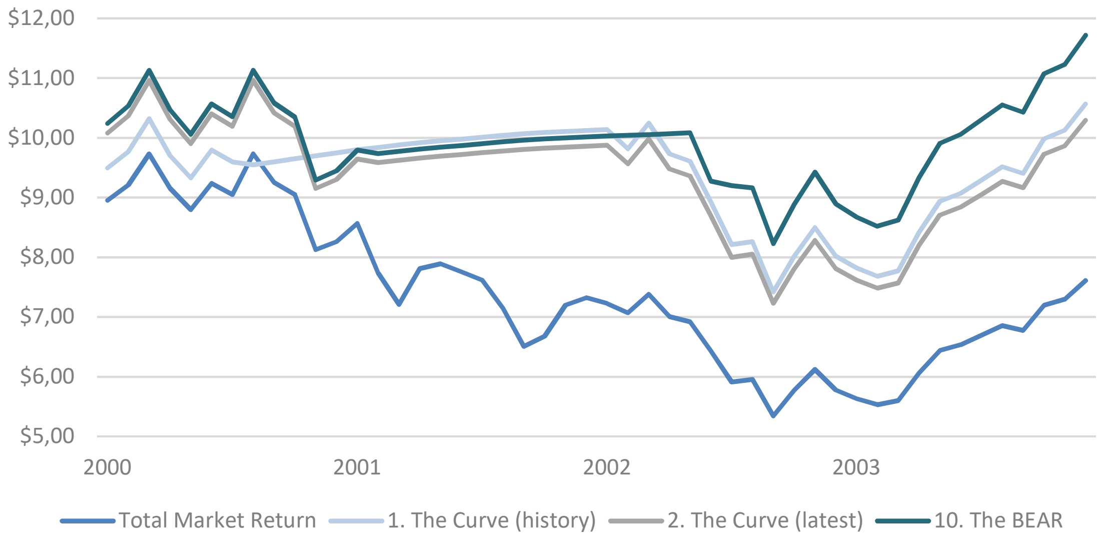

Figure 6.2 Assumed top 3 strategies' performance 2000–2003

Once again, the weights of the factors of *The BEAR* are rebalanced according to the performance of the weight allocations during the second part of the evaluation period (1993–2003). The best performing allocations of the period are presented in Table 6.4.

Table 6.4 *The BEAR* allocations 1993–2003

Ranking	***The Curve***	***The Fed***	***The Spread***	***The S&P***	***Annualized returns % (1993-2003)***
1.	50 %	5 %	40 %	5 %	14,70 %
2.	45 %	10 %	40 %	5 %	14,70 %
3.	40 %	15 %	40 %	5 %	14,70 %
4.	50 %	10 %	35 %	5 %	13,72 %
5.	50 %	5 %	35 %	10 %	13,72 %
6.	45 %	10 %	35 %	10 %	13,72 %
7.	35 %	20 %	40 %	5 %	13,61 %

The new rebalanced formula for *The BEAR* can be presented as:

$$\begin{aligned} The\ BEAR_t = {} & 0{,}5 * The\ Curve_t + 0{,}05 * The\ Fed_t + 0{,}4 * The\ Spread_t \\ & + 0{,}05 * The\ S\&P_t \\ & + (\ TRC5Y_t < 0{,}5\ \rightarrow\ -30\%\ \ until\ \ t = T_{inv}) \end{aligned} \tag{6.2}$$

The rebalanced weights will now be used for the next evaluation period of 2004–2009. It should be noted that many of the top performing allocations are the same as in the earlier part of the evaluation period. Furthermore, it should be noted that the factors related to the history of the inverted yield curve and the credit spread really seem to dominate the other factors as their weights only grow larger even though it could be possible to have 25% weights for each of the four factors.

6.3 The Great Recession (2004–2009)

The third part of the evaluation period starts in the beginning of 2004, the inversion of the yield curve takes place in August 2006 and the end of the period is 40 months after the inversion in the end of 2009. The later part of the 2000s is characterized by loose credit conditions as illustrated best by subprime mortgages and collateralized debt obligations, that lost most of their value in 2007–2008. The problems within the financial system culminated in the bankruptcy of Lehman Brothers, which resulted in widespread panic in financial markets. Many banks needed governmental assistance afterwards, which lead to public outrage.

Most of the proposed strategies performed well during the period and managed to beat the market, which had annualized returns (2,73%) that were way below historical averages. The best performing strategy of the training period and the first two evaluation periods also topped the list during the third part of the evaluation period, which further solidifies the efficiency of the best performing strategy, *The BEAR*. Also, the strategy using the latest inversion of the yield curve continued its strong performance against other strategies and the benchmark.

The value of the $1 invested in the beginning of 1986 until the end of 2009 and the annualized returns during the third part (2004–2009) for the strategies and the benchmark are presented in Table 6.5.

Table 6.5 Strategy performance 2004–2009

Ranking	*Strategy*	*Annualized returns (%) 2004-2009*	*$1 (1.1.1986 - 31.12.2009)*
T-1.	*10. The BEAR*	12,78 %	$24,12
T-1.	*2. The Curve (latest)*	12,78 %	$21,19
3.	*3. The Spread (history)*	11,28 %	$15,31
4.	*6. The S&P 500 (Viterbi)*	10,27 %	$11,21
5.	*5. The S&P 500 (weighted averages)*	9,76 %	$14,98
6.	*9. TRCAPE*	9,56 %	$16,96
7.	*4. The Spread (latest)*	4,67 %	$11,97
8.	*Total market return (Benchmark)*	2,73 %	$8,95
9.	*1. The Curve (history)*	2,59 %	$12,32
T-10.	*7. The NY Fed model (history)*	1,64 %	$9,73
T-10.	*8. The NY Fed model (latest)*	1,64 %	$9,67

The period out of the stock market proposed by *The BEAR* is exactly the same as in the 2nd strategy, which is not a complete surprise as 50% of *The BEAR* consists of the same strategy after the reweighing of 2003. With the weight allocations chosen according to the best performance during 1993–2003 as presented in Equation 6.2, the model suggested to stay out of the stock market between 2006:8 and 2009:2. The backstop $TRC5Y_t$ reaches 49,17% in 2009:3, but *The BEAR* had already gone below 50% so it has no real effect just as in the previous part of the evaluation period.

The top 3 strategies of the training period and their performance during 2006–2009 are presented visually in Figure 6.3. *The BEAR* continues to outperform the other strategies, whereas the 1st strategy loses to the benchmark. The 2nd and 10th strategies manage to get back to the stock market at the exact bottom, which could be considered a bit lucky, but had they got back to the market a month earlier or later, their performance would have been rather solid anyway.

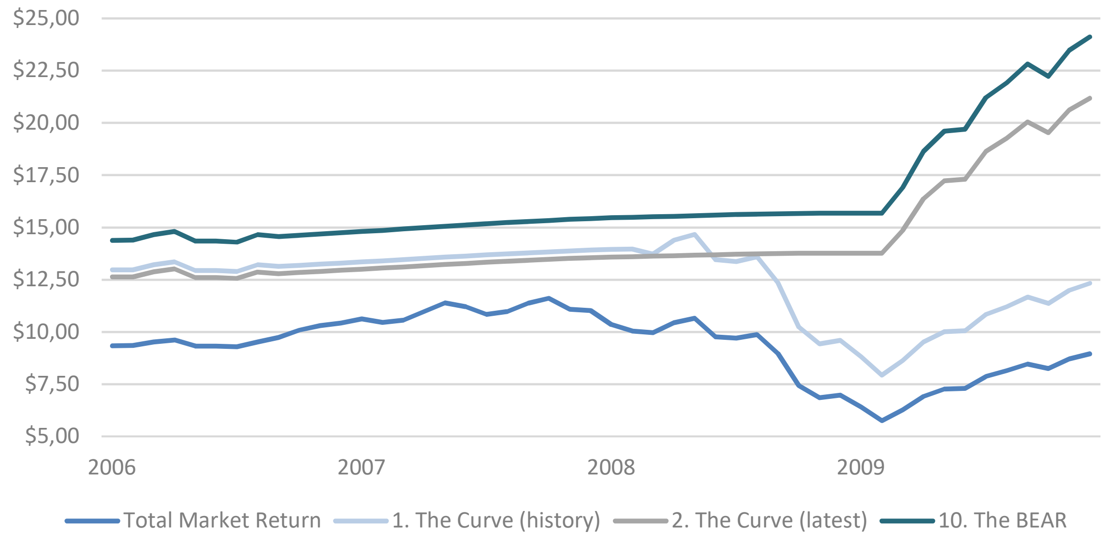

Figure 6.3 Assumed top 3 strategies' performance 2006–2009

Once again, the weights of the factors of *The BEAR* are rebalanced according to the performance of the weight allocations during the third part of the evaluation period (2004–2009). The best performing allocations of the period are presented in Table 6.6.

