ERASMUS UNIVERSITY ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS
BSc Economics & Business Economics
Major Financial Economics

Forecasting Economic Growth and Predicting Recessions:

An analysis of yield spread, consumer confidence, and news sentiment

Author: J.H. Kranenbarg
Student number: 447890
Thesis supervisor: dr. J.J.G. Lemmen
Second assessor: dr. N.L. van der Sar
Finish date: June 28, 2019

PREFACE AND ACKNOWLEDGEMENTS

I would like to express my gratitude to everyone who supported me throughout the last chapter of my studies. I am proud to have finished within three years, and this Bachelor's thesis is the crown on my degree. Writing the thesis was a very intense, but enlightening process, in which I have learned to properly structure an analysis. I hope the thesis is enjoyable for the reader, and the results are useful to whomever it may concern. A special thanks to my thesis supervisor dr. J.J.G. Lemmen, who provided valuable feedback and guided me in the right direction whenever necessary. Finally, I would like to thank my family and friends who always stood by me throughout my academic career.

ABSTRACT

This Bachelor's thesis provides a structured comparison of the indicators yield spread, consumer confidence and news sentiment in their usefulness to forecast economic growth and predict recessions in the United States and Germany. The variables are tested for normality, unit roots and cointegration. For economic growth forecasts, simple regressions and vector autoregression models are used. To predict recessions, Probit models are fitted. The in-sample fit is evaluated using AIC and out-of-sample forecasts are made and evaluated using RMSE. Which combination of variables performs best is highly reliant on the country of interest and the model estimated.

Keywords:
recession forecasts, yield spread, consumer confidence, news sentiment, Probit model

JEL Classification:
E37, E43, E21, G12, C25

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 Introduction

Many studies document the predictive power of the slope of the Treasury yield curve for forecasting economic growth and predicting recessions (e.g., Kessel, 1965; Fama, 1986; Harvey, 1988; Estrella and Hardouvelis, 1991; Estrella and Mishkin, 1996; Wright, 2006). The yield curve shows the interest rates for different maturities of debt instruments. It is generally upward-sloping, which means that the yield increases as the time to maturity rises. The reason for this is that investors expect future economic growth, leading to inflation. This implies a greater risk with maturity, and thus a higher rate of return is required. However, sometimes the yield curve inverts, meaning interest rates on long-term debt become lower than those on short-term debt. Since 1970, every time this happened, a recession followed (Carrera, 2019). On March 22, 2019, the 10-year Treasury yield fell below the three month T-bill yield for the first time since the Great Recession of 2008-2009 (Wigglesworth & Rennison, 2019). Does this mean the next recession is around the corner?

Besides the yield curve, many other economic variables have been investigated to improve recession forecasts. Especially after the 1990 recession, when the predictive power of the yield curve seemed to decline, researchers started to innovate. One variable that was mentioned a lot and seems to be effective is consumer confidence. This indicator measures how optimistic the population is about the state of the economy. Logically, this increases when the economy is doing well and decreases when it is not. If the economy lags behind consumer confidence and not vice versa, it could be a very useful predictor of recessions. Furthermore, the predictive power of the news will also be investigated in this research to see if it adds to the accuracy of the forecasts. Therefore, the research question will be:

Which combination of yield spread, consumer confidence, and news sentiment produces the most accurate forecasts of economic growth and recession predictions in the US and Germany?

This paper's contribution to existing literature is threefold. First of all, this paper provides a structured comparison of the indicators yield spread, consumer confidence and news sentiment, as well as a combination of the three, in their usefulness to forecast economic growth and predict recessions, using a single dataset and framework. Secondly, contrary to previous studies, this Bachelor's thesis uses a longer time horizon and therefore includes the 2007 recession. Finally, the direct predictive power of news sentiment for forecasting real GDP and the likelihood of a recession has not been studied before.

To answer the research question, time series analysis will be conducted for the period 1977 until 2018. For the forecasts of economic growth, this will be done using simple regressions and vector autoregression (VAR) models. To predict the likelihood of a recession, the Probit model will be used. Forecasts will be made for 1 to 8 quarters ahead. The models will be assessed based on both in-sample fit and out-of-sample performance.

Which combination of variables predicts most accurately varies between the models, the countries, and whether in-sample fit or out-of-sample performance is evaluated. In general, most models that perform well far ahead include at least yield spread, and there are several cases where including consumer confidence and/or news sentiment improves the model. Furthermore, the forecasting performance of the models is much better for the United States than for Germany.

This paper commences with an extensive review of the existing literature, together with the proposed hypotheses, in the Theoretical Framework. Next, in the Data section, the data sources are mentioned, descriptive statistics are given and the data is tested for normality. The Methodology section consists of an explanation of the preliminary tests, the statistical methods that will be used, and how the results will be evaluated. These results are then presented and evaluated in the Results section, after which the research question will be answered in the Conclusion. This section also contains the limitations of the research and recommendations for further research.

CHAPTER 2 Theoretical Framework

The aim of this Theoretical Framework is to present previous theories considering the predictive power of the variables. An overview can be found in table 1 at the end of this section. Furthermore, the hypotheses to be studied in this thesis are put forward. The section consists of five subsections. In the first subsection, the main findings of previous literature regarding usefulness of yield spread for forecasting economic growth and predicting recessions are summarized. Next, research on the consumer confidence index and news sentiment is summarized. In the last two sections, different theoretical models are analysed and the hypotheses are mentioned.

2.1 Yield spread

Since the late 1980s, numerous studies on the usefulness of the yield curve for forecasting future economic activity and predicting recessions have been conducted. A vast number of these studies has focused on the predictive power of the yield spread, which is the spread between the interest rates on the ten-year Treasury note and the three-month Treasury bill. The literature can be split into studies that focus on the forecasting of economic growth and those that try to predict recessions directly.

2.1.1 Forecasting economic growth

Kessel (1965) was the first to describe how the term structure of interest rates varies with the business cycle. Fama (1986) found that forward rates increase before economic expansions and decrease before recessions. However, he does not perform a detailed statistical analysis. Laurent (1988) and Harvey (1988) analyse the data more thoroughly. They regress the percentage change in real GNP on lags of the spread between a longer-term rate and the federal funds rate and real consumption growth on the expected real yield spread, respectively. Both authors conclude that the term structure contains informational content to predict future growth.

Estrella and Hardouvelis (1991) stop to think whether this information adds to what is available from other statistics, to assess whether it should be part of the leading indicators used for monetary policy. They conclude that to assess whether the yield curve will be useful to the monetary authorities and private forecasters in the future, one must determine the extent to which historical correlations found are specific to the sample period or are crucial elements in the process of agents when making decisions about the future.

Haubrich and Dombrosky (1996) do this by using out-of-sample forecasts. They find that over the time period from 1985 until 1995, there is a decline in how well the yield curve predicts real growth. This is caused by a change in the relationship between the yield curve and real economic activity. Dotsey (1998) also addresses the fact that the yield curve failed in predicting the recession of 1990-1991, but he cannot conclude whether the change in predictive power is temporary or permanent.

All of the researchers mentioned thus far have primarily focused on the U.S. Hu (1993) analysed the Group of Seven (G-7) industrial countries. He compares the yield spread model to a stock price model and a univariate time forecasting model, and finds that the yield spread model is superior for most countries. Plosser and Rouwenhorst (1994) study three industrialized countries, the US, UK and Germany. They also consider the relationship between growth in one country and term spread in another, and find that the term structures of a foreign country often predicts domestic real economic growth rates very well.

Davis and Fagan (1997) focused on forecasting output growth in nine countries of the European Union and Duarte, Venetis & Paya (2004), too, examine the relationship in the Euro area. However, they find contrasting results. Whereas Davis and Fagan (1997) find that term spread is not a good indicator of future output growth, Duarte et al. (2004) conclude the opposite.

To summarize, the existing literature has not managed to reach a general consensus concerning the usefulness of the yield spread in forecasting economic growth, especially over time and across countries.

2.1.2 Predicting recessions

Rather than forecasting economic growth, more recently, most studies have focused on the direct prediction of recessions, as pioneered by Stock and Watson (1993). They created a binary variable indicating whether the economy is in a recession (1) or not (0). Their model, too, failed to predict the recession starting in 1990. Estrella and Mishkin (1996) find that the yield curve spread does better than the Stock-Watson index to predict this particular recession.

Bernard and Gerlach (1998) extend the analysis by using quarterly data for eight countries – Belgium, Canada, France, Germany, Japan, the Netherlands, the United Kingdom and the United States. They find that though the relationship exists in all eight countries, the evidence is stronger for Canada, Germany and the United States than for the other five. Furthermore, the US and German spreads are significant for predicting recessions in the other countries as well, especially in the UK and Japan.

Moneta (2005) focuses on the Euro area specifically, and tests ten different spreads (combinations of the 3-month, 1-year, 2-year, 5-year and 10-year interest rates). Conform previous literature, the yield spread between the 10-year government bond rate and the 3-month interbank rate has the largest predictive power for recessions in the Euro area.

Wright (2006) examines different Probit models, taking into account more information the yield curve provides than simply the term spread. He finds that a model that includes the level of the federal funds rate in addition to the term spread is superior to the simple model with term spread alone.

2.2 Consumer confidence

With the decline in the ability of the yield spread to predict recessions, more recent studies have focused on additional variables that could increase the performance of the model. Many economists believe that pessimism amongst consumers was one of the causes of the 1990-91 recession (Walsh, 1993). This led to researchers considering the Consumer Confidence Index (CCI) as an indicator of recessions. When plotting the CCI against recessions after the Second World War, it turns out before all recessions, there was a decline in consumer confidence.

Garner (1991) argues that the CCI is not a reliable indicator by itself, and does not complement other economic variables either in forecasting, under normal circumstances. However, in exceptional cases such as abrupt changes in the economy, consumer confidence may add value in predictions. Therefore, it is certainly useful to consider it for predicting recessions.

Fuhrer (1993) asks the important question whether consumer sentiment foreshadows the future, and helps predict consumption growth. Using simple forecasting regressions, he finds that over the past 30 years, consumer sentiment has systematically been able to predict the macro economy. Matsusaka and Sbordonne (1995) use vector autoregressions to study the relationship between consumer confidence and fluctuations in the economy, and find that consumer sentiment Granger-causes GNP.

Batchelor and Dua (1997) do not assume the economy moves cyclically, but instead assume it switches between up and down states. They find that the CCI is useful in predicting these switching points, especially in the recovery from the 1990 recession. Batchelor (2001) uses a Markov Switching Model to establish the relationship between consumer and business confidence and the likelihood of a recession. However, the results of this study are mixed.

Howrey (2001) uses consumer confidence in combination with other variables, amongst which the yield spread, to see whether it augments the predictive power. Using the quadratic probability score (QPS) to evaluate the forecasts, out of all equations, the best one is the equation that includes the Index of Consumer Sentiment (ICS) and the interest rate spread.

Mourougane and Roma (2002) investigate whether the relationship holds in Europe as well, specifically in Belgium, Spain, Germany, France, Italy and the Netherlands. In all countries but Spain, consumer confidence is a good indicator for real GDP growth in the short run.

A very recent study on the predictive power of consumer confidence concludes that even though the CCI can be considered a reasonable predictor of economic growth, there are some limitations regarding short-term deviations (Mazurek & Mielcová, 2017). Therefore, they suggest combining the CCI with additional variables to improve the forecasts.

2.3 News sentiment

In 1990, the Washington Post published an article in which the author asks the question whether the economy is hurt by the so-called "media malady", meaning negative news articles speed up the economy going into recession (Kurtz, 1990). The fact that news can have a significant impact on stock prices has already been established (e.g. Niederhoffer, 1971; Chan, 2003; Davis, Piger & Sedor, 2006; Tetlock, 2007; Fang & Peress, 2008; Garcia, 2013). However, the predictive power of news sentiment regarding recessions directly has not yet been researched.

Several researchers have tried to find a link between news articles and the state of the economy, though. Blood and Philips (1995) examine the relationships amongst economic news coverage, consumer sentiment, the state of the economy and presidential popularity. Regarding the link between economic news and the state of the economy, they find some evidence of cointegration, but it is not very strong. Wu, Stevenson, Chen and Güner (2002) use a vector autoregression to find the association between recession news and the state of the economy. When the entire period (1987 – 1996) is considered, news does not significantly affect the economy. However, when looking at 1987 until 1990 separately, the period in which the economy was in a downturn, a causal effect is found. Hester and Gibson (2003) analyse print news from *The New York Times* and broadcast news from *ABC World News Tonight* to find the consequences of their coverage on the economy. Although the evidence is not very strong, their findings do suggest there is an influence of news on the economy.

A recently published article takes a different perspective by studying the impact of news sentiment on the yield curve (Gotthelf & Uhl, 2018). This article does thereby focus on the predictive power of news sentiment, but not for a recession directly. However, the findings are promising for this thesis too, since there seems to be informational content which is not captured by the traditional yield curve factors.