Table 6.6 *The BEAR* allocations 2004–2009

Ranking	*The Curve*	*The Fed*	*The Spread*	*The S&P*	*Annualized returns % (2004-2009)*
1.	45 %	5 %	45 %	5 %	14,56 %
2.	45 %	5 %	40 %	10 %	14,56 %
3.	45 %	5 %	35 %	15 %	14,56 %
4.	45 %	5 %	30 %	20 %	14,56 %
5.	40 %	10 %	45 %	5 %	14,56 %
6.	40 %	5 %	45 %	10 %	14,56 %
7.	40 %	10 %	40 %	10 %	14,56 %

The results of the different allocations in the third part of the evaluation period shows notable differences between the chosen factor weight allocation and the best ones. In total there were 22 weight allocations that performed better than the used weight allocation. This is the first indication that the underlying factors are not the same for every period, even though the weight differences are not large. The 50% reliance on the latest yield curve inversion seems to be problematic in this case as the strategy is not dynamic enough to take the changing landscape into account. It should be noted that *The BEAR* with the chosen factor weight allocation was anyway the best performing strategy and it outperformed the benchmark by a clear margin.

The new rebalanced formula for *The BEAR* can be now presented as:

$$The\ BEAR_t = 0{,}45 * The\ Curve_t + 0{,}05 * The\ Fed_t + 0{,}45 * The\ Spread_t + 0{,}05 * The\ S\&P_t + (\ TRC5Y_t < 0{,}5\ \rightarrow\ -30\%\ \ until\ \ t = T_{inv}) \tag{6.3}$$

These rebalanced weights will now be used for the next inversion period of 2010–2022. It should be noted, however, that the data set of this study ends in 2019 and therefore there will be no further analysis concluded for the period in question. The results for the whole evaluation period of 1986–2019 will be presented in the following section.

As a conclusive remark to the different factor weight allocations, it should be noted that the factors related to the history of the inverted yield curve and the credit spread really seemed to dominate the other factors during the whole period and their weights only grew larger with time. The main reason behind this might be just that the past inversions combined with the dynamic element of credit spread condition signals manage to time the market better or, alternatively, the rescaling of *The S&P* and *The Fed* was not done in an effective way. This topic is further discussed in the following chapter.

6.4 Total returns and investor indications

The proposed strategies did very well during the whole evaluation period as they all managed to beat the stock market, which was used as the benchmark of the analysis. The US stock market had annualized returns of 10,72% during the evaluation period of 1986–2019, which implies that it was a good period for stocks if compared to the long-term average annualized stock market returns (9,74% p.a., see Section 6.1).

The same strategies performed the best in all inversions, even though there were some differences between different periods. *The BEAR* outperformed other strategies in every inversion, whereas the 2nd and 9th strategies proved to have consistent results in every window.

The value of the $1 invested in the beginning of 1986 until the end of 2019 and the annualized returns during the whole evaluation period (1986–2019) for the strategies and the benchmark are presented in Table 6.7. It should be noted that because of the inversion of the yield curve in 2019:5 the total returns of the proposed strategies are not totally comparable, because some of the strategies have already made their allocation shifts and some have not. The corrected (no transactions, only stock market returns in 2019) returns are presented in parenthesis for the strategies which have already made their allocation shifts.

Table 6.7 Strategy performance 1986–2019

Ranking	***Strategy***	***Annualized returns (%) 1986-2019***	***$1 (1.1.1986 - 31.12.2019)***
1.	*10. The BEAR*	14,00 %	$85,95
2.	*2. The Curve (latest)*	13,56 %	$75,51
3.	*9. TRCAPE*	12,51% (12,82%)	$54,93 ($60,45)
4.	*3. The Spread (history)*	12,48 %	$54,56
5.	*5. The S&P 500 (weighted averages)*	12,30% (12,41%)	$51,57 ($53,38)
6.	*4. The Spread (latest)*	11,67 %	$42,68
7.	*1. The Curve (history)*	11,45% (11,77%)	$39,91 ($43,92)
8.	*6. The S&P 500 (Viterbi)*	11,36% (11,46%)	$38,83 ($39,96)
9.	*7. The NY Fed model (history)*	10,99 %	$34,69
10.	*8. The NY Fed model (latest)*	10,97 %	$34,46
11.	*Total market return (Benchmark)*	10,72 %	$31,90

If we take our pseudo out-of-sample analysis further, it would seem obvious that the rational investor would have chosen the best performing strategy after every part of the evaluation period. As *The BEAR* dominated each section, the rational investor would have chosen the strategy in question for every part of the evaluation period and hence the total return for the rational investor would have been the one shown on top of the list in Table 6.7. This whole analysis is, obviously, hypothetical, but the consistency of *The BEAR* is rather convincing in terms of the efficiency and continuity of the strategy.

The top 3 strategies of the training period and their performance during 1986–2019 are shown in Figure 6.4. The three strategies performed really well and the only newcomer in the top 3 during the evaluation period was the 9th strategy, which had solid returns in every part of the evaluation period, even though it never excelled, which shows the strength of compound interest. Apart from the 1st strategy ending in the 4th place, the same strategies, which performed the best during the training period, also performed the best during the whole evaluation period. The consistency of the strategies – and especially that of *The BEAR* – solidifies the fact, that the effects of the inverted yield curve do not seem to be totally random and that those effects can be exploited.

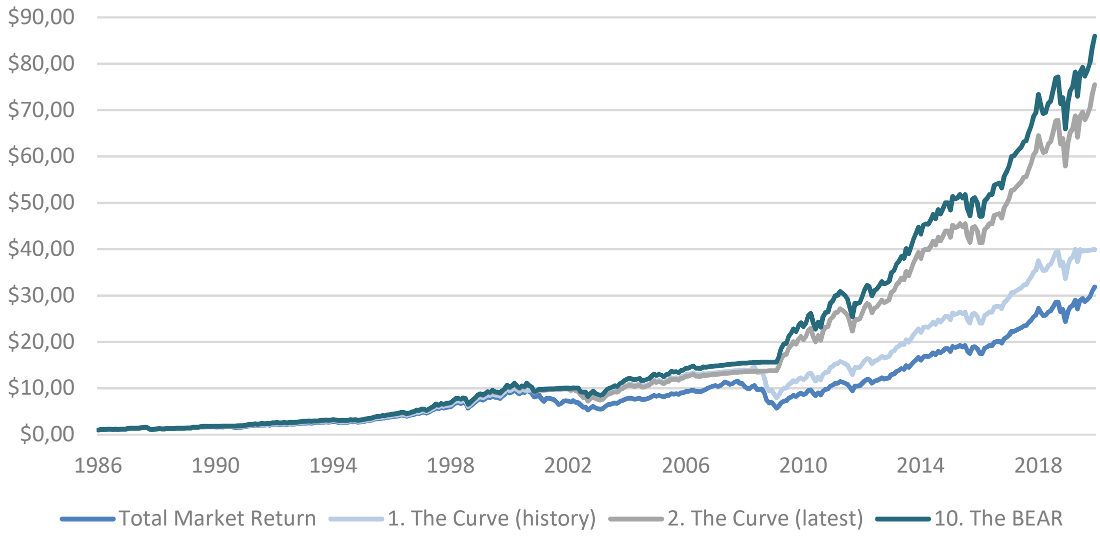

Figure 6.4 Assumed top 3 strategies' performance 1986–2019

Even though the long-term returns of *The BEAR* are clearly higher than the returns of the benchmark or the other strategies, past performance is no guarantee of future results. The evidence gathered in the analysis is, however, quite strong and *The BEAR* has proven to be a viable option going forward. We will conduct further analysis and testing in the next section, where we will discuss the choices made and test different assumptions.

6.5 Further testing and validation

This section is basically divided into three parts: first, we will discuss transaction costs; second, we will test *The BEAR* with shorting the stock market and investing in cash as opposed to the risk-free rate; and third, we will discuss the shortcomings of rescaling *The Fed* and *The S&P* when used as factors in *The BEAR*.

6.5.1 Transaction costs

The empirical part of the study is made with the assumption of 0,5% transaction costs so every trade basically costs 1% (0,5% for selling stocks/risk-free rate and 0,5% for buying stocks/risk-free rate). As different strategies trade at different frequencies, they are affected by transaction costs in different ways, which is why a simple sensitivity analysis is performed with various transaction cost levels for the strategies. The results are shown in Table 6.8.

The annualized returns which fall below the benchmark are highlighted in red. As we can see from the table, the strategies using the S&P 500 HMMs and the NY Fed model suffer more from higher transaction costs as they trade more frequently. It should be noted that the examined high transaction costs (>1%) are way higher than the ones usually used in previous studies (see e.g. Guidolin & Timmermann 2007; Xin et al. 2016, 67), but the high transaction costs are used to demonstrate the usefulness of the simple strategies as their excess returns are not dependable on minimal transaction costs. Also, the strategies usually trade in market conditions in which fear and uncertainty prevail, which leads to wide spreads and thus high transaction costs.