2.4 Theoretical models

To be able to use a model to predict recessions, it is first necessary to establish the exact definition of a recession. A recession is commonly understood to be a negative GDP growth rate, two quarters in a row. The official dates of the beginnings and ends of U.S. recessions are given by the Business Cycle Dating Committee of the National Bureau of Economic Research (NBER). They believe GDP is a too narrow measure, so they consider a recession as "*a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales*" (Hall et al., 2003, p. 1). The Centre for Economic Policy Research (CEPR) adopts a similar definition for the Euro area. However, Germany might be in a recession when the Euro area is not, or vice versa. Therefore, the general definition is used instead, means a recession implies a consecutive 2-quarter decline in GDP.

Filardo (1999) describes five popular business cycle models: simple rules of thumb using the Conference Board's composite index of leading indicators (CLI), the Neftçi's model (Neftici, 1982), a Probit model, a GDP forecasting model, and the Stock-Watson model. The Neftçi's model improves on the CLI rules of thumb by incorporating past information better, thus being able to signal a recession sooner. Furthermore, different beliefs can be incorporated into the model and additional variables can be included. In turn, the Probit model has advantages over the Neftçi's model. First of all, several indicators can be included at the same time. Secondly, the model yields more precise forecasts. Furthermore, different indicators can be used for different time periods. Therefore, the Probit model has several advantages over the CLI rules of thumb and the Neftçi model. It does not make sense then to study the first two, so they will be disregarded. The Probit model provides a probability of recession. The closer this probability is to 1, the more likely a recession will happen, and the closer to 0, the more unlikely. This is very useful to signal the turning points in the economy as well, which should happen when the probability is close to 0.5.

Rather than estimating only the probability of a recession, it might also be useful for governments to forecast the level of real GDP, which is needed for preparing the budget. The GDP forecasting model and Stock-Watson model can be used in this case. The models are similar except the second uses a broader measure of economic activity rather than GDP. Even though this aligns better with the definition of the NBER of a recession, the model is a bit too sophisticated for this thesis so the GDP forecasting model will be used instead. Besides, this broader measure is not what the government needs when setting the budget either. Therefore, to forecast economic growth, the GDP forecasting model will be used.

To sum up, the ability of the variables to forecast economic growth and predict recessions will be compared using two models, the Probit model and the GDP forecasting model. The specifics of the models can be found in the Methodology section.

2.5 Hypotheses

Based on the previously discussed research and theory, the following specific hypotheses are proposed:

H1: *On a stand-alone basis, yield spread is the strongest predictor for recessions.*

H2: *Adding consumer confidence and news sentiment to the models significantly improves the forecasting power.*

H3: *The forecasting models perform better for the United States than for Germany.*

Table 1. Overview of previous research

Author	Period	Region	Method	Variable(s) used	Relevant Results
Laurent (1988)	1961-1986	US	• Regression	• Fed funds rate • Real fed funds rate • Real M2 • Yield spread	• The spread between a long-term bond rate and the federal funds rate gives the best forecasts
Harvey (1988)	1953-1987	US	• IMA(1,1) • Regression • GMM estimation	• Expected real yield spread • Expected ST interest rate • Lagged consumption growth • Lagged stock returns	• Real term structure contains information about future consumption growth • Interest rate variables forecast better than lagged consumption growth and lagged stock returns
Estrella & Hardouvelis (1991)	1955-1988	US	• Regression • Probit	• Yield spread • Lagged output growth • Lagged inflation • Index of leading indicators • Level of real ST interest rates	• Yield spread has predictive power for cumulative changes in real GNP for 4 years into the future • Yield spread has predictive power for marginal changes in real GNP for 1.5 years into the future • Yield spread has additional predictive power above the other variables
Hu (1993)	1957-1991	G-7	• Regression	• Yield spread • Stock price • Lagged output growth • Inflation	• Yield spread forecasts better than changes in stock prices • Yield spread has additional predictive power above the other variables
Plosser & Rouwenhorst (1994)	1957-1991	US, UK & DE	• Regression	• Yield spread • M1	• In the US and Germany, yield spread has predictive power for long-term economic growth • In the UK, the predictive power of yield spread is weak
Haubrich & Dombrosky (1996)	1961-1995	US	• Regression	• Yield spread	• Over the last three decades, yield spread is one of the best predictors of real growth • Over the last ten years, yield spread performed worse than other forecasting models • The relationship between yield spread and real economic activity has changed
Davis & Fagan (1997)	1970-1992	BE, DK, DE, ES, FR, IR, IT, NL & UK	• VAR	• Slope of the yield curve • Reverse yield gap/stock price • Credit quality spread • Foreign bond yield differential	• Yield spread is not predictive of output growth and inflation in EU countries
Dotsey (1998)	1955-1997	US	• Regression • VAR • Probit	• Yield spread • Monetary tightening • Lagged output growth • Lagged T-bill rate	• Yield spread contains useful information • Recently, yield spread has become less informative

Duarte, Venetis & Paya (2004)	1970-2000	Euro area	• Regression • SC model • Threshold model • Probit	• EMU yield spread • US yield spread • Lagged real growth	• EMU yield spread is a good indicator of future output growth and recessions in the Euro area • There is a structural break in the model in 1992 • US yield spread is also a useful indicator
Stock & Watson (1993)	1959-1988	US	• DSIM	• Leading indicators • Coincident variables	• Selecting the right indicators to include in the model is difficult • The chosen indicators failed to predict the recession of 1990
Estrella & Mishkin (1996)	1960-1995	US	• Probit	• Yield spread	• Yield spread is useful in macroeconomic prediction, especially with longer lead times
Bernard and Gerlach (1998)	1972 - 1993	BE, CA, FR, DE, JP, NL, UK & US	• Probit	• Domestic yield spread • US yield spread • German yield spread • Leading indicators	• Domestic yield spread provides information about the likelihood of a recession in all countries • In some countries, the yield spread is useful for predicting 6-8 quarters ahead • US and German yield spread add limited information • Leading indicators are only useful for predicting in the near future
Moneta (2005)	1970-2002	Euro area	• Probit	• 10 yield spreads • Lagged indicator variable	• The 10Y-3M spread is the best predictor of recession • Adding the lagged indicator variable improves the forecasts
Wright (2006)	1964-2005	US	• Probit	• Yield spread • Nominal federal funds rate • Real federal funds rate • Return forecasting factor	• Yield curve contains more information about the likelihood of a recession than just the spread • Yield curve contains more information about the likelihood of a recession than just the spread
Garner (1991)	1977-1991	US	• Regression • Bayesian VAR	• Conference Board index • Michigan index • Lagged durable goods purchases • Lagged real disposable income • Lagged CPI • Lagged unemployment rate	• CCI is not a reliable indicator by itself • CCI does not complement other variables in forecasting under normal circumstances • CCI may add value in predictions in exceptional cases
Fuhrer (1993)	1960-1990	US	• Cointegration/Error-Correction model • VAR	• Michigan index • Real personal disposable income • Civilian unemployment rate • Annual rate of inflation in personal consumption deflator • 3-month Treasury bill rate	• Independently, sentiment is not very useful in forecasting consumption • Sentiment's predictive power, though small, is statistically significant because it is systematic • The relationship between sentiment and consumption is not stronger in the 1990s than before
Matsusaka and Sbordonne (1995)	1953-1988	US	• VAR	• Michigan index • Leading indicators • Government expenditure • Default risk	• Exogenous changes in consumer sentiment have real effects on output
Batchelor & Dua (1997)	1970-1995	US	• Two-state Markov switching model	• Conference Board index • Expectations index	• The consumer confidence index helps predict switching points in economic activity • The consumer confidence index predicted the 1990 recession better than leading indicators

Batchelor (2001)	1965-2001	US & UK	• Markov switching model	• Business confidence • Consumer confidence • Index of coincident indicators • Default risk	• There are significant differences between the US and the UK • A decrease in business confidence leads to a decreased probability of staying in a good state • An increase in consumer confidence leads to a decreased probability of staying in a bad state
Howrey (2001)	1962-2000	US	• VAR	• ICS • Yield spread • NYSE price index • Leading indicators	• Out of all equations, the equation that includes the ICS and the yield spread is strongest
Mourougane and Roma (2003)	1995-2000*	BE, ES, DE, FR, IT & NL	• Regression	• ESI • ICI • Lagged real GDP growth • Industrial production growth • Lags of the confidence indicators	• Confidence indicators, especially ESI, are useful to forecast real GDP growth in the short run • Results are strong in all countries but Spain
Mazurek & Mielcová (2017)	1960-2015	US	• Regression • VAR	• CCI	• The Consumer Confidence Index is a suitable predictor of economic growth in the US • Non-systematic shocks may cause short-term estimations to deviate from the long-term trend
Blood & Philips (1995)	1989-1993	US	• VAR	• Recession headlines • Consumer sentiment • Leading indicators • Presidential approval	• Weak evidence of cointegration between economic news and the state of the economy is found
Wu, Stevenson, Chen & Güner (2002)	1987-1996	US	• VAR	• Recession news • Conference Board index • Economic Index	• When the entire period is considered, news does not significantly affect the economy • When looking at a down period separately, news does significantly affect the economy
Hester & Gibson (2003)	1998-2002	US	• Prais-Winsten procedure	• Economic news coverage • CPI • Unemployment • DJIA	• News about the economy was more likely to be framed as negative than as positive • Weak evidence that there is an influence of news on the economy
Gotthelf and Uhl (2018)	2003-2014	US	• Regression	• News sentiment • Macroeconomic control variables	• News sentiment is able to explain and predict movements in the yield curve

* This is the out-of-sample period, full period not mentioned

CHAPTER 3 Data

In this section, information about the data used is presented. In the first subsection, the different variables, the time period that is taken into consideration, and the databases through which the data is extracted are described. In the second subsection, some information about the descriptive statistics of the data is provided and a brief analysis is made. The last subsection contains normality tests.

3.1 Data collection

The research focuses on the United States and Germany, so for each variable, data is collected from these two countries. When possible, the data is collected from the same source. Initially, all available data is collected. Later, the research period is established based on the data availability of all variables. Since the two countries are analysed separately, and there are separate models for forecasting economic growth and predicting recessions, this research deals with time-series data. An overview of the variables and its sources can be found in table 8 at the end of the section.

3.1.1 Real GDP

The quarterly growth rate of real GDP is obtained from the OECD's Quarterly National Accounts database. It is the change in GDP, which is adjusted for inflation, over the previous quarter. The data is available from 1947 onwards, and since it is already quarterly, no transformations have to be made.

3.1.2 Recession dummy

Recession data is provided by the Business Cycle Dating Committee of the National Bureau of Economic Research for the US and the OECD for Germany. Both can be obtained from the Federal Reserve Bank of St. Louis. There are three options to choose from, which are three different interpretations of when a date is included in a recession. Relevant terms for these interpretations are "peak" and "trough", which are the negative and positive turning points of GDP growth, respectively. The interpretation used in this research is the trough method, which excludes the peak, but includes the trough. This interpretation is chosen because it is the one used by the FRED to shade areas on their graphs (Federal Reserve Bank of St. Louis, 2019). Data is available from 1854 onwards for the USA and from 1960 onwards for Germany. Again, no transformations are necessary.

3.1.3 Yield spread

The OECD also provides data on both the long-term interest rate and the short-term interest rate for both countries. The first are based on government bonds maturing in ten years, and the latter refers to three-month money market rates. Therefore, the short-term interest rate is subtracted from the long-term rate to obtain the yield spread for both countries. Data are available quarterly from 1964Q3 for the United States and 1960Q1 for Germany.

3.1.4 Consumer confidence

To quantify consumer confidence, the Consumer Confidence Index (CCI) is used. This indicator is constructed by using the answers to consumer opinion surveys. It is adjusted so the average is 100, and a value above 100 illustrates optimism whereas a value below 100 signals pessimism (OECD, 2019). The data obtained from the OECD Main Economic Indicators is monthly, so it has to be adjusted for this research. This is done by taking the average of January, February and March for Q1, the average of April, May and June for Q2, and so on.

3.1.5 News sentiment

For the news sentiment indicator, the database *Nexis Uni* is used. This offers the full text of many articles from international newspapers and newsmagazines. For the USA, *The Washington Post* is analyzed, using the keyword “economy”. This specific newspaper is chosen because it is one of the biggest daily US newspapers, and it dated back furthest in the database. For Germany, the *Süddeutsche Zeitung* will be analysed, using the keyword “Wirtschaft”. The database has a "Negative Business News" feature. When language is identified as negative by Nexis Uni, that search result is included in this category. To represent news sentiment, the fraction of articles per month that are included in this category will be used to predict whether a recession is likely. This is obtained by dividing the number of articles in the category by the total number of articles in that quarter. For *The Washington Post*, articles are available from 1977Q1 onwards. For the *Süddeutsche Zeitung*, articles only date back to 1994. Furthermore, after 2006, no articles are included in the “Negative Business News” category anymore for this paper. However, since media are biased towards presenting negative news (Soroka, 2004), a simple count of the number of articles mentioning “Wirtschaft” might already have predictive power. Therefore, two variables will be included for news sentiment: one representing the fraction of articles in the “Negative Business News” section and one representing the total number of articles in a quarter containing the keyword “economy” or “Wirtschaft”.

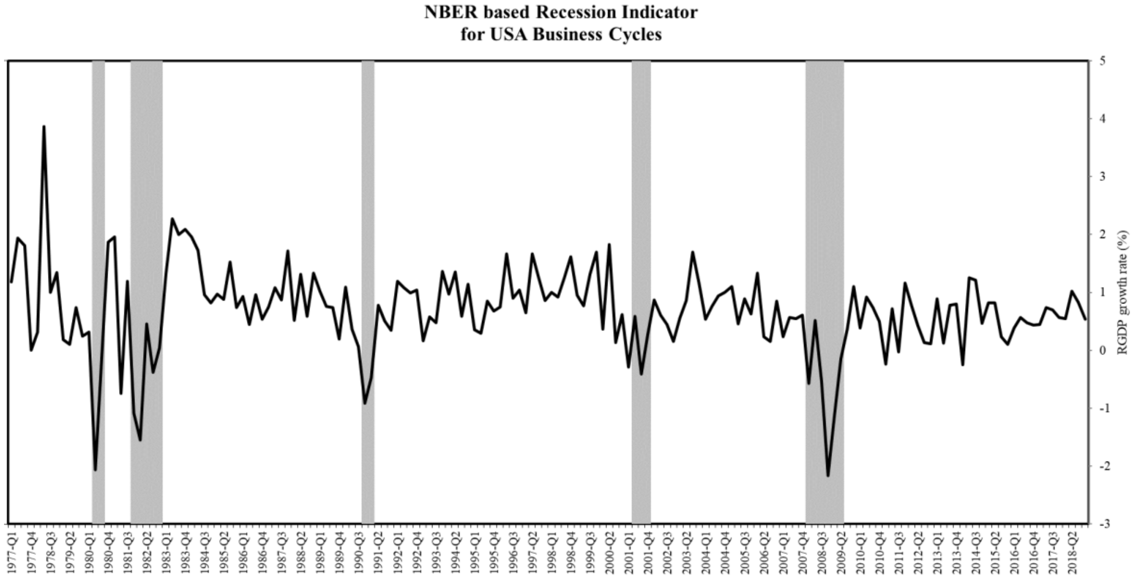
3.2 Descriptive statistics

During the period of interest in this research (1977 – 2018), both the United States and Germany have seen several recessionary and expansionary periods.

3.2.1 United States

Figure 1 shows the recessionary periods in the United States, which are shaded grey. During the period of interest, the USA has been in recession five times. The two recessions in the early 1980s were caused by the Federal Reserve, the central bank of the US, when they raised interest rates to counter the high inflation. The origin of the recession in 1990-1991 was the preceding Savings and Loan Crisis. The next recession happened in 2001, caused by the dot-com bubble. The most recent recession took place in 2008 – 2009, due to the subprime mortgage crisis (Amadeo, 2019).

Figure 1. Recessionary periods and real GDP growth rates in the USA (1977 – 2018)



Source: Federal Reserve Bank of St. Louis

When taking a look at the descriptive statistics of the variables, it can be seen that the USA was in a recession for 10.7% of the sample time period. For the period 1977 until 2018, on average, the economy grew by 0.683% per quarter, with the biggest growth of 3.864% in the second quarter of 1978 and the biggest decline in the last quarter of 2008. The difference between the yield on a ten-year Treasury note and the three-month Treasury bill was 1.184% on average, with the biggest inversion in the last quarter of 1980, when the short-term interest rate was 3.333% higher than then long-term interest rate. On average, 17% of the news was seen as negative, and the number of articles containing the keyword "economy" in a quarter ranged from 1022 in the third quarter of 1977 to 3218 in the first quarter of 2009.

Table 2. Descriptive statistics US

Variable	**Mean**	**Median**	**Standard Deviation**	**Min**	**Max**	**N**
Quarterly real GDP growth (%)	0.683	0.740	0.683	-2.164	3.864	168
Recessionary period (0 or 1)	0.107	0	0.310	0	1	168
10Y – 3M yield spread (%)	1.184	1.345	1.184	-3.333	3.510	168
Consumer Confidence Index	99.93	100.40	1.417	96.37	102.66	168
News sentiment (fraction)	0.171	0.168	0.030	0.050	0.269	168
News sentiment (absolute)	1802	1756	367	1022	3218	168

Table 3 shows the correlation matrix for the US variables. The correlation between the independent variables is assessed to find out whether multicollinearity could pose a problem. A rule of thumb for detecting multicollinearity is if correlation is higher than 0.80. This is not the case for any of the variables, so none of them have to be dropped. Interesting, however, is the strong negative correlation between the CCI and news sentiment. This means when consumers are more confident, there is less negative news, and vice versa.

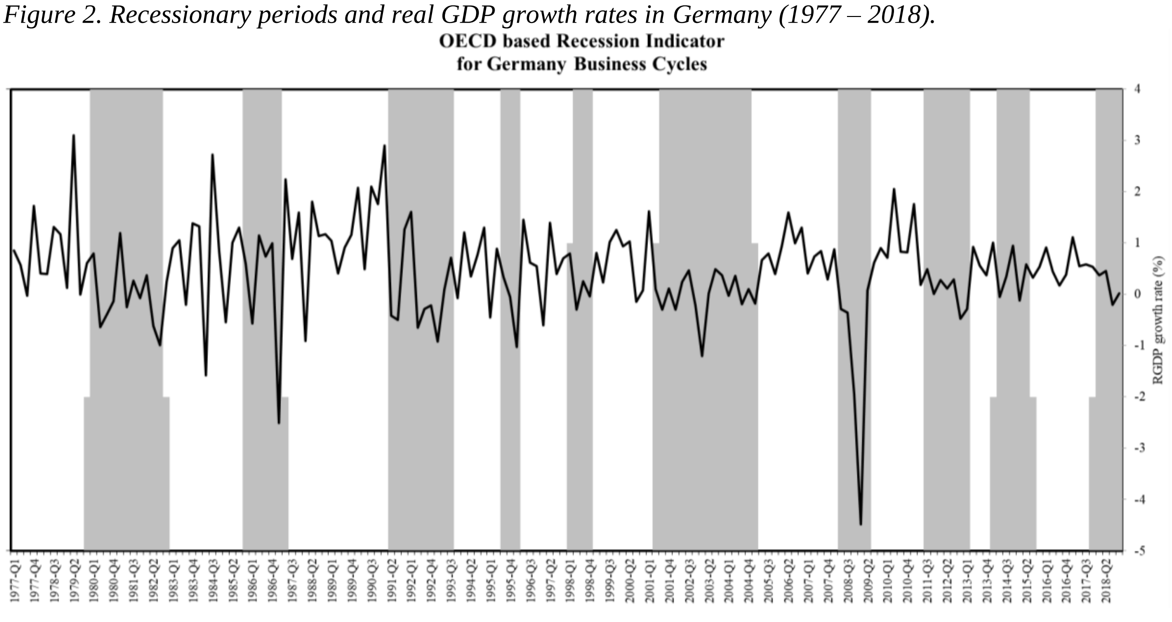
Table 3. Correlation matrix US variables

Variable	10Y – 3M yield spread (%)	Consumer Confidence Index	News sentiment (fraction)	News sentiment (absolute)
10Y – 3M yield spread (%)	1			
Consumer Confidence Index	0.0495	1		
News sentiment (fraction)	0.2152	-0.6062	1	
News sentiment (absolute)	0.2451	-0.1800	0.3792	1

3.2.2 Germany

Figure 2 shows the recessionary periods in Germany, which are again shaded grey. It is immediately apparent that Germany has been in a lot more, and longer, recessions than the United States. Additional recessions to those in the US took place in 1985-1986, 1995, 1998, 2011-2013 2014-2015 and 2018. In general, this is the case because the United States tends to recover more quickly from a recession. Because of the country's large size, there is more demand from within, and the country relies less on exports and imports of other countries that may also be in crisis (Harlan, 2014). The recession in quarter four of 2018 is due to shocks to various sectors, such as motor vehicles, chemicals and pharmaceuticals. Since these are one-off, they are expected to reverse soon (Davies, 2019).

Figure 2. Recessionary periods and real GDP growth rates in Germany (1977 – 2018).



Source: Federal Reserve Bank of St. Louis

The descriptive statistics for Germany are given in table 4. Average quarterly real GDP growth is lower than in the US. Just like the United States, Germany had the biggest economic decline during the Great Recession, in the first quarter of 2009. Germany was in recession for a striking 42.9% of the period. The average yield spread is smaller than in the US. Furthermore, the number of articles containing "Wirtschaft" has a much larger range than the number of articles containing "economy". The number of observations is lower for the news variables because of data availability, see 3.1.5.

Table 4. Descriptive statistics Germany

Variable	**Mean**	**Median**	**Standard Deviation**	**Min**	**Max**	**N**
Quarterly real GDP growth (%)	0.458	0	0.905	-4.486	3.105	168
Recessionary period (0 or 1)	0.429	0.457	0.484	0	1	168
10Y – 3M yield spread (%)	0.975	1.143	1.235	-2.827	3.127	168
Consumer Confidence Index	99.91	100.10	1.385	96.74	103.29	168
News sentiment (fraction)	0.002	0.001	0.002	0	0.007	52
News sentiment (absolute)	4183	4306	1553	165	7234	100

For Germany, too, a correlation matrix is used to detect if there is multicollinearity between the independent variables. Again, there does not seem to be a problem so all variables can be used simultaneously. It is interesting to note that there is a strong negative correlation between the yield spread and the CCI in Germany, while this is not at all the case for the US. Furthermore, yield spread is also negatively correlated to the news variables. As was the case for the United States, there is a negative correlation coefficient for the CCI and news sentiment, though it is much smaller.

Table 5. Correlation matrix German variables

Variable	10Y – 3M yield spread (%)	Consumer Confidence Index	News sentiment (fraction)	News sentiment (absolute)
10Y – 3M yield spread (%)	1			
Consumer Confidence Index	-0.3910	1		
News sentiment (fraction)	-0.1421	-0.1235	1	
News sentiment (absolute)	-0.0316	-0.2700	0.1877	1

There do not seem to be any problems with the data so far. However, it is important that all variables are normally distributed. Therefore, the paper proceeds with tests for skewness and kurtosis.

3.3 Normality testing

In this section, normality tests will be performed in this section to test whether the normal distribution assumption holds. This is done using two numerical measures of shape: *skewness* and *excess kurtosis*.

Whether there is skewness can be detected if the median differs greatly from the mean of a variable. This means there is asymmetry about the mean. Looking at the descriptive statistics from the previous section, this seems to be the case for a number of variables. However, skewness can be analysed more accurately by finding the skewness value. Perfect symmetry means a value of 0, and skewness can be both positive or negative. To assume normality, we require skewness to be between -0.5 and 0.5.

Kurtosis is related to the central peak, and how it compares to the normal bell curve. For a normal distribution, the kurtosis is 3. If kurtosis is less than 3, the distribution has a flat top and there are relatively few outliers. If kurtosis is greater than 3, the peak is very sharp with many outliers.

The *Jarque-Bera* test is widely used to test for normality. The tool combines skewness and kurtosis into a single test (Jarque & Bera, 1987). The test is performed using the following formula:

$$JB = \frac{n}{6}\left(S^2 + \frac{K^2}{4}\right) \sim X^2(2)$$

where n represents the number of observations, S denotes the sample skewness and K denotes the sample excess kurtosis (total kurtosis measured minus 3). The null hypothesis of normality is rejected at a 5% significance level if the value of the test is greater than 5.99. The test is performed for all variables except the recession dummy. The results are shown in tables 6 and 7.

For most of the variables in the sample, the null hypothesis of normal distribution has to be rejected. Only for the German variables of GDP growth and yield spread, normality can be assumed. Most of the variables have kurtosis higher than 3, which means the tails are fat and extreme events occur more often than normal. Furthermore, skewness is not within the [-0.5, 0.5] range for most variables either. Unfortunately, not a lot can be done about this, since the data have to be taken as they are. However, according to the Central Limit Theorem, when the sample is sufficiently large, a normal distribution can still be assumed. A sample is considered sufficiently large if it has more than 30 observations, which is the case for each of the variables in this research. Therefore, the problem of non-normality is ignored and after an overview of the variables on the next page, the paper proceeds with the Methodology section.