Table 6.8 Transaction cost sensitivity analysis

Ranking	*Strategy (Annualized returns 1986-2019)*	*0 % Transaction costs*	*0,5 % Transaction costs*	*1 % Transaction costs*	*1,5 % Transaction costs*	*2 % Transaction costs*
1.	*10. The BEAR*	14,27 %	14,00 %	13,72 %	13,45 %	13,17 %
2.	*2. The Curve (latest)*	13,76 %	13,56 %	13,36 %	13,15 %	12,95 %
3.	*9. TRCAPE*	12,74 %	12,51 %	12,27 %	12,03 %	11,79 %
4.	*3. The Spread (history)*	12,68 %	12,48 %	12,28 %	12,08 %	11,87 %
5.	*5. The S&P 500 (weighted averages)*	12,96 %	12,30 %	11,63 %	10,96 %	10,28 %
6.	*4. The Spread (latest)*	11,87 %	11,67 %	11,47 %	11,27 %	11,07 %
7.	*1. The Curve (history)*	11,68 %	11,45 %	11,22 %	10,99 %	10,75 %
8.	*6. The S&P 500 (Viterbi)*	11,82 %	11,36 %	10,90 %	10,43 %	9,96 %
9.	*7. The NY Fed model (history)*	11,26 %	10,99 %	10,73 %	10,46 %	10,19 %
10.	*8. The NY Fed model (latest)*	11,23 %	10,97 %	10,71 %	10,44 %	10,17 %
11.	*Total market return (Benchmark)*	10,72 %	10,72 %	10,72 %	10,72 %	10,72 %

The best performing strategies can clearly endure the higher transaction costs, which practically makes the strategies available also for private investors, who have to pay larger transaction costs than institutional investors in practice. All in all, most of the strategies seem feasible for the examined reasonable transaction cost levels.

6.5.2 Shorting the stock market

Another interesting aspect to do further testing on is replacing the risk-free rate with shorting the stock market or just keeping cash at hand while not investing in the stock market. As *The BEAR* has proven to be the best performing strategy for every part of the evaluation period, this analysis is conducted only for the strategy in question. Some assumptions are made regarding the analysis and the short positions:

- 0,5% transaction cost for every trade (just as before)
- 2% p.a. cost for shorting the market (0,1652% per month)
- 0,5% trade cost for cash/stock market allocations (single buy/sell)

The annualized returns for shorting the market or liquidating all of the assets according to *The BEAR* are shown in Table 6.9. In order to really compare the results, we divide the evaluation period into three parts – just as before – and look at each part separately, but also compare the strategies during the whole evaluation period.

Table 6.9 Alternative allocations (*The BEAR*)

Strategy	***Annualized returns 1986-1992***	***Annualized returns 1993-2003***	***Annualized returns 2004-2009***	***Annualized returns 1986-2019***
The BEAR	16,14 %	13,72 %	12,78 %	14,00 %
The BEAR/SHORT	14,56 %	15,55 %	18,32 %	15,24 %
The BEAR/CASH	14,73 %	13,51 %	11,51 %	13,42 %
Total market return	13,93 %	10,69 %	2,73 %	10,72 %

Using the strategy to decide the periods to short the stock market seems to be the best option two times out of three. Also, the choice to short outperforms the choice to invest in the risk-free rate during the whole evaluation period. As risk-free rates have become lower and lower, the option to keep cash in hand have become more and more tempting, but the total returns are naturally lower with the assumed 0,5% transaction costs.

Shorting the market is, obviously, the more volatile option as it mirrors the turbulent stock market and it can lead to significant losses, when the market reverts to more bullish regimes. For example, in 1990:11 (+6,92%) and 1990:12 (+3,06%) *The BEAR* misses the start of the bull market, which is not so serious if investing in the risk-free rate, whereas shorting the market in such conditions is devastating to portfolio performance.

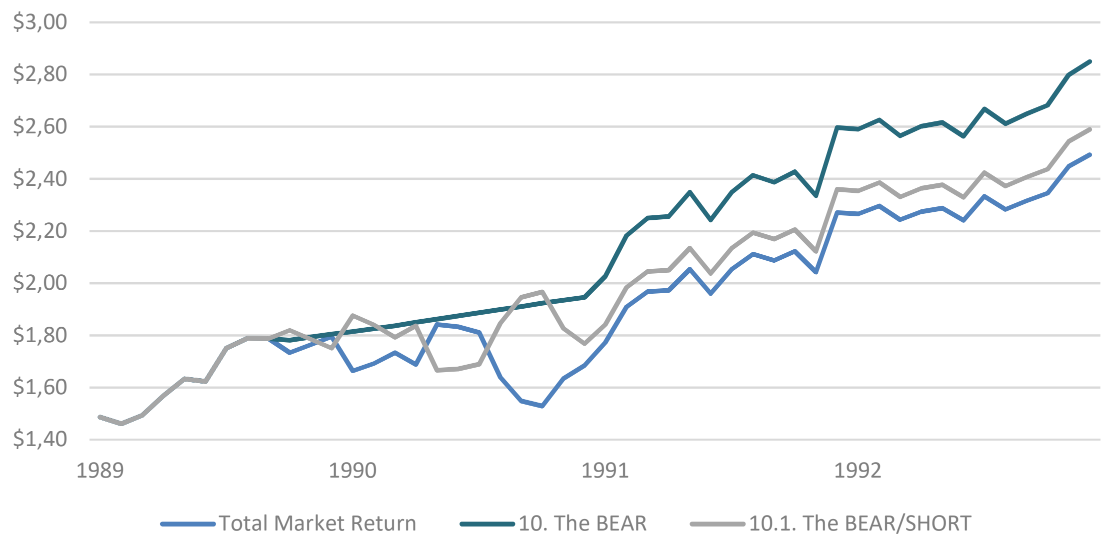

Figure 6.5 ***The BEAR*** **vs.** ***The BEAR/SHORT*** **1989–1992**

Obviously, the better the strategy manages to time the market, the larger the returns that can be achieved by shorting the market. The periods of turbulence and volatility are also periods of great opportunities, in other words the upside risk should not be totally ignored.

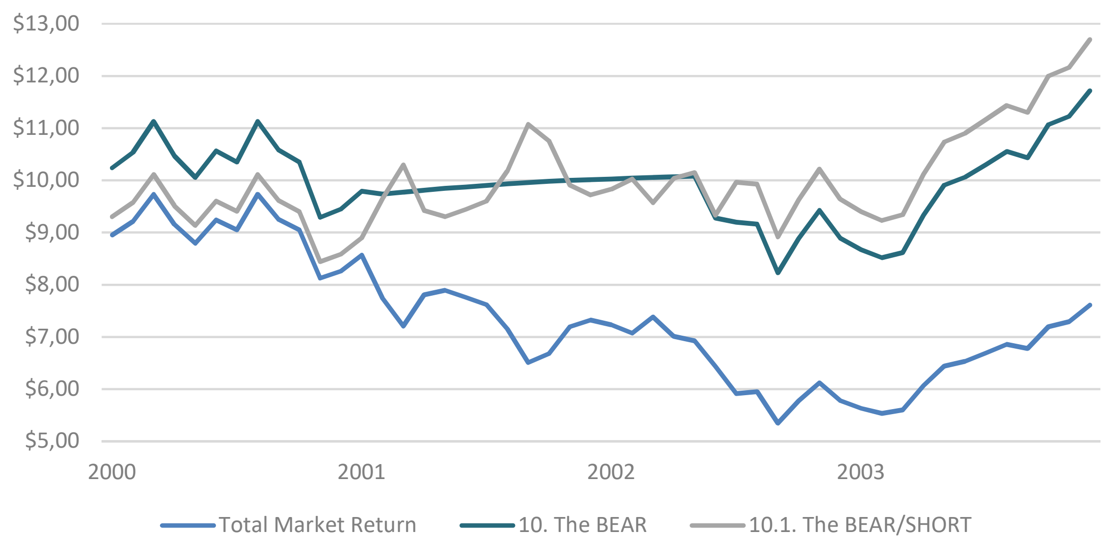

Figure 6.6 ***The BEAR*** **vs.** ***The BEAR/SHORT*** **2000–2003**

Whether or not to use shorts to boost the performance of the strategy depends on the risk tolerance of the investor, even though the results do seem to favor the option to short. On the other hand, the phenomenon of the inverted yield curve is dynamic and shorting the market gives less margin for error and increases the volatility of the portfolio.