Table 6. Normality check US

Variable	**Skewness**	**Kurtosis**	**Jarque-Bera**	**P-value**
Quarterly real GDP growth (%)	0.502	6.497	91.536	0.000
10Y – 3M yield spread (%)	-1.307	14.617	980.721	0.000
Consumer Confidence Index	-0.372	4.088	242.299	0.002
News sentiment (fraction)	-0.847	8.671	70.569	0.000
News sentiment (absolute)	0.612	5.950	1397.140	0.000

Table 7. Normality check Germany

Variable	**Skewness**	**Kurtosis**	**Jarque-Bera**	**P-value**
Quarterly real GDP growth (%)	0.002	2.952	0.005	0.998
10Y – 3M yield spread (%)	-0.027	3.330	0.233	0.890
Consumer Confidence Index	-0.941	5.361	18.996	0.000
News sentiment (fraction)	1.213	4.812	19.096	0.000
News sentiment (absolute)	0.333	5.945	18.998	0.000

Table 8. Overview of variables

Variable	**Description**	**Period**	**Frequency**	**Source**
Prob_USA	1 – recessionary period 0 – expansionary period	1854Q4 – 2019Q1	Quarterly	Federal Reserve Bank of St. Louis
Prob_DEU	1 – recessionary period 0 – expansionary period	1960Q1 – 2018Q4	Quarterly	Federal Reserve Bank of St. Louis
RGDP_USA	Quarterly growth rates of real GDP	1947Q1 – 2018Q4	Quarterly	OECD
RGDP_DEU	Quarterly growth rates of real GDP	1947Q1 – 2018Q4	Quarterly	OECD
Yield_USA	Long-term interest rate – short-term interest rate	1964Q3 – 2018Q4	Quarterly	OECD
Yield_DEU	Long-term interest rate – short-term interest rate	1960Q1 – 2018Q4	Quarterly	OECD
Conf_USA	Consumer Confidence Index (CCI)	Jan 1960 – Mar 2019	Monthly	OECD
Conf_DEU	Consumer Confidence Index (CCI)	Jan 1973 – Mar 2019	Monthly	OECD
News_USA	Fraction of "Negative Business News" articles containing *economy* per quarter	1977Q1 – 2018Q4	Quarterly	Nexis Uni
News_DEU	Fraction of "Negative Business News" articles containing *Wirtschaft* per quarter	1994Q1 – 2006Q4	Quarterly	Nexis Uni
NewsTot_USA	Total number of articles in *The Washington Post* containing *economy* per quarter	1977Q1 – 2018Q4	Quarterly	Nexis Uni
NewsTot_DEU	Total number of articles in *Süddeutsche Zeitung* containing *Wirtschaft* per quarter	1994Q1 – 2018Q4	Quarterly	Nexis Uni

CHAPTER 4 Methodology

As explained in section 2.4, the Probit model and the GDP forecasting model will be used in this research. This section describes the methodology of the empirical data study. In section 4.1, stationarity is explained, an important concept when working with time series data. In section 4.2, the tests for cointegration are put forward. The GDP forecasting model is explained in section 4.3, and section 4.4 details the workings of the Probit model.

4.1 Stationarity

Before any tests are performed or models are estimated, it is necessary to check the data for non-stationarity. This is a problem because non-stationarity can lead to spurious regressions. This means a relationship is found even though none exists. Furthermore, the data does not follow a t-distribution so the standard statistical tools can no longer be used. Stationarity can be tested for using an Augmented Dickey-Fuller (ADF) test. Consider an AR model of order *p*:

$$y_t = \sum_{i=1}^{p} \beta_i \, y_{t-1} + u_t$$

where y_t is the variable of interest, *t* the time index, β the coefficient and u_t the error term. The model is non-stationary if there is a unit root, which is the case if $\alpha = 1$. There are three versions of the (Augmented) Dickey-Fuller test. Which one is used depends whether there seems to be a drift or deterministic time trend in the series. This will be checked using graphs, after which the appropriate ADF test is conducted. The appropriate lag length is chosen based on the partial autocorrelations given by the correlograms. Under the null hypothesis, a unit root is present. The alternative hypothesis states there is no unit root. If the null hypothesis cannot be rejected, the first difference of the series is taken (Dickey & Fuller, 1981).

4.2 Cointegration

If several variables turn out to be non-stationary, the estimators can still be valid if two or more of the variables are cointegrated. Cointegration means two or more time series share a common stochastic drift. If this is the case, there is a long-run relationship between the variables. If only one cointegration relationship is possible, the Engle-Granger test is used (Engle & Granger, 1987). If several variables might cointegrate, the Johansen test is used instead, because it allows for more than one cointegrating relationship (Johansen, 1988). If a cointegrating relationship is found, a Vector Error Correction Model (VECM) is estimated instead of the VAR model explained in section 4.3.2.

4.3 Forecasting GDP

4.3.1 Simple regressions

First, to evaluate the forecast performance at different forecast horizons, simple regressions will be run with only one lag of the independent variable. These regressions are of the form

$$RGDP_{t+k} = \alpha_0 + \alpha_1 X_t + \varepsilon_t$$

where α_0 is a constant term, α_1 the coefficient for the independent variable, X_t the independent variable (yield spread, consumer confidence or news sentiment), ε_t a normally distributed error term and k the forecast horizon measured in number of quarters. Following previous research (Estrella & Mishking, 1996; Davis & Fagan, 1997), forecasts for horizons from 1 to 8 quarters ahead will be made. Additionally, regressions will be run combining more than one of the explanatory variables. Newey-West (1987) standard errors are applied to control for heteroskedasticity and autocorrelation.

4.3.2 Vector autoregression model

When including more than one lag in the equation, the different time series might be dependent upon each other. Therefore, a vector autoregression model will be specified. In Filardo (1999), the GDP forecasting model is specified as a vector autoregression (VAR) model using several indicators. This model will be used, but changed slightly in this thesis to include the variables previously determined. Since it includes real GDP growth, yield spread, consumer confidence and news sentiment, it is a 4-variable vector autoregression. Each equation contains lags (up to 4 quarters) of the dependent variable, as well as the independent variables. The real GDP growth function is:

$$RGDP_t = \gamma_0 + \sum_{i=1}^{4} \gamma_i^{RGDP} RGDP_{t-i} + \sum_{i=1}^{4} \gamma_i^{Yield} Yield_{t-i} + \sum_{i=1}^{4} \gamma_i^{Conf} Conf_{t-i} + \sum_{i=1}^{4} \gamma_i^{News} News_{t-i} + \varepsilon_t$$

where γ_0 is a constant term, $RGDP_t$ is quarterly real GDP growth, γ_i^{RGDP} is the coefficient for the i^{th} lag of quarterly real GDP growth, $RDGP_{t-i}$ is the value of real GDP growth at the i^{th} lag, $\gamma_i^{Yield}$is the coefficient for the i^{th} lag of yield, $Yield_{t-i}$is the value of the independent variable at the i^{th} lag, etc. and ε_t is a normally distributed error term. The other equations in the system describe the dynamics of yield spread, consumer confidence and news sentiment, with the same explanatory variables:

$$Yield_t = \gamma_0 + \sum_{i=1}^{4} \gamma_i^{RGDP} RGDP_{t-i} + \sum_{i=1}^{4} \gamma_i^{Yield} Yield_{t-i} + \sum_{i=1}^{4} \gamma_i^{Conf} Conf_{t-i} + \sum_{i=1}^{4} \gamma_i^{News} News_{t-i} + \varepsilon_t$$

$$Conf_t = \gamma_0 + \sum_{i=1}^{4} \gamma_i^{RGDP} RGDP_{t-i} + \sum_{i=1}^{4} \gamma_i^{Yield} Yield_{t-i} + \sum_{i=1}^{4} \gamma_i^{Conf} Conf_{t-i} + \sum_{i=1}^{4} \gamma_i^{News} News_{t-i} + \varepsilon_t$$

$$News_t = \gamma_0 + \sum_{i=1}^{4} \gamma_i^{RGDP} RGDP_{t-i} + \sum_{i=1}^{4} \gamma_i^{Yield} Yield_{t-i} + \sum_{i=1}^{4} \gamma_i^{Conf} Conf_{t-i} + \sum_{i=1}^{4} \gamma_i^{News} News_{t-i} + \varepsilon_t$$

4.4 Predicting recessions

The Probit model, as proposed by Estrella and Mishkin (1998), does not focus on forecasting economic activity but solely on predicting recessions. A Probit model is a type of regression where the dependent variable can take only two values, in this case whether there is a recession or not. Mathematically, the model is defined as follows:

$$y_{t+k} = \beta' x_t + \varepsilon_t$$

where y_t determines the occurrence of a recession at time t, k is the length of the forecast horizon, β is a vector of coefficients, x_t is a vector of values of the independent variables, and ε_t is a normally distributed error term. To estimate the recession indicator R_t, where

$$R_t \begin{cases} 1, & \text{if } y_t > 0 \\ 0, & \text{otherwise}' \end{cases}$$

the cumulative normal distribution function (φ) corresponding to -ε is used:

$$P(R_{t+k} = 1) = \varphi(\beta' x_t)$$

The model is estimated using maximum likelihood. The likelihood function is defined as

$$L = \prod_{\{R_{t+k}=1\}} \varphi(\beta' x_t) \prod_{\{R_{t+k}=0\}} [1 - \varphi(\beta' x_t)].$$

4.5 Model evaluation

The goal of this thesis is to analyse the predictive power of the different variables and find the model that is most useful in forecasting GDP and predicting recessions. Therefore, it has to be determined which model performs best. The GDP forecasting model and the Probit model have two different purposes, so these will not be compared with each other. However, within each category, the different specifications are compared. This is done by assessing in-sample fit and out-of-sample performance.

4.5.1 In-sample fit

Model selection criteria based on information theory deal with the trade-off between a model's goodness of fit and how complex it is (Wang & Chaovalitwongse, 2010). The two most well-known are Akaike's Information Criterion (AIC) and the Bayesian Information Criterion or Schwarz criterion (BIC). The difference is in the penalty coefficient for the number of parameters, which is larger for the BIC than the AIC. In this research, the AIC will be used to assess the models' in-sample fit. The criterion is defined as

$$AIC = -2\ln(L) + 2k$$

where $\ln(L)$ is the maximized log-likelihood of the model and k the number of parameters.

4.5.2 Out-of-sample performance

A good in-sample fit does not automatically imply the model also forecasts well. Due to overfitting, in-sample error may be very small, but the forecasts very inaccurate. Therefore, it is essential to measure out-of-sample performance as well. Since there is no time to wait and see how well the forecasts did, the sample is split into an estimation period and a forecasting period. The sample is split into an estimation period of 80% and a forecasting period of 20%. This leads to 134 and 34 observations, representing the periods of 1977Q1 – 2010Q2 and 2010Q3-2018Q4, respectively.

A lot of different criteria exist to asses forecasting models. The four direct measures are mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). Since RMSE squares the errors before taking the average, larger errors have a relatively large weight. Large errors are to be avoided in forecasting GDP, since it might lead to large deficits after budgeting. Therefore, the RMSE is a useful measure to compare the models in this research. The formula is

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}}$$

where y_i the real observation and $\hat{y}_i$ the forecast value for observation i. The smaller the value, the more accurate the forecast.

The RMSE is useful to determine the best-performing model, but it does not prove that one model is significantly better than the other. A test of significance relevant for this is the Diebold-Mariano (DM) test (Diebold & Mariano, 1995). The test assesses whether the difference between the forecast error functions,

$$d_t = e\left(h_{1,t}\right) - e(h_{2,t})$$

is significantly different from 0, where $t = 1, \ldots, T$, and $e\left(h_{i,t}\right)$ the error function of forecast i at time t. The test uses the DM-statistic:

$$\frac{E(d_t)}{V(d_t)/\sqrt{T}} \sim N(0,1)$$

The MSE is used as the forecast error function in this case, since RMSE is not available in the software package for the Diebold-Mariano test.

CHAPTER 5 Results

This section provides the results of all statistical tests and models put forward in the methodology section. Part 5.1 establishes whether the time series are stationary or contain a unit root. If the time series are non-stationary, a test for cointegration will be conducted in section 5.2. Next, in sections 5.3 and 5.4, the results of the GDP forecasting model and the Probit model will be given. Section 5.5 evaluates the two models based on in-sample fit and out-of-sample performance.

5.1 Stationarity

First, it is determined whether the time series contain a unit root, since this can cause problems with the statistical inference. To find out whether this is the case, the Augmented Dickey-Fuller (1981) is performed for all variables except the recession dummy, since dummies are stationary by construction. The null hypothesis states a unit root is present, while the alternative hypothesis states there is no unit root present. If the time series is non-stationary, the test is repeated using the first difference of the series until the series achieve stationarity. The results of the tests are shown in tables 9 and 10.

Table 9. ADF test for the US

Variable	**Test type**	**Lag length**	**Test statistic**	**Critical values**		
				1%	**5%**	**10%**
RGDP_USA	2	1	-6.672	-3.488	-2.886	-2.576
Yield_USA	2	2	-2.982	-3.488	-2.886	-2.576
Conf_USA	2	4	-2.362	-3.489	-2.886	-2.576
ΔConf_USA	1	3	-6.599	-2.592	-1.950	-1.614
News_USA	2	3	-2.830	-3.489	-2.886	-2.576
ΔNews_USA	1	2	-9.390	-2.591	-1.950	-1.614
NewsTot_USA	2	1	-3.802	-3.488	-2.886	-2.576

Table 10. ADF test for Germany

Variable	**Test type**	**Lag length**	**Test statistic**	**Critical values**		
				1%	**5%**	**10%**
RGDP_DEU	2	0	-11.767	-3.488	-2.886	-2.576
Yield_DEU	2	2	-3.980	-3.488	-2.886	-2.576
Conf_DEU	2	4	-3.920	-3.489	-2.886	-2.576
News_DEU	2	1	-3.586	-3.580	-2.930	-2.600
NewsTot_DEU	2	1	-2.441	-3.513	-2.892	-2.581
ΔNewsTot_DEU	1	0	-10.655	-2.601	-1.950	-1.610

Only the time series for consumer confidence and news sentiment in the USA, and the absolute number of news articles in Germany contain a unit root. However, the first differences of these time series are stationary.

5.2 Cointegration

Variables can only be cointegrated if they are non-stationary. Therefore, only one cointegrating relationship is possible, between consumer confidence and news sentiment in the USA. To test whether the time series are indeed cointegrated, the Engle-Granger (1987) test is used. The first step is to estimate the following regression equation

$$Conf\_USA_t = \beta_1 + \beta_2 News\_USA_t + u_t$$

The equation for the estimated residuals is then

$$\hat{u}_t = Conf\_USA_t - \hat{\beta}_1 - \hat{\beta}_2 News\_USA_t$$

If consumer confidence and news sentiment are cointegrated, the estimated residuals, $\hat{u}_t$, must be stationary. Under the null hypothesis, consumer confidence and news sentiment are not cointegrated. If the null hypothesis is rejected, the residual is stationary and the time series are cointegrated. The results of the test are shown in table 11.

Table 11. Engle-Granger test for cointegration of Conf_USA and News_USA

Variable	**Test statistic**	**Critical values**		
		1%	**5%**	**10%**
$\hat{u}_t$	-4.384	-3.963	-3.373	-3.070

Since the null hypothesis is rejected, the series are cointegrated and a VECM is estimated instead of a VAR in 5.4.

5.3 Simple regressions

This subsection provides the results of the simple regressions for the USA as well as Germany. The equations are estimated using different combinations of the variables for the forecast horizons of 1 to 8 quarters.

5.3.1 United States

The results of the regressions for the United States are presented in table 12. For every variable, the coefficients are most significant at a short forecast horizon, either k = 1 or k = 2. However, the coefficient for yield spread stays significant up to seven lags. This is not the case for consumer confidence and news sentiment, for which the explanatory power dies out after two lags. For the United States, the absolute number of news articles is slightly more informative than the fraction of negative articles, so the *NewsTot* variable is included in the further regressions. Including additional variables logically increases the R^2, but this does not mean anything. The lower RMSE does indicate an improvement of the model. However, to analyse the in-sample fit more accurately, the AIC is calculated for the different models, and it is given in table 12.

Table 12. Different specifications of the simple regressions for the US

Probit model	**Forecast horizon (*k* = number of quarters)**							
	k=1	***k=2***	***k=3***	***k=4***	***k=5***	***k=6***	***k=7***	***k=8***
Yield								
R^2	0.065	**0.100**	0.060	0.073	0.081	0.053	0.039	0.016
α	.526***	**.482*****	.519***	.509***	.502***	.523***	.539***	.575***
coefficient yield	.130***	**.160*****	.123**	.135**	.142***	.108**	.092*	.059
RMSE	0.721	**0.704**	0.716	0.712	0.710	0.680	0.687	0.695
∆Conf								
R^2	**0.112**	0.027	0.012	0.056	0.004	0.004	0.004	0.014
α	**.669*****	.664***	.668***	.669***	.651***	.649***	.645***	.649***
coefficient conf	**.502*****	.243*	.160	.352***	.087	.093	.092	-.166
RMSE	**0.699**	0.729	0.735	0.720	0.698	0.699	0.699	0.697
∆News								
R^2	**0.038**	0.021	0.000	0.024	0.000	0.001	0.002	0.001
α	**.673*****	.665***	.669***	.671***	.652***	.650***	.646***	.648***
coefficient news	**-6.33****	-4.69*	-.102	-4.97	.331	-.971	1.49	1.04
RMSE	**0.728**	0.731	0.739	0.732	0.699	0.700	0.700	0.701
NewsTot								
R^2	**0.040**	0.015	0.006	0.000	0.001	0.017	0.025	0.015
α	**1.41*****	1.12***	.952***	.744**	.592*	.202	.098	.219
coefficient newstot	**-.000****	-.000	-.000	-.000	.000	.000	.000**	.000
RMSE	**0.731**	0.736	0.736	0.739	0.741	0.693	0.692	0.695
Yield/∆Conf								
R^2	**0.145**	0.107	0.066	0.108	0.077	0.052	0.038	0.039
α	**.557*****	.492***	.527***	.531***	.497***	.524***	.541***	.560***
coefficient yield	**.095****	.145***	.120**	.117**	.130**	.105**	.087	.075
coefficient conf	**.444*****	.153	.087	.280***	.006	.028	.038	-.213*
RMSE	**0.688**	0.700	0.716	0.702	0.673	0.684	0.689	0.690
Yield/NewsTot								
R^2	0.138	**0.142**	0.080	0.080	0.083	0.059	0.052	0.024
α	1.50***	**1.22*****	1.03***	.827***	.677**	.260	.145	.249
coefficient yield	.164***	**.185*****	.140***	.146***	.148***	.099**	.079*	.047
coefficient newstot	-.001***	**-.000*****	-.000	-.000	-.000	.000	.000*	.000
RMSE	0.695	**0.689**	0.711	0.711	0.712	0.680	0.684	0.694
Yield/∆Conf/NewsTot								
R^2	**0.194**	0.140	0.090	0.114	0.077	0.060	0.055	0.044
α	**1.37*****	1.15***	1.08***	.805***	.483**	.228	.080	.324
coefficient yield	**.128****	.172***	.142***	.128**	.130**	.094**	.069	.066
coefficient conf	**.398*****	.116	.055	.264**	.007	.044	.064	-.200*
coefficient newstot	**-.000*****	-.000**	-.000*	-.000	.000	.000	.000*	.000
RMSE	**0.670**	0.690	0.710	0.702	0.676	0.684	0.685	0.691

*** p-value < 0.1, **p-value < 0.05, *** p-value < 0.01**

Table 13. Akaike Information Criteria (AIC) for the different regressions in the US

Probit model	**Forecast horizon (*k* = number of quarters)**							
	k=1	***k=2***	***k=3***	***k=4***	***k=5***	***k=6***	***k=7***	***k=8***
Yield								
AIC	329.0	320.7	329.0	328.2	**328.1**	**332.2**	334.4	337.9
ΔConf								
AIC	317.9	334.1	337.6	330.3	340.0	340.0	339.9	338.3
ΔNews								
AIC	334.6	335.9	340.4	336.7	340.6	340.5	340.2	340.4
NewsTot								
AIC	337.0	339.6	340.4	340.3	339.3	337.2	336.1	338.4
Yield/ΔConf								
AIC	312.6	319.5	329.8	**323.4**	330.0	334.2	336.3	**336.3**
Yield/NewsTot								
AIC	322.2	316.9	**329.0**	330.0	330.0	333.0	**334.1**	338.7
Yield/ΔConf/NewsTot								
AIC	**308.9**	**316.9**	330.2	325.4	332.0	334.9	335.7	337.5

As indicated by the AIC, there is not a single model that performs best at each forecast horizon. For short lead times (k = 1 and k = 2), the full model with all three variables is preferable. When predicting 3 or 7 quarters ahead, the model with yield spread and the number of news articles has the lowest AIC. At k = 4 and k = 8, this is the model with yield spread and consumer confidence. At average lead times (k = 5 and k = 6), the model with solely yield spread performs best. Overall, the model with yield spread, consumer confidence and the news factor at a lead time of only 1 quarter, has the lowest AIC.

Table 14. Root Mean Squared Error (RMSE) for the out-of-sample regressions in the US

Probit model	**Forecast horizon (*k* = number of quarters)**							
	k=1	***k=2***	***k=3***	***k=4***	***k=5***	***k=6***	***k=7***	***k=8***
Yield								
RMSE	0.212	0.221	0.208	0.227	0.242	0.222	0.212	0.195
ΔConf								
RMSE	**0.197**	0.192	0.203	0.186	0.178	0.184	0.178	**0.172**
ΔNews								
RMSE	0.203	**0.176**	0.178	**0.185**	**0.174**	**0.176**	**0.175**	0.175
NewsTot								
RMSE	0.200	0.182	**0.176**	0.177	0.178	0.183	0.202	0.208
Yield/ΔConf								
RMSE	0.214	0.222	0.223	0.229	0.230	0.224	0.211	0.198
Yield/ΔNews								
RMSE	0.228	0.217	0.211	0.232	0.230	0.222	0.213	0.198
Yield/ΔConf/ΔNews								
RMSE	0.227	0.221	0.224	0.232	0.230	0.224	0.214	0.199

To assess the out-of-sample performance of the models, the sample is split into an estimation and forecasting period, and the RMSE is calculated. The results are given in table 14. Evaluating the models by out-of-sample performance provides rather different outcomes than by in-sample fit. For none of the forecast horizons, any model that combines the variables performs better than the models with a single variable. For most lead times, news sentiment provides the most accurate forecasts. For k = 1 and k = 8, this is consumer confidence. The most accurate forecast overall is the one provided by consumer confidence 8 quarters ahead, with a RMSE of 0.172. To test whether this model performs significantly better than the one selected using the AIC criterion for in-sample fit, the Diebold-Mariano test is used. The output is given in table 15.

Table 15. Diebold-Mariano test for the simple regressions in the US

Probit model	**Difference**	**DM statistic**	**P-value**
Yield/ΔConf/ΔNews (k = *1*) **- ΔConf** (k = 8)	0.06901	1.290	0.197

The output shows that even though the RMSE is lower for the model including only consumer confidence, this model does not perform *significantly* better than the model containing yield spread, consumer confidence and news sentiment, which was the best model based on in-sample fit.

5.3.2 Germany

The same regressions are performed with the data for Germany. The results are given in table 16. Notable is the news factor, which becomes significant at a long lead time (k = 7). Contrary, yield spread is only significant for lead times up to k = 5, while consumer confidence is only informative up to 3 quarters ahead. Whether the inclusion of additional variables improves the model cannot be concluded based on these results, since the number of observations is not the same for every model. Therefore, the AIC is calculated for each model and forecast horizon using the same number of observations (44). Those results are given in table 17.

When forecasting up to four quarters ahead, the full model with all three variables has the best in-sample fit. However, with longer lead times, the models without consumer confidence do better. For k = 6 and k = 8, the model with solely yield spread has the lowest AIC. For k = 5, this is the model with the number of news articles as explanatory variable and for k = 7 the model with yield spread and news sentiment. Again, this is only the in-sample fit, which can be biased due to overfitting. Therefore, the out-of-sample performance is also evaluated. Because there are only 52 observations for the models including news sentiment, the estimation period includes the first 42 observations and the forecasting period includes the last 10 observations. This means the estimation period runs from 1994:1 to 2004:2 and the forecasts are made for 2004:3 until 2006:4. The results from calculating the RMSE are shown in table 18.