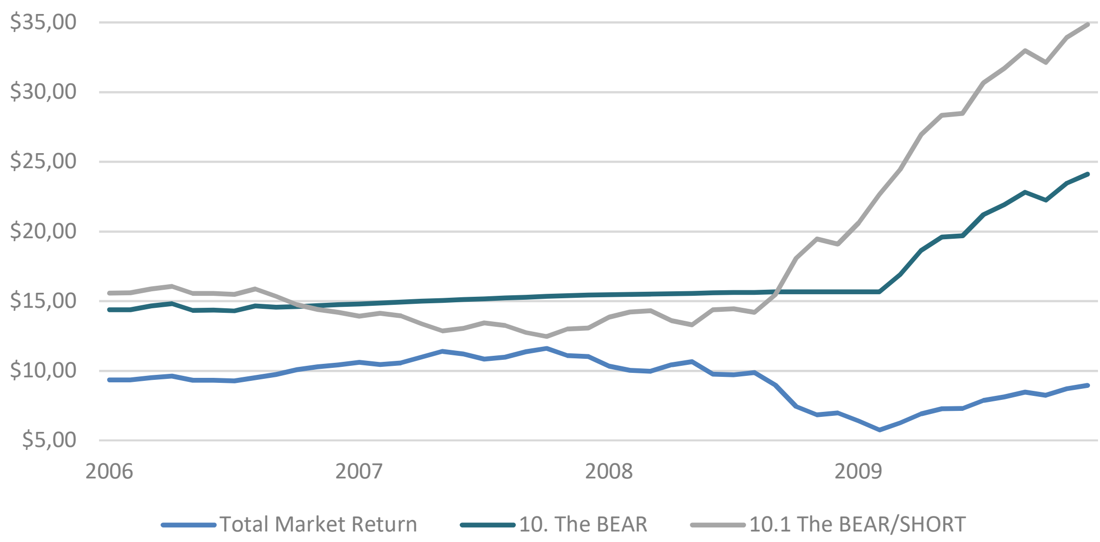

Figure 6.7 ***The BEAR*** **vs.** ***The BEAR/SHORT*** **2006–2009**

6.5.3 Rescaling the factors & Risk-adjusted returns

Third topic in this section is the possible shortcomings of rescaling *The Fed* and *The S&P* as their weight allocation in *The BEAR* are very small and only diminishing. Table 6.10 shows the annualized returns of the rescaled *The Fed* and *The S&P*, in other words the returns as if they were the only components (100% weight allocation, respectively) of *The BEAR*.

Table 6.10 The rescaled strategies' performance (***The Fed*** **and** ***The S&P*****)**

Strategy	*Annualized returns 1986-1992*	*Annualized returns 1993-2003*	*Annualized returns 2004-2009*	*Annualized returns 1986-2019*
The Rescaled Fed	13,99 %	8,93 %	1,15 %	9,59 %
The Rescaled S&P	13,23 %	8,04 %	1,69 %	8,83 %
Total market return	13,93 %	10,69 %	2,73 %	10,72 %

The below-the-market returns present a serious question, whether the rescaling of the components was done in a sensible way, because the original strategies do perform quite well. The rescaling is clearly not optimal, but combining different indicators and strategies require compromises. Finding better ways to rescale and combine the original strategies is left for further research.

Finally, risk-adjusted returns for the strategies and the benchmark are presented in Table 6.11. The ranking of the strategies remains basically the same as all the strategies try to avoid the most volatile periods of the stock market and the standard deviations of the strategies vary between 13,14% and 14,79%, which are all below the benchmark.

The excess returns are calculated with the previous returns deducted with the annualized return (3,16%) of the risk-free rate (French 2020a) between 1986 and 2019. The (annualized) standard deviations are calculated from the simple monthly returns of the strategies. Furthermore, skewness and kurtosis of the series are presented in the table to further demonstrate the differences between the strategies. *The BEAR* manages to keep the top position as it provides the best risk return ratio, but it should be noted that it has rather fat tails.

Table 6.11 Risk-adjusted returns 1986-2019

Ranking	***Strategy***	***Annualized excess returns (%) 1986-2019***	***Standard deviation 1986-2019***	***Sharpe ratio***	***Skewness***	***Kurtosis***
1.	*10. The BEAR*	*10,83 %*	*13,14 %*	0,824	-0,881	7,599
2.	*10.1 The BEAR/SHORT*	*12,07 %*	*14,79 %*	0,817	-0,588	5,790
3.	*2. The Curve (latest)*	*10,40 %*	*13,31 %*	0,781	-0,865	7,279
4.	*9. TRCAPE*	*9,34 %*	*13,46 %*	0,694	-0,858	7,084
5.	*3. The Spread (history)*	*9,32 %*	*13,72 %*	0,679	-0,876	6,770
6.	*5. The S&P 500 (weighted averages)*	*9,13 %*	*14,01 %*	0,652	-0,825	6,259
7.	*4. The Spread (latest)*	*8,51 %*	*14,05 %*	0,606	-1,026	7,150
7.	*1. The Curve (history)*	*8,29 %*	*13,89 %*	0,597	-1,012	7,317
8.	*6. The S&P 500 (Viterbi)*	*8,20 %*	*14,37 %*	0,571	-0,838	5,904
9.	*8. The NY Fed model (latest)*	*7,81 %*	*14,66 %*	0,533	-0,954	6,245
10.	*7. The NY Fed model (history)*	*7,83 %*	*14,74 %*	0,531	-0,911	6,117
11.	*Total market return (Benchmark)*	*7,56 %*	*15,15 %*	0,499	-0,907	5,698

7 CONCLUSION

Previous studies have shown very clearly that there is a connection between the inversion of the yield curve and future recessions measured as contractions in the GDP. The focus of this study was the connection between financial markets and the inverted yield curve, which has not been as studied as the production related dependences. Even though the mechanisms of Wall Street and Main Street differ from another, it is obvious that there are some similarities, which can also be seen in the rather similar consequences of the inversion of the yield curve.

To return to the original research question, how does the inverted yield curve predict future stock market crashes? The empirical part of the study would seem to indicate that there is, in fact, some predictive power in the inverted yield curve concerning stock market crashes. Although it is difficult to draw a clear line between stock market crashes and the preceding inversions of the yield curve, the data suggests that the inverted yield curve is a sign of more volatile times ahead, which clearly heightens the risk of a major stock market downturn. It is, however, very hard to address cause and effect: Does the inverted yield curve *per se* predict future stock market crashes or is it just a consequence of larger changes in the market? Whatever the deeper reason behind the phenomenon, the inverted yield curve seems to be a reliable indicator also in the stock market as noted in the empirical part of the study.

The connections between the inverted yield curve, real economy and financial markets seems to be rather strong, but one of the main objectives of the study was to find factors and tools that could be combined with the inverted yield curve in order to improve the predictability of future bear markets. Various alternatives were introduced, but the best options seemed to be credit spread HMMs and HMMs related to the stock market itself. The credit spread especially seems to be a good measure of market sentiment that reflects the expectations and risk appetite of credit investors that has clear implications on the stock market and can thus improve the predictability of future stock market crashes, if combined with the inverted yield curve.

The results obtained by modeling the stock market (S&P 500) itself were not so clear, but it is no surprise that the stock market itself has at least some indications on future bear markets. The improvements on the predictability, however, remained rather small probably because of the noise in the stock market, which affects the efficiency of the HMM and thus prevents clear signals.

The proposed strategies would seem to indicate that it is indeed possible to construct successful investment strategies using the inverted yield curve and other related components. All of the proposed strategies managed to beat the benchmark and the conducted sensitivity analysis proves that the strategies are feasible at reasonable transaction cost levels. The greater annualized returns and lesser volatility of the strategies lead to significantly better risk-adjusted returns than what one would achieve using the simple buy and hold strategy. The obtained Sharpe ratios range from 0,531 to 0,824, whereas the Sharpe ratio of the benchmark total stock market return is 0,499 during the (pseudo) out-of-sample forecasting period of 1986–2019. It should be noted that the analysis is conducted using past data and the true acid test of the strategies will be during the current inversion period.

The best performing strategy is the comprehensive bear market predictor (The *BEAR*), which relies heavily on the latest inversion of the yield curve and the credit spread HMM. Why are inverted yield curve and credit spread strategies so successful when combined? One proposition could be that the two factors may very well be the most reliable and constant indicators of general market sentiment, which is why they perform the best together. History mostly repeats itself, which is captured by the factor exploiting the latest inversion, but the phenomenon is not static by its nature, which is captured by the credit spread HMM that manages to detect the dynamic changes in the atmosphere.

The big lingering question throughout the study has been the potential excess returns that do not seem to fade away with time. If the inverted yield curve predicts the future reliably and thus renders excess returns possible, why has this opportunity not been already captured by efficient markets? It simply makes no sense that the phenomenon stubbornly persists despite of excessive research and various occurrences during the last 30 years. Rudebusch & Williams (2009, 492) refer to this as the inverted yield curve puzzle: economists seem to downplay the predictive power of the inverted yield curve despite the strong evidence since late 1980s. Possible explanations include shorter explanation horizons (inverted yield curve works well at horizons beyond two quarters); systematic underestimations of the effects of changes in monetary policies as proxied by the yield curve; and finally, simply arguing that the formidable past performance did not apply in the current situation (Rudebusch & Williams 2009, 501–502).

Rudebusch & Williams focus on economists predicting output and recessions, but the same explanations could apply to financial markets as well. Especially the reluctance to see similarities between different situations seems to be catastrophic for financial markets and economy in general. Reinhart & Rogoff (2009, 15) refer to this as the "this-time-is-different syndrome": otherwise competent people seem to miss the obvious signs of a bubble, because they truly believe that the circumstances have changed and that the old rules of valuation no longer apply. Maybe people refuse to consider the inverted yield curve as a reliable and consistent leading indicator because of its simplicity: it does seem rather confusing that a simple indicator could have so much predictive power on both the economy and financial markets. On the other hand, it could also be that people do recognize the predictive power and act accordingly: the inversion of the yield curve can also be a self-fulfilling prophecy that causes economic distress, but this theory has not been really studied.

Even though the inverted yield curve has been shown to have some predictive power during different times, it is not a static phenomenon. In other words, the recessions and stock market crashes following the inversion do share some qualities but differ on another. For example, as presented in the empirical part of the study, the worst performing stock market months occur much later after the inversion than before. The hypothesis of this study is that the delayed effects are because of more active monetary policies: the bigger the interest rate cuts and quantitative easing, the longer it takes to face the inevitable. However, this study is not focused on monetary policies and the topic is left for future research.

As stated before, one of the main objectives was to find new tools to enhance the predictive power of the inverted yield curve and the main method chosen were hidden Markov models, which were applied to the BAA to 10Y credit spread and the S&P 500. The relevant question is whether the applied models add real value or could the same conclusions have been made with common sense: Could it be possible to achieve same results with simple common sense rules? For example, if simplified, *The BEAR* invests in the risk-free rate once the yield curve has inverted and the chosen credit spread widens significantly and it stays out of the stock market until it is sensible to invest in the stock market according to the latest inversion and recession. So, one could ask, is it necessary to use HMMs?

This study argues that HMMs are necessary, since they take out most of the human input that is vulnerable to behavioral biases. For example, what is a significant change in credit spreads? The used HMMs analyze the whole data set, identify different regimes and choose the best fitting one, whereas we as human beings always find ways to make our own interpretations, for better or for worse. In other words, it is impossible for a human being to be completely objective, which is usually problematic in financial markets, which is why HMMs are used in this study.

The conducted empirical study and the proposed strategies were quite simple in their nature and there are obviously a lot of ways to improve the analysis and conduct further research. First, the chosen indicators of market sentiment might not have been the optimal ones in trying to predict future bear markets. Good alternatives or additions could be e.g. search word analysis with Google Trends or similar, NYSE turnover as a measure of stock market volume or VIX as a measure of risk appetite as soon as there is more data available. Plus, there are certainly more indicators or measures that could be combined with the inverted yield curve to improve predictability.

On the technical side, there are also a lot of options for further research. With the modern computational power, one way to possibly improve the strategies could be the use of daily data instead of monthly data. Also, the used HMMs could be constructed on a rolling basis with a fixed data set length, in other words always dropping the last data point with new data points, which could produce different results. Alternatively, using some other method to emphasize newer data or high-order HMMs (see e.g. Zhang et al. 2019) could also improve the signals given by the HMMs.

More specifically, the proposed comprehensive bear market predictor (*The BEAR*) could be improved significantly by rescaling the factors better. For example, average means are used to rescale the S&P 500 HMM, but better results could be presumably achieved with probability distributions, which would give a more detailed picture, that would be easier to scale (e.g. the relative amount of probability mass below zero could be used directly). This was just one suggested option and there should be more ways to improve results. All in all, this study is in no way complete and there is plenty of room for further research.

To summarize, the inverted yield curve seems to have clear implications also for financial markets, even though the connections and causalities may not be as clear as in the GDP-measured real economy. If combined with other factors, the inverted yield curve can be used to predict future bear markets to some extent. The proposed strategies based on the inverted yield curve, other factors and HMMs outperformed the benchmark total stock market return during 1986–2019 in terms of both annualized and risk-adjusted returns, but the true acid test of the proposed strategies will be made during the current inversion in the near future.

Financial markets are characterized by rather long stable periods, which are succeeded by sudden instability and uncertainty that can unwind past returns in matter of months, weeks, days or even hours. In other words, as Myron Scholes puts it (Pedersen 2015, 268): “Finance is in volatility time, not calendar time.” Although greater volatility leads to greater opportunities in the market, the effectiveness of the proposed strategies in this study is based on prudence and patience: the best option for the average investor could very well be to weather the storm in safer assets and then return to the stock market once the volatility is lower again.

REFERENCES

Adrian, Tobias – Estrella, Arturo – Shin, Hyung Son (2019) Risk-taking channel of monetary policy. *Financial management*, 48, 725–738.

Ang, Andrew – Piazzesi, Monika – Wei, Min (2006) What does the yield curve tell us about GDP growth? *Journal of Econometrics*, 131(1), 359–403.

Anh, Nguyen Xuan – Nataraja, Ramesh Mark – Chauhan, Sunita (2020) Towards near real-time assessment of surgical skills: A comparison of feature extraction techniques. *Computer Methods and Programs in Biomedicine*, 187.

Baum, Leonard E. – Petrie, Ted (1966) Statistical Inference for Probabilistic Functions of Finite State Markov Chains. *The Annals of Mathematical Statistics*, 37(6), 1554–1563.

Baum, Leonard E. – Petrie, Ted – Soules, George – Weiss, Norman (1970) A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains. *The Annals of Mathematical Statistics*, 41(1), 164–171.

Bektic´, Demir – Regele, Tobias (2018) Exploiting uncertainty with market timing in corporate bond markets. *Journal of Asset Management*, 19 (2), 79–92.

Bordalo, Pedro – Gennaioli, Nicola – Shleifer, Andrei (2018) Diagnostic Expectations and Credit Cycles. *Journal of Finance*, 73(1), 199–227.

Campbell John Y. – Lo, Andrew W. – MacKinlay, A. Craig (1997) *The econometrics of financial markets*. Princeton University Press; Princeton N.J.

Cappé, Olivier – Moulines, Eric – Rydén, Tobias (2005) *Inference in Hidden Markov Models*. Springer; New York, NY.

Chen, Shiu-Sheng (2009) Predicting the bear stock market: Macroeconomic variables as leading indicators. *Journal of Banking & Finance*, 33, 211–223.

Chicago Board Options Exchange (CBOE) (2020). VIX Index Historical Data. <http://www.cboe.com/products/vix-index-volatility/vix-options-and-futures/vix-index/vix-historical-data>, read 10.1.2020.

Chinn, Menzie – Kucko, Kavan (2015) The Predictive Power of the Yield Curve Across Countries and Time. *International Finance*, 18(2), 129–156.

Choudry, Moorad (2013) *An introduction to value-at-risk*. Wiley; New York. 5th edition.

Choudry, Moorad (2018[2011]) *An introduction to banking: principles, strategy and risk management*. Wiley; Hoboken. N.J. 2nd edition.

Ding, Zhuanxin – Granger, Clive W.J. – Engle, Robert F. (1993) A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1(1), 83–106

Estrella, Arturo – Hardouvelis, Gikas A. (1991) The Term Structure as a Predictor of Real Economic Activity. *The Journal of Finance*, 46 (2), 555–576.

Estrella, Arturo –Mishkin, Frederic S. (1997) The predictive power of the term structure of interest rates in Europe and the United States: Implications for the European Central Bank. *European Economic Review*, 41, 1375–1401.

Estrella, Arturo – Trubin, Mary (2006) The Yield Curve as a Leading Indicator: Some Practical Issues. *Current Issues in Economics and Finance*, 12(5), 1–7.

Federal Reserve Bank of New York (2019a) The Yield Curve as a Leading Indicator. <https://www.newyorkfed.org/medialibrary/media/research/capital_ markets/Prob_Rec.pdf >, read 26.10.2019.

Federal Reserve Bank of New York (2019b) The Yield Curve as a Leading Indicator. <https://www.newyorkfed.org/research/capital_markets/ ycfaq.html#/faq>, read 24.10.2019.

Federal Reserve Bank of New York (2019c) The Yield Curve as a Leading Indicator: Monthly data. <https://www.newyorkfed.org/medialibrary/media/ research/capital_markets/allmonth.xls>, downloaded 26.10.2019.

Federal Reserve Economic Database (2019a) 10-Year Treasury Constant Maturity Rate (DGS10). <https://fred.stlouisfed.org/series/DGS10>, downloaded 13.1.2020.

Federal Reserve Economic Database (2019b) Moody's Seasoned Baa Corporate Bond Yield (BAA). <https://fred.stlouisfed.org/series/BAA>, downloaded 13.1.2020.