Table 16. Different specifications of the simple regressions for Germany

Probit model	**Forecast horizon (*k* = number of quarters)**							
	k=1	***k=2***	***k=3***	***k=4***	***k=5***	***k=6***	***k=7***	***k=8***
Yield								
R^2	**0.057**	0.055	0.043	0.033	0.023	0.014	0.012	0.009
α	**286*****	.287***	.309***	.321***	.343***	.366***	.368***	.373***
coefficient yield	**.174*****	.172***	.152**	.132**	.110*	.087	.079	.069
RMSE	**0.883**	0.887	0.894	0.897	0.904	0.911	0.912	0.915
Conf								
R^2	**0.060**	0.035	0.020	0.002	0.000	0.001	0.915	0.006
α	**-15.5*****	-11.8**	-8.81*	-2.53	.606	1.99	5.58	5.76
coefficient conf	**.160*****	.123**	.093**	.030	-.002	-.015	-.051	-.053
RMSE	**0.882**	0.896	0.905	0.911	0.915	0.917	0.915	0.916
News								
R^2	0.013	0.030	0.070	0.092	0.009	0.020	**0.109**	0.006
α	.337***	.308**	.252**	.213*	.367***	.314***	**.186***	.296**
coefficient news	42.9	65.5	100.8**	112.9***	34.0	52.9	**123.8*****	33.0
RMSE	0.636	0.632	0.620	0.601	0.620	0.620	**0.600**	0.710
ΔNewsTot								
R^2	0.009	**0.012**	0.005	0.000	0.003	0.012	0.002	0.000
α	.365***	**.360*****	.350***	.358***	.353***	.351***	.356***	.372***
coefficient newstot	-.000	**-.000**	-.000	-.000	-.000	.000	.000	-.000
RMSE	0.795	**0.797**	0.799	0.800	0.801	0.802	0.810	0.802
Yield/Conf								
R^2	**0.096**	0.075	0.053	0.033	0.024	0.017	0.023	0.020
α	**-13.0*****	-9.16*	-6.38	-.176	2.74	3.76	7.40	7.42
coefficient yield	**.144****	.150***	.137**	.131**	.116*	.095*	.096*	.086*
coefficient conf	**.134*****	.095**	.067	.005	-.024	-.034	-.071	-.071
RMSE	**0.867**	0.880	0.892	0.899	0.906	0.913	0.910	0.913
Yield/News								
R^2	0.013	0.034	0.082	0.093	0.009	0.047	**0.216**	0.134
α	.327	.403*	.396*	.169	.402*	.103	**-.247**	-.232
coefficient yield	.006	-.057	-.087	.027	-.021	.128	**.263****	.320**
coefficient news	43.3	61.6	94.9**	114.7***	32.6	61.5	**141.5*****	54.6
RMSE	0.643	0.637	0.622	0.607	0.626	0.618	**0.568**	0.669
Yield/Conf/News								
R^2	0.080	0.048	0.091	0.135	0.010	0.049	**0.229**	0.134
α	-19.5*	-8.44	-7.03	-15.4	-.582	-3.64	**-8.71**	-.882
coefficient yield	.102	-.015	-.051	.102	-.017	.146	**.303*****	.324**
coefficient conf	.198*	.088	.074	.156	.010	.037	**.085**	.006
coefficient news	63.1	70.5	102.4*	130.3***	33.6	65.3	**150.0*****	55.3
RMSE	0.627	0.639	0.626	0.599	0.632	0.623	**0.569**	0.676

*** p-value < 0.1, **p-value < 0.05, *** p-value < 0.01**

Table 17. Akaike Information Criteria (AIC) for the different regressions in Germany

Probit model	**Forecast horizon (*k* = number of quarters)**							
	k=1	***k=2***	***k=3***	***k=4***	***k=5***	***k=6***	***k=7***	***k=8***
Yield								
AIC	85.14	85.22	85.21	85.22	84.91	**83.52**	82.44	**83.47**
Conf								
AIC	81.71	83.16	83.34	83.88	84.91	85.23	84.83	83.81
News								
AIC	84.83	83.47	82.84	81.79	84.87	84.15	75.28	84.23
ΔNewsTot								
AIC	84.68	85.11	85.23	85.01	**82.58**	84.38	85.16	83.73
Yield/Conf								
AIC	79.69	83.54	84.74	85.03	85.93	85.17	84.42	85.03
Yield/News								
AIC	86.67	85.40	84.84	83.65	86.45	84.22	**73.95**	84.50
Yield/Conf/News								
AIC	**79.23**	**81.53**	**82.25**	**80.85**	86.85	85.14	74.86	86.30

Table 18. Root Mean Squared Error (RMSE) for the out-of-sample regressions in Germany

Probit model	**Forecast horizon (*k* = number of quarters)**							
	k=1	***k=2***	***k=3***	***k=4***	***k=5***	***k=6***	***k=7***	***k=8***
Yield								
RMSE	0.304	0.286	**0.281**	0.296	0.286	0.287	0.272	0.266
Conf								
RMSE	**0.263**	**0.277**	0.282	0.293	**0.284**	0.285	0.285	0.282
News								
RMSE	0.276	0.277	0.293	**0.238**	0.314	0.323	0.215	0.282
ΔNewsTot								
RMSE	0.279	0.284	0.287	0.283	0.286	**0.283**	0.285	0.288
Yield/Conf								
RMSE	0.291	0.284	0.284	0.311	0.287	0.289	0.278	0.279
Yield/News								
RMSE	0.298	0.282	0.291	0.248	0.316	0.323	0.203	**0.265**
Yield/Conf/News								
RMSE	0.278	0.277	0.290	0.249	0.317	0.325	**0.198**	0.277

As was the case for the United States, the in-sample fit gives a very different view than the out-of-sample performance. When looking at table 18, the full model does not perform as well as it seemed from table 17. Now, the models with only a single variable look better. Combining the variables only leads to better models for long forecast horizons (k = 7 and k = 8). The model with the lowest RMSE overall is the model with yield spread, consumer confidence and news sentiment, forecasting 7 quarters ahead. Whether this model performs significantly better than the model with solely yield spread is evaluated using the Diebold-Mariano test. The results are presented in table 19.

Table 19. Diebold-Mariano tests for the simple regressions in Germany

Probit model	**Difference**	**DM statistic**	**P-value**
Yield - Yield/Conf/News	0.03605	1.100	0.272

Similar to the US, the output shows that even though the RMSE is lower for the model including all variables, it does not perform *significantly* better than the model containing solely yield spread.

5.4 VAR

5.4.1 United States

For the United States, since consumer confidence and news sentiment cointegrate, a vector error correction model is estimated instead of a vector autoregression. The results are presented in table 20.

Table 20. VECM model for the US

Effect of	**Lags (*i = number of quarterly lags*)**		
	i = 1	*i* = 2	*i = 3*
On RGDP			
CointEq1	-.640***		
RGDP	-.221*	-.127	-.058
Yield	-.060	.208***	-.043
Conf	.279**	-.076	.095
News	-7.28**	-5.59*	-1.64
On yield			
CointEq1	-.385***		
RGDP	.215*	.011	.040
Yield	.018	-.241***	.037
Conf	.-.248*	.022	-.224*
News	-3.02	-6.20*	-2.70
On conf			
CointEq1	-.098		
RGDP	.147*	.098	.068
Yield	.110*	-.070	.044
Conf	.212**	-.255***	.188**
News	-3.46	1.05	1.15
On news			
CointEq1	-.000		
RGDP	-.003	-.003	-.005**
Yield	.000	.003	-.002
Conf	-.012***	-.003	-.007*
News	-.614***	-.483***	-.270**

*** p-value < 0.1, **p-value < 0.05, *** p-value < 0.01**

As can be seen in the table, quarterly real GDP growth is influenced by yield spread (at lag 2), consumer confidence (at lag 1) and news sentiment (at lag 1), but not by its own lags. Furthermore, yield spread is only influenced by the cointegration equation and its own lag. For consumer confidence, its own lags are highly significant, but none of the other variables influence consumer confidence. News sentiment is influenced by real GDP growth (at lag 3), consumer confidence (at lag 1) and its own lags.

5.4.2 Germany

The results of the vector autoregression for Germany are presented in table 21.

Table 21. VAR model for Germany

Effect of	**Lags (*i = number of quarterly lags*)**			
	i = 1	*i* = 2	*i* = 3	*i* = 4
On RGDP				
RGDP	-.224	.090	.017	.205
Yield	.468	.101	-.958**	.678***
Conf	.977***	-1.24**	.582	-.005
News	-13.9	49.0	88.0	125.6**
On yield				
RGDP	-.082	-.172**	-.026	-.005
Yield	1.30***	-.379*	.042	-.070
Conf	-.040	.186	-.140	.053
News	44.2*	-48.1*	65.2**	-98.9***
On conf				
RGDP	-.021	.159**	.043	.064
Yield	.052	.090	-.359*	.229**
Conf	1.88***	-1.76***	1.10***	-.393***
News	24.5	-33.4	19.6	8.20
On news				
RGDP	.000	-.000	-.001	-.001
Yield	.000	-.002*	.002	-.000
Conf	-.000	-.000	-.001	-.001*
News	.255*	-.088	.222	.133

*** p-value < 0.1, **p-value < 0.05, *** p-value < 0.01**

Based on these results, quarterly real GDP growth is influenced by yield spread, consumer confidence and news sentiment, but not by its own lags. Furthermore, yield spread is influenced by news sentiment (at all lags), but not by consumer confidence. On the other hand, consumer confidence is influenced by the yield spread (at lag 3 and 4), and its own lags are also very significant. News sentiment is not significantly influenced by any of the lags at a 5% level.

5.5 Probit model

This subsection provides the results of the different specifications of the Probit model for the USA as well as Germany. The Probit model is estimated using different combinations of the variables for the forecast horizons of 1 to 8 quarters.

5.5.1 United States

Table 22. Different specifications of the Probit model for the US

Probit model	**Forecast horizon (*k* = number of quarters)**							
	k=1	***k=2***	***k=3***	***k=4***	***k=5***	***k=6***	***k=7***	***k=8***
Yield								
Pseudo R^2	0.088	0.227	0.303	0.372	**0.378**	0.364	0.301	0.243
t-stat yield	-3.10***	-4.55***	-5.02***	-5.29***	**-5.29*****	-5.29***	-5.02***	-4.63***
Log likelihood	-52.05	-44.05	-39.63	-35.62	**-35.23**	-35.94	-39.13	-42.578
ΔConf								
Pseudo R^2	**0.207**	0.126	0.103	0.067	0.000	0.001	0.001	0.004
t-stat conf	**-4.22*****	-3.50***	-3.22***	-2.65***	-0.08	-0.25	-0.37	-0.62
Log likelihood	**-45.21**	-49.70	-50.91	-52.86	-56.51	-56.36	-56.21	-55.96
ΔNews								
Pseudo R^2	**0.106**	0.070	0.026	0.003	0.007	0.000	0.000	0.000
t-stat news	**3.24*****	2.67***	1.68*	0.54	0.88	-0.00	-0.01	-0.21
Log likelihood	**-50.93**	-52.91	-55.29	-56.48	-56.12	-56.39	-56.27	-56.13
NewsTot								
Pseudo R^2	**0.106**	0.034	0.001	0.005	0.016	0.040	0.080	0.092
t-stat newstot	**3.30*****	1.96*	0.29	-0.71	-1.29	-2.00**	-2.73***	-2.91***
Log likelihood	**-51.03**	-55.04	-56.82	-56.48	-55.74	-54.25	-51.91	-51.09
Yield/ΔConf								
Pseudo R^2	0.263	0.333	0.380	**0.419**	0.392	0.375	0.311	0.268
t-stat yield	-2.49	-4.28***	-4.73***	**-5.06*****	-5.25***	-5.25***	-5.04***	-4.74***
t-stat conf	-3.95***	-3.19***	-2.83***	**-2.25****	1.27	1.14	0.81	1.66*
Log likelihood	-42.01	-38.13	-35.17	**-32.88**	-34.38	-35.26	-38.78	-41.11
Yield/NewsTot								
Pseudo R^2	0.281	0.350	0.341	**0.385**	0.380	0.366	0.330	0.281
t-stat yield	-4.10***	-4.82***	-5.01***	**-5.11*****	-5.13***	-5.07***	-4.68***	-4.20***
t-stat newstot	4.04***	3.41***	2.08**	**1.19**	0.42	-0.50	-1.55	-1.94*
Log likelihood	-41.04	-37.03	-37.45	**-34.92**	-35.15	-35.81	-37.81	-40.44
Yield/ΔConf/NewsTot								
Pseudo R^2	0.391	0.414	0.400	**0.423**	0.396	0.376	0.333	0.307
t-stat yield	-3.48***	-4.47***	-4.65***	**-4.87*****	-4.99***	-4.99***	-4.66***	-4.29***
t-stat conf	-3.16***	-2.54**	-2.47**	**-2.04****	1.38	1.08	0.70	1.65*
t-stat newstot	3.37***	2.87***	1.48	**0.68**	0.68	-0.36	-1.51	-1.95*
Log likelihood	-34.71	-33.34	-34.08	**-32.66**	-34.16	-35.20	-37.53	-38.94

*** p-value < 0.1, **p-value < 0.05, *** p-value < 0.01**

The presented results confirm the first hypothesis that on a stand-alone basis, the yield spread is the strongest predictor for recessions. The pseudo R^2 is much higher for the model containing yield spread than those containing consumer confidence and news sentiment, for every forecast horizon except $k = 1$. Furthermore, the coefficient for yield spread is statistically significant at every forecast horizon, while the coefficient for confidence is only significant up to $k = 4$ and for news up to $k = 2$. This means the lead time is much longer for yield spread, so the forecasts are more timely. However, it does seem to be the case that adding consumer confidence and news sentiment to the models significantly improves the forecasting power. The pseudo R^2 is even higher for these forecasts. To test whether this is true in-sample, the AIC criterion is calculated for the different models using the same number of observations (159). This is done for forecast horizons $k = 4$ and $k = 5$, because these seem to give the strongest forecasts. The results are shown in table 23.

Table 23. Akaike Information Criteria (AIC) for the different Probit models in the US

Probit model	**Forecast horizon (*k* = number of quarters)**							
	***k* = 1**	***k* = 2**	***k* = 3**	***k* = 4**	***k* = 5**	***k* = 6**	***k* = 7**	***k* = 8**
Yield								
AIC	106.12	91.10	82.72	74.95	**74.28**	**75.80**	82.21	89.13
Yield/ΔConf								
AIC	87.21	81.31	**75.82**	**71.47**	74.64	76.47	83.53	88.23
Yield/NewsTot								
AIC	87.90	80.01	80.81	75.74	76.16	77.49	**81.45**	86.75
Yield/ΔConf/NewsTot								
AIC	**76.86**	**74.63**	76.01	73.15	76.25	78.31	82.94	**85.88**

Among the different forecast horizons, it varies which model provides the best forecast. For $k = 1$, $k = 2$ and $k = 8$, this is the model containing all variables. For $k = 3$ and $k = 4$, the model containing yield spread and consumer confidence forecasts best. For $k = 5$ and $k = 6$, the model containing solely yield spread provides the best forecasts, and for $k = 7$ the best-performing model combines yield spread and news sentiment. However, overall, the model with the lowest AIC is the one combining yield spread and consumer confidence at a forecast horizon of 4 quarters ahead.

Since the in-sample fit can be biased due to overfitting, the out-of-sample performance of the models is evaluated next. The sample is split into an estimation period and a forecasting period. The model based on the estimation period is then used to forecast out-of-sample, and these forecast values are compared to the actual values of the forecasting period. This is done by calculating the RMSE of the different forecasts at different horizons. The models without yield spread are disregarded, since these seem to perform far worse than those containing yield spread. The results are given in table 24.

Table 24. Root Mean Squared Error (RMSE) for the Probit models in the US out-of-sample

Probit model	**Forecast horizon (*k* = number of quarters)**							
	k=1	***k=2***	***k=3***	***k=4***	***k=5***	***k=6***	***k=7***	***k=8***
Yield								
RMSE	0.038	0.024	0.018	0.012	0.011	**0.011**	**0.017**	**0.022**
Yield/ΔConf								
RMSE	0.047	0.022	0.014	0.009	0.013	0.013	0.018	0.026
Yield/NewsTot								
RMSE	**0.034**	0.020	0.013	0.009	**0.010**	0.013	0.021	0.026
Yield/ΔConf/NewsTot								
RMSE	0.034	**0.014**	**0.010**	**0.007**	0.012	0.013	0.022	0.029

The results for the out-of-sample performance are quite similar to those for the in-sample fit. Again, the model with solely the yield curve as a predictor performs best 5 quarters ahead, while the other models perform best 4 quarters ahead. However, whereas the AIC criterion selected the model combining yield spread and consumer confidence as the best model, when using RMSE, adding the news factor does seem to improve the forecasting power. Furthermore, it seems that the models containing news perform best at short lead times (k = 1 to k = 5), whereas the model containing solely yield spread performs best at longer lead times (k = 6 to k = 8). Overall, the model with the lowest RMSE is lowest for the model containing yield spread, consumer confidence and the number of news articles, at a forecast horizon of 4 quarters ahead.

As mentioned before, the RMSE statistic does not prove whether one model is significantly better than the other. To test whether the models containing additional variables are significantly better (at $k = 4$) than the model with solely yield spread, the Diebold-Mariano test is used. The output is given in table 25.

Table 25. Diebold-Mariano tests for the Probit models in the US

Probit model	**Difference**	**DM statistic**	**P-value**
Yield - Yield/ΔConf	0.000085	1.098	0.272
Yield - Yield/ΔConf/NewsTot	0.0001136	1.084	0.278
Yield/ΔConf - Yield/ΔConf/NewsTot	0.0000285	1.044	0.296

The output shows that even though the RMSE is lower for the models including consumer confidence and/or news sentiment, these models do not perform *significantly* better than the model containing solely yield spread as a predictor. Therefore, the second hypothesis, which states that adding consumer confidence and news sentiment to the models significantly improves the forecasting power, cannot be accepted based on these results. To assess whether the third hypothesis is true, the same results for Germany are analysed next.

5.5.2 Germany

Table 26. Different specifications of the Probit model for Germany

Probit model	**Forecast horizon (*k* = number of quarters)**							
	k=1	***k=2***	***k=3***	***k=4***	***k=5***	***k=6***	***k=7***	***k=8***
Yield								
Pseudo R^2	**0.122**	0.109	0.074	0.044	0.020	0.008	0.001	0.000
t-stat yield	**-4.92*** **	-4.67***	-3.93***	-3.09***	-2.12**	-1.33	-0.45	0.13
Log likelihood	**-101.20**	-102.19	-105.62	-108.37	-110.45	-111.20	-111.34	-110.78
Conf								
Pseudo R^2	0.040	0.011	0.000	0.006	0.016	0.024	0.034	**0.047**
t-stat conf	-2.99***	-1.58	-0.16	1.16	1.89*	2.31**	2.71***	**3.15*** **
Log likelihood	-110.63	-113.37	-113.99	-112.69	-110.91	-109.36	-107.64	**-105.60**
News								
Pseudo R^2	0.000	0.020	0.058	0.062	0.063	0.076	**0.080**	0.020
t-stat news	-0.06	-1.16	-1.87*	-1.92*	-1.94*	-2.12**	**-2.18** **	-1.18
Log likelihood	-35.695	-34.99	-33.64	-33.49	-33.46	-33.15	**-33.01**	-35.16
∆NewsTot								
Pseudo R^2	0.000	0.001	0.000	0.001	0.000	0.000	**0.002**	0.000
t-stat newstot	0.04	0.27	0.20	0.29	0.11	-0.22	**-0.53**	0.03
Log likelihood	-67.85	-67.15	-66.50	-65.80	-65.15	-64.43	**-63.61**	-63.03
Yield/Conf								
Pseudo R^2	**0.146**	0.111	0.076	0.060	0.047	0.040	0.039	0.048
t-stat yield	**-4.60*** **	-4.52***	-3.99***	-3.41***	-2.58**	-1.87*	-1.05	-0.54
t-stat conf	**-2.32** **	-0.76	0.68	1.88*	2.41**	2.65***	2.87***	3.20***
Log likelihood	**-98.41**	-101.90	-105.39	-106.58	-107.49	-107.59	-107.09	-105.46
Yield/News								
Pseudo R^2	0.000	0.021	0.061	0.067	0.068	0.123	**0.146**	0.076
t-stat yield	-0.13	-0.27	-0.49	-0.60	-0.62	-1.79*	**-2.11** **	-1.96*
t-stat news	-0.08	-1.18	-1.91*	-1.98**	-1.99**	-2.32**	**-2.43** **	-1.47
Log likelihood	-35.69	-34.95	-33.52	-33.31	-33.26	-31.47	**-30.66**	-33.17
Yield/Conf/News								
Pseudo R^2	0.080	0.108	0.136	0.116	0.113	0.234	**0.236**	0.081
t-stat yield	-1.12	-1.30	-1.40	-1.29	-1.28	-2.73***	**-2.83*** **	-2.01**
t-stat conf	-2.31**	-2.41**	-2.25**	-1.85*	-1.77*	-2.59**	**-2.32** **	-0.61
t-stat news	-0.54	-1.65*	-2.24**	-2.24**	-2.24**	-2.76***	**-2.78*** **	-1.56
Log likelihood	-32.85	-31.86	-30.85	-31.55	-31.65	-27.49	**-27.43**	-32.98

*** p-value < 0.1, **p-value < 0.05, *** p-value < 0.01**

The results for Germany seem to be quite different from those for the United States. Here, on a stand-alone basis, yield spread is not clearly the strongest predictor for recessions. Furthermore, the lead times are very different. For yield spread, the best model is the one predicting only one quarter ahead, while for consumer confidence and news sentiment these are 8 and 7 quarters ahead, respectively.

Furthermore, for the United States, the second news factor counting the total number of articles was more accurate. However, for Germany, the news sentiment factor, containing the fraction of articles in the negative business news section is significantly more accurate. This is unfortunate, since this is the variable with a lot of missing observations. Adding consumer confidence and news sentiment to the model seems to improve the forecasts, especially at longer lead times. This would confirm the second hypothesis of the paper. The model with the highest pseudo R^2 overall is the model containing all three variables with the lead time of 7 quarters ahead.

Since it is less clear for Germany which are the best models, the AIC is calculated for every model, at every forecast horizon, using the same number of observations (45). The results are presented in table 27.

Table 27. Akaike Information Criteria (AIC) for the different Probit models in Germany

Probit model	**Forecast horizon (*k* = number of quarters)**							
	k=1	***k=2***	***k=3***	***k=4***	***k=5***	***k=6***	***k=7***	***k=8***
Yield								
AIC	64.37	63.76	62.92	62.14	62.35	62.06	62.36	**62.43**
Conf								
AIC	59.02	58.80	60.08	61.91	63.16	63.79	64.30	64.54
News								
AIC	64.63	63.88	62.44	60.72	61.16	60.46	59.20	64.31
ΔNewsTot								
AIC	66.05	65.98	66.18	66.18	65.35	65.31	65.52	64.32
Yield/Conf								
AIC	**56.68**	55.08	56.80	59.55	62.78	63.24	63.43	64.73
Yield/News								
AIC	68.06	66.53	63.98	60.88	61.95	62.40	60.15	64.83
Yield/Conf/News								
AIC	57.20	**50.79**	**48.41**	**47.86**	**55.43**	**58.61**	**55.30**	66.20

For almost every forecast horizon, the model containing all three variables has the lowest AIC. Only for k = 1, the model without news sentiment performs better, and for k = 8 the model with solely yield spread as a predictor forecasts best. When looking at the three models containing a single variable, consumer confidence is most accurate at shorter lead times (k = 1 to k = 3), news sentiment for average lead times (k = 4 to k = 7) and yield spread for the long lead times (k = 8). Therefore, hypothesis 1 is not true for Germany. The second hypothesis does seem to be confirmed, since the full model performs best at most forecast horizons. However, this is only the in-sample fit, so out-of-sample performance has to be evaluated again as well. Since there are only 52 observations, the sample is split into an estimation period of 42 and a forecasting period of 10 observations. The results from calculating the RMSE are shown in table 28.

Table 28. Root Mean Squared Error (RMSE) for the Probit models in Germany out-of-sample

Probit model	**Forecast horizon (*k* = number of quarters)**							
	k=1	***k=2***	***k=3***	***k=4***	***k=5***	***k=6***	***k=7***	***k=8***
Yield								
RMSE	0.231	0.232	0.230	0.222	0.214	0.206	0.198	0.195
Conf								
RMSE	**0.187**	**0.197**	0.208	0.214	0.218	0.232	0.231	0.217
News								
RMSE	0.231	0.198	**0.192**	**0.211**	**0.208**	0.199	0.185	0.205
ΔNewsTot								
RMSE	0.204	0.208	0.211	0.213	0.217	0.215	0.212	0.210
Yield/Conf								
RMSE	0.225	0.232	0.239	0.235	0.230	0.269	0.260	0.230
Yield/News								
RMSE	0.259	0.226	0.215	0.223	0.211	**0.190**	**0.172**	**0.191**
Yield/Conf/News								
RMSE	0.252	0.213	0.209	0.229	0.224	0.245	0.204	0.227

Again, the results are much less consistent than for the United States. For shorter lead times (k = 1 and k = 2), the model with solely consumer confidence predicts most accurately, and at average lead times (k = 3 to k = 5), the model with just news sentiment performs best. For longer lead times (k = 6 to k = 8), the model combining yield spread and news sentiment provides the best forecast. This is also the model with the lowest RMSE overall, at a horizon of 7 quarters ahead. To see whether this model performs significantly better than the model with solely yield spread, the Diebold-Mariano test is performed again. The results are shown in table 29.

Table 29. Diebold-Mariano test for the Probit models in Germany

Probit model	**Difference**	**DM statistic**	**P-value**
Yield - Yield/News	0.01111	1.043	0.297

Similar as for the United States, the model including news sentiment does not perform *significantly* better than the model containing solely yield spread as a predictor.

Overall, the RMSE is much lower for the United States than for Germany. Furthermore, the coefficients are more significant and the pseudo R^2 higher. The AIC cannot be compared directly, because the number of observations differs. However, it is quite clear that the forecasts are much more accurate for the United States than for Germany. Therefore, the third hypothesis can be accepted.

5.6 Using the models in practice

The results of this research are very useful for practitioners, because forecasts usually have to be made for different forecast horizons. Tables 30 and 31 give a concise overview of which model to use for a specific forecast horizon, based on whether the goal is to predict recessions (Probit model) or accurately forecast economic growth (simple regressions). The chosen models are those with the lowest RMSE for the out-of-sample forecasts.