Feldman, Todd – Liu, Shuming (2018). A New Predictive Measure Using Agent-Based Behavioral Finance. *Computational Economics,* 51, 941–959.

French, Kenneth R. (2020a) Current Research Returns. <https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html>, visited 15.1.2020.

French, Kenneth R. (2020b) Description of Fama/French Factors. < https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f_factors.html>, visited 15.1.2020.

Friedman, Milton (1970) Controls on Interest Rates Paid by Banks. *Journal of Money, Credit and Banking*, 2 (1), 15–32.

Gogas, Periklis – Papadimitriou, Theophilos – Matthaiou, Maria – Chrysanthidou (2015) Yield Curve and Recession Forecasting in a Machine Learning Framework. *Computational Economics*, Vol.45(4), 635–645.

Greenwood, Robin – Hanson, Samuel G. (2013) Issuer Quality and Corporate Bond Returns. *The Review of Financial Studies*, 26(6), 1483–1525

Guidolin, Massimo – Timmermann, Allan (2007) Asset allocation under multivariate regime switching. *Journal of Economic Dynamics and Control*, Vol.31(11), 3503–3544.

Harte, David (2017) HiddenMarkov: Hidden Markov Models. R package. <https://cran.r-project.org/web/packages/HiddenMarkov/index.html>, visited 7.1.2019.

Harvey, Campbell R. (1986) *Recovering expectations of consumption growth from an equilibrium model of the term structure of interest rates*. Ph.D. dissertation. The University of Chicago, Chicago, Illinois. <https://faculty.fuqua.duke.edu/~charvey/Research/Thesis/Thesis.htm>, read 23.10.2019.

Joshi, J – Tankeshwar, K – Srivastava, Sunita (2017) Hidden Markov Model for quantitative prediction of snowfall and analysis of hazardous snowfall events over Indian Himalaya. *Journal of Earth System Science*, 126(3), 1–12.

Jurafsky, Daniel – Martin, James H. (2019) *Speech and Language Processing.* Draft of October 2, 2019. <https://web.stanford.edu/~jurafsky/slp3/A.pdf>, read 14.2.2020.

Keen, Steve (2013) A monetary Minsky model of the Great Moderation and the Great Recession. *Journal of Economic Behavior & Organization*, 86, 221–235.

Kindleberger, Charles P. – Aliber, Robert Z. (2011) *Manias, panics and crashes: a history of financial crises*. Palgrave Macmillan, Basingstoke.

Kirschenmann, Karolin – Malinen, Tuomas – Nyberg, Henri (2016) The risk of financial crises: Is there a role for income inequality? *Journal of International Money and Finance,* 68, 161–180.

Little, Roderick J. A. – Rubin, Donald B. (2002) *Statistical analysis with missing data*. John Wiley & Sons, Inc.; Hoboken, New Jersey.

López-Salido, David – Stein, Jeremy C. – Zakrajšek, Egon (2017) Credit-Market Sentiment and the Business Cycle, *The Quarterly Journal of Economics*, 132(3), 1373–1426.

Malkiel, Burton G. (2007) *Sattuman kauppaa Wall Streetillä (A Random Walk Down Wall Street: the time tested strategy for successful investing).* Talentum, Helsinki.

Mamon, Rogemar S. (ed.) – Elliott, Robert J. (ed.) (2007) Hidden Markov Models in Finance. *International Series in Operations Research & Management Science*. Springer; Boston, MA.

Mandelbrot, Benoit (1963) The Variation of Certain Speculative Prices. *The Journal of Business*, 36(4), 394–419.

Mishkin, Frederic S. (2019) *The economics of money, banking and financial markets*. Pearson Education, Harlow. 12th edition.

NBER (The National Bureau of Economic Research) (2019) US Business Cycle Expansions and Contractions. <https://www.nber.org/cycles.html>, read 24.10.2019.

Nguyen, Nguyet (2018) Hidden Markov Model for Stock Trading. *International Journal of Financial Studies*, Vol.6(2)

Nyberg, Henri (2013) Predicting bear and bull stock markets with dynamic binary time series models. *Journal of Banking & Finance*, 37, 3351–3363.

Pedersen, Lasse Heje (2015) *Efficiently Inefficient: How Smart Money Invests and Market Prices Are Determined*. Princeton University Press, US.

Phillips, Peter C. B. – Shi, Shuping – Yu, Jun (2015) Testing for Multiple bubbles: Historical episodes of exuberance and collapse in the S&P 500. *International Economic Review*, 56(4), 1043–1078.

Rabiner, Lawrence R. (1989) A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2), 257–286.

Reinhart, Carmen M. – Rogoff, Kenneth S. (2009) *This Time Is Different: Eight Centuries of Financial Folly*. Princeton University Press; Princeton, N.J.

Resnick, Bruce G. – Shoesmith, Gary L. (2002) Using the Yield Curve to Time the Stock Market. *Financial Analysts Journal,* 58 (3), 82–90.

Rudebusch, Glenn D. – Williams, John C. (2009) Forecasting Recessions: The Puzzle of the Enduring Power of the Yield Curve. *Journal of Business & Economic Statistics*, 27(4), 492–503.

Shen, Lucinda (2019) Are We Near a Recession? The Godfather of the Inverted Yield Curve Says It's 'Code Red'. *Fortune.* <https://fortune.com/2019/09/09/recession-stock-inverted-yield-curve-market-great-what-is-economy-us/>, read 17.10.2019.

Shiller, Robert J. (2015 [2000]) *Irrational Exuberance*. Princeton University Press; Princeton, N.J.

Shiller, Robert J. (2020) Online data Robert Shiller. < http://www.econ.yale.edu/~shiller/data.htm>, read 10.2.2020.

Sornette, Didier (2017) *Why stock markets crash: critical events in complex financial systems*. Princeton University Press; Princeton, New Jersey.

Stock, James H. – Watson, Mark W. (1989) New Indexes of Coincident and Leading Economic Indicators. *NBER Macroeconomics Annual*, 4, 351–394.

Tsay, Ruey S. (2005) *Analysis of financial time series*. Wiley; Hoboken, N.J. 2nd ed.

Tsay, Ruey S. – Chen, Rong (2019) *Nonlinear time series analysis*. John Wiley & Sons; Hoboken, NJ.

U.S. Department of the Treasury (2019) Daily Treasury Yield Curve Rates. <https://www.treasury.gov/resource-center/data-chart-center/interest-rates/pages/textview.aspx?data=yield>, read 25.10.2019.

Veronesi, Pietro (2010) *Fixed income securities: valuation, risk and risk management*. Wiley; Hoboken, N.J.

Viterbi, Andrew J. (1967) Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory*, 13(2), 260–269.

Whaley, R.E. (2000) The investor fear gauge: Explication of the CBOE VIX. *Journal of Portfolio Management*, Spring 2000,26(3), 12–17

World Federation of Exchanges (2020) World Federation of Exchanges Statistics Portal. <https://statistics.world-exchanges.org/>, visited 14.1.2020.

Guo, Xin – Lai, Tze Leung – Shek, Howard – Po-Shing Wong, Samuel (2016) *Quantitative Trading: Algorithms, Analytics, Data, Models, Optimization*. E-book, CRC Press.

Zhang, Mengqi – Jiang, Xin – Fang, Zehua – Zeng, Yue – Xu, Ke (2019) High-order Hidden Markov Model for trend prediction in financial time series.

Physica A: Statistical Mechanics and its Applications, 3/2019, Vol.517, C, 1–12.

Zucchini, Walter – MacDonald, Iain L. (2009) *Hidden Markov models for time series: an introduction using R*. CRC Press, Boca Raton.

APPENDICES

Appendix 1: S&P 500 HMM signals (weighted averages)

Colum	Date	Inversi	SP500	Prob1	Prob2	Prob3	Prob4	Mean1	Mean2	Mean3	Mean4	Expected mean
224	1966:08	7	-8,09 %	0,16	0,84	0,00	0,00	-0,0653	0,0026	0,0187	0,0714	-0,0083
259	1969:07	7	-6,21 %	0,31	0,69	0,00	0,00	-0,0647	0,0034	0,0189	0,0725	-0,0178
265	1970:01	13	-7,96 %	0,12	0,88	0,00	0,00	-0,0655	0,0029	0,0187	0,0724	-0,0056
268	1970:04	16	-9,48 %	0,12	0,88	0,00	0,00	-0,0660	0,0029	0,0185	0,0719	-0,0055
269	1970:05	17	-6,29 %	0,58	0,42	0,00	0,00	-0,0678	0,0030	0,0185	0,0717	-0,0384
270	1970:06	18	-5,13 %	0,85	0,15	0,00	0,00	-0,0666	0,0034	0,0185	0,0717	-0,0559
281	1971:05	29	-4,24 %	0,01	0,99	0,00	0,00	-0,0668	-0,0062	0,0279	0,0726	-0,0067
311	1973:11	5	-12,09 %	0,21	0,79	0,00	0,00	-0,0675	0,0046	0,0178	0,0705	-0,0108
317	1974:05	11	-3,41 %	0,02	0,97	0,01	0,00	-0,0664	-0,0017	0,0207	0,0733	-0,0025
318	1974:06	12	-1,48 %	0,00	0,95	0,05	0,00	-0,0665	-0,0017	0,0206	0,0733	-0,0006
319	1974:07	13	-8,10 %	0,15	0,85	0,00	0,00	-0,0666	-0,0026	0,0203	0,0734	-0,0123
320	1974:08	14	-9,46 %	0,80	0,20	0,00	0,00	-0,0701	-0,0025	0,0203	0,0733	-0,0569
321	1974:09	15	-12,71 %	1,00	0,00	0,00	0,00	-0,0762	-0,0063	0,0157	0,0444	-0,0762
323	1974:11	17	-5,46 %	0,00	0,00	0,00	1,00	-0,0673	0,0055	0,0183	-0,0001	-0,0002
331	1975:07	25	-7,01 %	0,32	0,61	0,00	0,06	-0,0176	0,0080	0,0314	0,0000	-0,0008
332	1975:08	26	-2,13 %	0,34	0,59	0,00	0,06	-0,0176	0,0080	0,0315	-0,0002	-0,0013
333	1975:09	27	-3,53 %	0,11	0,78	0,00	0,12	-0,0450	0,0033	0,0394	0,0126	-0,0007
382	1979:10	11	-7,11 %	0,20	0,77	0,00	0,03	-0,0727	0,0042	0,0168	0,0130	-0,0105
387	1980:03	16	-10,74 %	0,51	0,31	0,00	0,18	-0,0748	0,0043	0,0166	0,0129	-0,0345
404	1981:08	10	-6,41 %	0,21	0,79	0,00	0,00	-0,0766	0,0066	0,0398	0,0212	-0,0111
405	1981:09	11	-5,53 %	0,41	0,57	0,00	0,02	-0,0753	0,0065	0,0397	0,0201	-0,0267
410	1982:02	16	-6,25 %	0,24	0,76	0,00	0,00	-0,0734	0,0065	0,0406	0,0181	-0,0122
505	1990:01	7	-7,13 %	0,15	0,82	0,00	0,03	-0,0680	0,0047	0,0401	0,0176	-0,0062
512	1990:08	14	-9,91 %	0,50	0,33	0,00	0,18	-0,0683	0,0050	0,0401	0,0165	-0,0294
513	1990:09	15	-5,25 %	0,39	0,24	0,00	0,37	-0,0681	0,0050	0,0401	0,0162	-0,0194
635	2000:11	4	-8,35 %	0,33	0,65	0,00	0,02	-0,0755	0,0081	0,0481	0,0234	-0,0190
638	2001:02	7	-9,68 %	0,47	0,45	0,00	0,08	-0,0758	0,0080	0,0479	0,0210	-0,0302
639	2001:03	8	-6,64 %	0,65	0,21	0,00	0,14	-0,0748	0,0080	0,0478	0,0208	-0,0443
644	2001:08	13	-6,63 %	0,13	0,84	0,00	0,02	-0,0747	0,0079	0,0494	0,0220	-0,0029
645	2001:09	14	-8,53 %	0,65	0,29	0,00	0,07	-0,0733	0,0079	0,0490	0,0207	-0,0437
652	2002:04	21	-6,34 %	0,12	0,88	0,00	0,01	-0,0726	0,0078	0,0481	0,0214	-0,0015
654	2002:06	23	-7,52 %	0,28	0,70	0,00	0,02	-0,0720	0,0077	0,0481	0,0213	-0,0144
655	2002:07	24	-8,23 %	0,77	0,17	0,00	0,06	-0,0700	0,0078	0,0476	0,0214	-0,0513
656	2002:08	25	0,49 %	0,21	0,47	0,16	0,16	-0,0695	0,0076	0,0467	0,0197	-0,0001
657	2002:09	26	-11,66 %	0,83	0,01	0,00	0,16	-0,0644	0,0079	0,0481	0,0150	-0,0509
726	2008:06	22	-8,99 %	0,06	0,24	0,00	0,70	-0,1235	0,0088	0,0470	-0,0075	-0,0110
727	2008:07	23	-0,99 %	0,00	0,33	0,00	0,66	-0,1253	0,0088	0,0470	-0,0074	-0,0025
729	2008:09	25	-9,52 %	0,01	0,03	0,00	0,96	-0,1264	0,0089	0,0469	-0,0078	-0,0089
730	2008:10	26	-18,56 %	0,73	0,00	0,00	0,27	-0,1230	0,0088	0,0480	-0,0092	-0,0921
731	2008:11	27	-7,78 %	0,36	0,00	0,00	0,64	-0,1115	0,0087	0,0484	-0,0119	-0,0475
732	2008:12	28	0,78 %	0,01	0,00	0,05	0,94	-0,1195	0,0089	0,0476	-0,0123	-0,0104
733	2009:01	29	-8,95 %	1,00	0,00	0,00	0,00	-0,0256	0,0081	0,0379	0,0784	-0,0256
860	2019:08	2	-1,83 %	0,79	0,17	0,00	0,04	-0,0012	0,0106	0,0546	-0,0339	-0,0004

Appendix 2: S&P 500 HMM signals (Viterbi)

Colum	Date	Inversi	SP500	Viterbi	Mean1	Mean2	Mean3	Mean4	Negative Viterbi
269	1970:05	17	-6,29 %	1	-0,0678	0,0030	0,0185	0,0717	1
270	1970:06	18	-5,13 %	1	-0,0666	0,0034	0,0185	0,0717	1
281	1971:05	29	-4,24 %	2	-0,0668	-0,0062	0,0279	0,0726	1
306	1973:06	0	-0,66 %	2	-0,0669	-0,0039	0,0274	0,0737	1
307	1973:07	1	3,73 %	2	-0,0667	-0,0013	0,0234	0,0734	1
317	1974:05	11	-3,41 %	2	-0,0664	-0,0017	0,0207	0,0733	1
318	1974:06	12	-1,48 %	2	-0,0665	-0,0017	0,0206	0,0733	1
319	1974:07	13	-8,10 %	2	-0,0666	-0,0026	0,0203	0,0734	1
320	1974:08	14	-9,46 %	1	-0,0701	-0,0025	0,0203	0,0733	1
321	1974:09	15	-12,71 %	1	-0,0762	-0,0063	0,0157	0,0444	1
323	1974:11	17	-5,46 %	4	-0,0673	0,0055	0,0183	-0,0001	1
387	1980:03	16	-10,74 %	1	-0,0748	0,0043	0,0166	0,0129	1
512	1990:08	14	-9,91 %	1	-0,0683	0,0050	0,0401	0,0165	1
513	1990:09	15	-5,25 %	1	-0,0681	0,0050	0,0401	0,0162	1
639	2001:03	8	-6,64 %	1	-0,0748	0,0080	0,0478	0,0208	1
645	2001:09	14	-8,53 %	1	-0,0733	0,0079	0,0490	0,0207	1
655	2002:07	24	-8,23 %	1	-0,0700	0,0078	0,0476	0,0214	1
657	2002:09	26	-11,66 %	1	-0,0644	0,0079	0,0481	0,0150	1
729	2008:09	25	-9,52 %	4	-0,1264	0,0089	0,0469	-0,0078	1
730	2008:10	26	-18,56 %	1	-0,1230	0,0088	0,0480	-0,0092	1
731	2008:11	27	-7,78 %	1	-0,1115	0,0087	0,0484	-0,0119	1
732	2008:12	28	0,78 %	4	-0,1195	0,0089	0,0476	-0,0123	1
733	2009:01	29	-8,95 %	1	-0,0256	0,0081	0,0379	0,0784	1
859	2019:07	1	1,30 %	1	-0,0013	0,0106	0,0544	-0,0339	1
860	2019:08	2	-1,83 %	1	-0,0012	0,0106	0,0546	-0,0339	1
861	2019:09	3	1,70 %	1	-0,0010	0,0106	0,0547	-0,0340	1

Appendix 3: Credit Spread HMM signals

Date	RM	1	2	3	4	Viterbi	mean1	mean2	mean3	mean4
1969:10	0,233378	0,09	0,00	0,00	0,91	4	-0,022	-0,027	0,037	0,072
1973:09	0,313464	0,00	0,00	0,00	1,00	4	-0,043	-0,017	0,053	0,085
1980:02	-0,33669	0,14	0,00	0,00	0,86	4	-0,006	-0,020	0,040	0,052
1980:12	0,160588	0,43	0,00	0,00	0,57	1	-0,020	-0,016	0,068	0,078
1981:06	0,229436	0,35	0,00	0,00	0,65	4	-0,023	-0,016	0,068	0,078
1989:06	0,090601	0,03	0,10	0,85	0,02	3	-0,026	-0,030	0,084	0,046
2001:09	0,13605	0,25	0,17	0,56	0,02	3	-0,016	-0,005	0,115	0,044
2007:08	0,182504	0,00	0,01	0,90	0,08	3	-0,006	-0,009	0,097	0,041