Table 30. Best-performing models for the US

	Forecast horizon (*k* = number of quarters)							
	k=1	***k=2***	***k=3***	***k=4***	***k=5***	***k=6***	***k=7***	***k=8***
Simple regressions	ΔConf	ΔNews	NewsTot	ΔNews	ΔNews	ΔNews	ΔNews	ΔConf
Probit model	Yield/ NewsTot	Yield/ ΔConf/ NewsTot	Yield/ ΔConf/ NewsTot	Yield/ ΔConf/ NewsTot	Yield/ NewsTot	Yield	Yield	Yield

Table 31. Best-performing models for Germany

	Forecast horizon (*k* = number of quarters)							
	k=1	***k=2***	***k=3***	***k=4***	***k=5***	***k=6***	***k=7***	***k=8***
Simple regressions	Conf	Conf	Yield	News	Conf	ΔNewsTot	Yield/ Conf/ News	Yield/ News
Probit model	Conf	Conf	News	News	News	Yield/ News	Yield/ News	Yield/ News

CHAPTER 6 Conclusion

The goal of this thesis was to provide a structured comparison of the indicators yield spread, consumer confidence and news sentiment, as well as combinations of the three, in their usefulness to forecast economic growth and predict recessions in the United States and Germany. After looking at the descriptive statistics and normality tests, the variables were tested for unit roots and cointegration, before proceeding with estimation of the actual models. For economic growth forecasts, simple regressions and vector autoregression models were used. To predict recessions, Probit models were fitted. The in-sample fit was evaluated using AIC and out-of-sample forecasts were made and evaluated using RMSE.

The first hypothesis states that on a stand-alone basis, yield spread is the strongest predictor for recessions. As seen in section 5.4.1, this is clearly true for the United States. The pseudo R^2 is much higher and the log likelihood lower for the model containing yield spread than those containing consumer confidence and news sentiment, for most forecast horizons. The coefficient for yield spread is significant for every lead time, which is not the case for consumer confidence and news sentiment. However, when taking a look at Germany, the hypothesis is not so easily accepted. When comparing the AIC values, consumer confidence is most accurate at shorter lead times, news sentiment for average lead times, and yield spread only for the longest lead time. Therefore, whether hypothesis 1 is accepted depends on the country of interest.

Hypothesis 2 states that adding consumer confidence and news sentiment to the models significantly improves the forecasting power. Looking at the simple regressions performed to forecast economic growth, the out-of-sample performance for the US is stronger for the models containing a sole variable than the models that combine multiple variables. For Germany, the same is true, except for very long lead times. For the Probit models, out-of-sample performance does become better in the US by adding consumer confidence and news sentiment, but only at short lead times. For Germany, it is the other way around. At long lead times, the full model performs better again. Therefore, it is not clear that adding the variables significantly improves the forecasting power, so hypothesis 2 is not accepted.

The third and last hypothesis states that the forecasting models perform better for the United States than for Germany. When looking at the simple regression, the model with the lowest RMSE for the US (0.172) performs slightly better than the model with the lowest RMSE for Germany (0.198). However, for the Probit models, the difference is much bigger. Where the RMSE for the US models is always below 0.05, for the German models it is around 0.2. Therefore, the third hypothesis can be accepted.

To give a good answer to the research question "*Which combination of yield spread, consumer confidence, and news sentiment produces the most accurate forecasts of economic growth and recession predictions in the US and Germany?"* is quite hard, because it varies between the models, the countries, and whether in-sample fit or out-of-sample performance is evaluated. Since out-of-sample performance is considered most important according to practitioner forecasters, the largest weight is given to these. In general, most models that perform well far ahead include at least yield spread, and there are several cases where including consumer confidence and/or news sentiment improves the model.

Naturally, there are several limitations to this research. First of all, because of the missing observations for the news variable in Germany, the remaining sample size is rather small. This reduces the power of the models and increases the margin of error. Furthermore, the forecasting period for the United States is 2010Q3-2018Q4. However, the country was not in recession in any of these quarters. Therefore, the recession predictions are only assessed whether they are close to 0, but not whether they are close to 1 when the country is actually in recession. Third, the simple regressions as well as the Probit models are essentially linear, whereas the relationship between the indicators and the dependent variable might be non-linear. Lastly, the frequency in which the variable of interest and the relevant information are available is different.

From this research flow several potential topics for further research. First of all, it might be very interesting to see whether the same relationship between yield spread (and consumer confidence and news sentiment) and recessions holds in developing countries. This was not attempted in this thesis due to data availability, but might be very interesting to explore in the future. It might also be interesting to look at a small, open country, such as The Netherlands. Furthermore, there is a lot of current research on modern economic indicators such as Bloomberg and Google trends search volume. These could be very useful as a substitute for the news sentiment factor or as an addition to the model as a whole. Lastly, to tackle the limitation that the relationships might not be linear, artificial neural network models could be estimated since these can more accurately represent the data. To tackle the different frequencies of the data, Mixed Data Sampling (MIDAS) regressions could be used (Ghysels, Santa-Clara & Valkanov, 2004).

REFERENCES

Amadeo, K. (2019, March 20). History of Recessions in the United States. Retrieved May 23, 2019, from https://www.thebalance.com/the-history-of-recessions-in-the-united-states-3306011

Batchelor, R. (2001). Confidence indexes and the probability of recession: a Markov switching model. *Indian Economic Review*, *36*(1), 107-124.

Batchelor, R., & Dua, P. (1997). Consumer Confidence And The Probability Of Recession: A Markov Switching Model. (No. 47). Centre for Development Economics, Delhi School of Economics

Bernard, H., & Gerlach, S. (1998). Does the term structure predict recessions? The international evidence. *International Journal of Finance & Economics*, *3*(3), 195-215.

Blood, J.D., & Philips, P.C.B. (1995). Recession Headline News, Consumer Sentiment, the State of the Economy and Presidential Popularity: A Time Series Analysis 1989-1993. *International Journal of Public Opinion Research*, *7*(1), 2-22

Carrera, J. (2019). The Shape of the U.S. Yield Curve: A Cautionary Tale. Retrieved May 29, 2019, from https://www.allianz.com/en/economic_research/publications/specials_fmo/Yieldcurve_15052019.html

Chan, W. S. (2003). Stock price reaction to news and no-news: drift and reversal after headlines. *Journal of Financial Economics*, *70*(2), 223-260.

Davies, G. (2019, February 03). Why is Germany flirting with recession? Retrieved May 23, 2019, from https://www.ft.com/content/15a2b9b2-2490-11e9-8ce6-5db4543da632

Davis, A. K., Piger, J. M., & Sedor, L. M. (2006). *Beyond the numbers: An analysis of optimistic and pessimistic language in earnings press releases* (No. 2006-005). St. Louis, MO: Federal Reserve Bank of St. Louis.

Davis, E. P., & Fagan, G. (1997). Are financial spreads useful indicators of future inflation and output growth in EU countries?. *Journal of Applied Econometrics*, *12*(6), 701-714.

Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica: Journal of the Econometric Society*, 1057-1072.

Dotsey, M. (1998). The predictive content of the interest rate term spread for future economic growth. *FRB Richmond Economic Quarterly*, *84*(3), 31-51.

Duarte, A., Venetis, I. A., & Paya, I. (2005). Predicting real growth and the probability of recession in the Euro area using the yield spread. *International Journal of Forecasting*, *21*(2), 261-277.

Engle, R.E., and Granger, C.W.J., (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica* 55, 251-276

Estrella, A., & Hardouvelis, G. A. (1991). The term structure as a predictor of real economic activity. *The Journal of Finance*, *46*(2), 555-576.

Estrella, A., & Mishkin, F. S. (1996). The yield curve as a predictor of US recessions. *Current issues in Economics and Finance*, *2*(7).

Fama, E. F. (1986). Term premiums and default premiums in money markets. *Journal of Financial Economics*, *17*(1), 175-196.

Fang, L., & Peress, J. (2008). Media Coverage and the Cross-Section of Stock Returns. *The Journal of Finance 64(5):2023-2052*.

Federal Reserve Bank of St. Louis. (2019). *NBER based Recession Indicators for the United States from the Period following the Peak through the Trough* [USREC], retrieved from FRED; https://fred.stlouisfed.org/series/USREC, May 12, 2019.

Filardo, A. J. (1999). How reliable are recession prediction models?. *Economic Review-Federal Reserve Bank of Kansas City*, 84, 35-56.

Fuhrer, J. C. (1993). What role does consumer sentiment play in the US macroeconomy?. *New England Economic Review*, (Jan), 32-44.

Garcia, D. (2013). Sentiment during recessions. *The Journal of Finance*, *68*(3), 1267-1300.

Garner, C. A. (1991). Forecasting consumer spending: Should economists pay attention to consumer confidence surveys?. *Economic Review*, *76*(3), 57-71.

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2004). The MIDAS touch: Mixed data sampling regression models. *UCLA: Finance*. Retrieved from https://escholarship.org/uc/item/9mf223rs

Gotthelf, N., & Uhl, M. W. (2018). News sentiment: A new yield curve factor. *Journal of Behavioral Finance*, 1-11.

Hall, R., Feldstein, M., Frankel, J., Gordon, R., Romer, C., Romer, D., & Zarnowitz, V. (2003). The NBER's recession dating procedure.

Harlan, C. (2014, October 30). The U.S. recovery is frustrating - but it's the envy of the advanced world. Retrieved May 23, 2019, from https://www.washingtonpost.com/news/wonk/wp/2014/10/30/the-u-s-recovery-is-frustrating-but-its-the-envy-of-the-advanced-world/?utm_term=.b1f9222c1f00

Harvey, C. R. (1988). The real term structure and consumption growth. *Journal of Financial Economics*, *22*(2), 305-333.

Haubrich, J. G., & Dombrosky, A. M. (1996). Predicting real growth using the yield curve. *Economic Review*, *32*(1), 26-35.

Hester, J. B., & Gibson, R. (2003). The economy and second-level agenda setting: A time-series analysis of economic news and public opinion about the economy. *Journalism & Mass Communication Quarterly*, *80*(1), 73-90.

Howrey, E. P. (2001). The predictive power of the index of consumer sentiment. *Brookings Papers on Economic Activity*, *2001*(1), 175-207.

Hu, Z. (1993). The yield curve and real activity. *Staff Papers*, *40*(4), 781-806.

Jarque, C. M., and A. K. Bera. 1987. A test for normality of observations and regression residuals. International Statistical Review 2: 163–172.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of economic dynamics and control*, *12*(2-3), 231-254.

Kessel, R.A. (1965). The Cyclical Behavior of the Term Structure of Interest Rates, *National Bureau of Economic Research, Occasional Paper No. 91*.

Kurtz, H. (1990). Is the economy suffering from media malady. *Washington Post*, *28*, 1.

Laurent, R. D. (1988). An interest rate-based indicator of monetary policy. *Economic Perspectives*, (Jan), 3-14.

Matsusaka, J. G., & Sbordone, A. M. (1995). Consumer confidence and economic fluctuations. *Economic Inquiry*, *33*(2), 296-318. ‘

Mazurek, J., & Mielcová, E. (2017). Is consumer confidence index a suitable predictor of future economic growth? an evidence from the USA. *E a M: Ekonomie a Management*, *20*(2), 30-45.

Moneta, F. (2005). Does the yield spread predict recessions in the Euro area?. *International Finance*, *8*(2), 263-301.

Mourougane, A., & Roma, M. (2003). Can confidence indicators be useful to predict short term real GDP growth?. *Applied Economics Letters*, *10*(8), 519-522.

Neftici, S. N. (1982). Optimal prediction of cyclical downturns. *Journal of Economic Dynamics and Control*, *4*, 225-241.

Newey, W. K., and K. D. West. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica 55*, 703–708.

Niederhoffer, V. (1971). The analysis of world events and stock prices. *The Journal of Business*, *44*(2), 193-219.

OECD (2019), Consumer confidence index (CCI) (indicator). doi: 10.1787/46434d78-en (Accessed on 13 May 2019)

Plosser, C. I., & Rouwenhorst, K. G. (1994). International term structures and real economic growth. *Journal of Monetary Economics*, *33*(1), 133-155.

Soroka, S. N. (2004). No news is good news (or at least most news is bad news): asymmetric responses to economic information. In *European consortium for political research joint sessions of workshops* (pp. 2-20).

Stock, J. H., & Watson, M. W. (1993). A procedure for predicting recessions with leading indicators: econometric issues and recent experience. In *Business cycles, indicators and forecasting*(pp. 95-156). University of Chicago Press.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, *62*(3), 1139-1168.

Walsh, C. E. (1993). What caused the 1990-1991 recession?. *Economic Review-Federal Reserve Bank of San Francisco*, (2), 33-48.

Wang, S., & Chaovalitwongse, W. A. (2010). Evaluating and Comparing Forecasting Models. *Wiley Encyclopedia of Operations Research and Management Science*.

Wigglesworth, R., & Rennison, J. (2019). Has the yield curve predicted the next US downturn? Retrieved May 29, 2019, from https://www.ft.com/content/cf9eb29a-5220-11e9-9c76-bf4a0ce37d49

Wright, J. H. (2006). The yield curve and predicting recessions. *FEDs Working Paper No. 2006-7*. Available at SSRN: https://ssrn.com/abstract=899538

Wu, H. D., Stevenson, R. L., Chen, H. C., & Güner, Z. N. (2002). The conditioned impact of recession news: A time-series analysis of economic communication in the United States, 1987–1996. *International Journal of Public Opinion Research*, *14*(1), 19-36.