Appendix 4: S&P 500 HMM

```
####
## S&P 500 HMM Dynamic
library(tidyverse)
library(moments)
library(readxl)
library(xlsx)
library(HiddenMarkov)


###-- four-regime --###
SP500history <- read_excel("SP5001928-.xlsx",
                    col_types = c("date", "numeric", "numeric", "numeric", "numeric"))

aloitus <-"1959-01-01"
dates <- cbind.data.frame(seq(as.Date("1928/1/1"), by = "month", to =
as.Date("2019/12/1")),SP500history$Date)
dates <- filter(dates, SP500history$Date>=aloitus)
df1 <- data.frame(Date = double()
            ,RM = double()
            ,State1Prob = double()
            ,State2Prob = double()
            ,State3Prob = double()
            ,State4Prob = double()
            ,ViterbiPath = double()
            ,State1Mean = double()
            ,State2Mean = double()
            ,State3Mean = double()
            ,State4Mean = double()
            ,State1Sd = double()
            ,State2Sd = double()
            ,State3Sd = double()
            ,State4Sd = double())

## ------ setting initial values
Pi <- diag(rep(0.8,4))+matrix(rep(1,16)/4*0.2,ncol=4,nrow=4)
delta <- rep(1,4)/4
dist <- 'norm'
```

```
pm <- list(mean=c(-0.1,0,0.05,0.1),sd=c(0.1,0.05,0.05,0.1))

# Skip 10/1987 (model results in an error)
virhe1 <- grep("1987-10-01",SP500history$Date)-grep(aloitus,SP500history$Date)
loppu <- nrow(dates)

# HMM estimations one window at a time
for (i in append(1:virhe1, (virhe1+2):loppu)) {

    SP500history1 <- filter(SP500history, Date>="1948-01-01", Date<=dates[i,2])
    RM <- as.vector(unlist(SP500history1$Diff))
    RM <- log(RM+1)
    tt <- seq(as.Date("1948-01-01"), by = "month", to = as.Date(dates[i,1]))

    ## HMM modeling
    require("HiddenMarkov")
    temp <- dthmm(RM,Pi,delta,dist,pm)   #setting model
    out4 <- BaumWelch(temp,control=bwcontrol(prt=FALSE))  #Baum Welch
    estimation

    v4 <- Viterbi(out4)   # most likely path: Viterbi
    pp <- 1:length(RM); for(i in 1:length(RM))pp[i] <- out4$u[i,v4[i]]
    # obtain smoothed probability of the Viterbi states

    # Data frame including relevant data
    haarukka<-cbind(SP500history1[,1], RM, out4$u, v4)
    rivi <- cbind.data.frame(haarukka[length(haarukka[,1]),], out4$pm$mean[1],
    out4$pm$mean[2], out4$pm$mean[3], out4$pm$mean[4],
    out4$pm$sd[1],out4$pm$sd[2], out4$pm$sd[3], out4$pm$sd[4])
    df1 <- rbind.data.frame(df1,rivi)

}
summary(out4)

# add error row (10/1987)
df1<-rbind.data.frame(df1[1:virhe1,],c(haarukka[grep("1987-10-01",haa-
rukka$Date),],df1[grep("1987-09-01",df1$Date),-(1:7)]),df1[-(1:virhe1),])
write.xlsx(df1,"4state_S&P500_1948.xlsx")
```

```
# State probabilities
tt <- seq(as.Date("1959/1/1"), by = "month", to = as.Date("2019/12/1"))
par(mfrow=c(4,1))
plot(tt,df1$`1`,type='l',xlab='Time',ylim=c(0,1),ylab='Prob',main='State 1 Prob')
plot(tt,df1$`2`,type='l',xlab='Time',ylim=c(0,1),ylab='Prob',main='State 2 Prob')
plot(tt,df1$`3`,type='l',xlab='Time',ylim=c(0,1),ylab='Prob',main='State 3 Prob')
plot(tt,df1$`4`,type='l',xlab='Time',ylim=c(0,1),ylab='Prob',main='State 4 Prob')
par(mfrow=c(1,1))

logLik(out4) # log likelihood value

par(mfrow=c(2,2))
res4 <- residuals(out4)  # get residuals
pacf(res4**2,main='PACF residual squared')    # pacf residual squared
pacf(RM**2, main='PACF Market return squared')    # pacf sp return squared
qqnorm(res4, main='residual')    # qq plot residual
qqnorm(RM,main='Market return')      # qq plot sp return

skewness(RM)
kurtosis(RM)
```

Appendix 5: Credit spread HMM

```
####
##  Credit Spread HMM dynamic
library(tidyverse)
library(moments)
library(readxl)
library(xlsx)
library(HiddenMarkov)


###-- four-regime --###
SP500history <- read_excel("SP5001928-.xlsx",
                    col_types = c("date", "numeric", "numeric", "numeric", "numeric"))

aloitus <-"1968-01-01"
```

```
dates <- cbind.data.frame(seq(as.Date("1928/1/1"), by = "month", to =
as.Date("2019/12/1")),SP500history$Date)
dates <- filter(dates, SP500history$Date>=aloitus)
df1 <- data.frame(Date = double()
        ,CS = double()
        ,State1Prob = double()
        ,State2Prob = double()
        ,State3Prob = double()
        ,State4Prob = double()
        ,ViterbiPath = double()
        ,State1Mean = double()
        ,State2Mean = double()
        ,State3Mean = double()
        ,State4Mean = double()
        ,State1Sd = double()
        ,State2Sd = double()
        ,State3Sd = double()
        ,State4Sd = double())

## ------ setting initial values
Pi <- diag(rep(0.8,4))+matrix(rep(1,16)/4*0.2,ncol=4,nrow=4)
delta <- rep(1,4)/4
dist <- 'norm'
pm <- list(mean=c(-0.1,0,0.05,0.1),sd=c(0.1,0.05,0.05,0.1))

loppu <- nrow(dates)
for (i in 1:loppu) {

    SP500history1 <- filter(SP500history, Date>="1962-01-01", Date<=dates[i,2])
    CS <- as.vector(unlist(SP500history1$Diff.BAAto10YSpread))
    CS <- log(CS+1)
    tt <- seq(as.Date("1962-01-01"), by = "month", to = as.Date(dates[i,1]))
    ## HMM modeling
    require("HiddenMarkov")
    temp <- dthmm(CS,Pi,delta,dist,pm)   #setting model
    out4 <- BaumWelch(temp,control=bwcontrol(prt=FALSE))   #Baum Welch
    estimation

    v4 <- Viterbi(out4)   # most likely path: Viterbi
```

```
        pp <- 1:length(CS); for(i in 1:length(CS))pp[i] <- out4$u[i,v4[i]]
        # obtain smoothed probability of the Viterbi states

        # Data frame including relevant data
        haarukka<-cbind(SP500history1[,1], CS, out4$u, v4)
        rivi  <-  cbind.data.frame(haarukka[length(haarukka[,1]),],  out4$pm$mean[1],
        out4$pm$mean[2], out4$pm$mean[3], out4$pm$mean[4],
        out4$pm$sd[1],out4$pm$sd[2], out4$pm$sd[3], out4$pm$sd[4])
        df1 <- rbind.data.frame(df1,rivi)

}
summary(out4)

write.xlsx(df1,"4state_CS_1968.xlsx")

# State probabilities
tt <- seq(as.Date("1968/1/1"), by = "month", to = as.Date("2019/12/1"))
par(mfrow=c(4,1))
plot(tt,df1$`1`,type='l',xlab='time',ylim=c(0,1),ylab='prob',main='State 1 Prob')
plot(tt,df1$`2`,type='l',xlab='time',ylim=c(0,1),ylab='prob',main='State 2 Prob')
plot(tt,df1$`3`,type='l',xlab='time',ylim=c(0,1),ylab='prob',main='State 3 Prob')
plot(tt,df1$`4`,type='l',xlab='time',ylim=c(0,1),ylab='prob',main='State 4 Prob')
par(mfrow=c(1,1))

logLik(out4) # log likelihood value

par(mfrow=c(2,2))
res4 <- residuals(out4)  # get residuals
pacf(res4**2,main='PACF Residual squared')   # pacf residual squared
pacf(CS**2, main='PACF Credit Spread Diff. squared')   # pacf sp return squared
qqnorm(res4, main='Residual')   # qq plot residual
qqnorm(CS,main='Credit Spread Diff.')     # qq plot sp return

skewness(CS)
kurtosis(CS)
```