KOF

Measuring Media Sentiment – Essays on Its Impact on the Economy and the Financial Markets

Matthias W. Uhl

KOF Dissertation Series No 12, 2011

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Imprint

Publisher

KOF Swiss Economic Institute, ETH Zurich

DISS. ETH NO. 20119

MEASURING MEDIA SENTIMENT – ESSAYS ON ITS IMPACT ON THE ECONOMY AND THE FINANCIAL MARKETS

A dissertation submitted to

ETH ZURICH

For the degree of

DOCTOR OF SCIENCES

presented by

MATTHIAS W. UHL

M.Sc. in Economic and Social History, University of Oxford, U.K.
Born 04 April 1981
Citizen of Germany

accepted on the recommendation of

Prof. Dr. Jan-Egbert Sturm, examiner
Prof. Dr. Didier Sornette, co-examiner

2011

Preface

This thesis was written while I was a researcher in the international business cycles section at the KOF Swiss Economic Institute at the Swiss Federal Institute of Technology (ETH) Zurich. I am particularly thankful to my advisor Prof. Jan-Egbert Sturm who gave me exceptional freedom for my research endeavors, while providing motivating and constructive feedback throughout my thesis. I am also indebted to my co-advisor Prof. Didier Sornette who supported this thesis with fruitful feedback from the early stages.

Further, I would like to thank my colleagues at the ETH Zurich and in particular of the international business cycles section of the KOF Swiss Economic Institute for great help and advice, stimulating discussions, and their kind collaboration, in particular Yngve Abrahamsen, Christian Busch, Edward Fridael, Georges Harras, Jochen Hartwig, Andrea Lassmann, Michael Lamla, Thomas Maag, and Boriss Siliverstovs. Further thanks go to Katharina Bloch for helping with many administrative issues. I also thank discussants and participants of conferences and seminars as well as referees during the past years. Thanks also go to participants of the advanced doctoral courses that I took at the Study Center Gerzensee for inspiring discussions. This empirical work would not have been possible without the collaboration of Thomson Reuters and MediaTenor. I thank both companies for supplying the respective datasets.

Last, but not least, I thank my family and friends for their encouragement and support throughout the years. Most of all, I would like to thank my wife Elise for her unbounded love, patience and trust in me. Without her, I would not be who and where I am today.

Matthias W. Uhl

Für Elise

Contents

List of Figures

List of Tables

Summary of the Thesis

This thesis considers sentiment in the print, TV and financial markets news media in order to evaluate if and how sentiment can explain and predict consumer and investor behavior. In the empirical section, we introduce novel and unique datasets in order to show with various models the explanatory and predictive power of media sentiment.

The first chapter introduces the subject matter, sets out the methodological framework, and introduces the datasets. The second chapter tests the models of Carroll et al (1994) and Sommer (2007) by adding news sentiment to their consumption behavior equation. It is shown that Autoregressive Moving Average (ARMA) models are suited for modeling private consumption with macroeconomic and sentiment variables. The introduced news sentiment index was created from over 100,000 newspaper articles, and it is shown how such an index can be created. The third chapter takes the nowcasting approach from Kholodilin et al (2010) and Schmidt and Vosen (2011) by showing that TV sentiment is a valid variable to nowcast private consumption. The TV sentiment index was created from over 10,000 TV news shows. In the last chapter, we examine the explanatory power of sentiment from Reuters news on stock returns with Vector Autoregression (VAR) models, accounting for the findings of Tetlock (2007), Brown and Cliff (2005) as well as Menkhoff and Rebitzky (2008). In total, we consider over 3.6 million Reuters news pieces.

The main results are the following: first, news sentiment is a valid variable to add to private consumption behavior models, and news sentiment has similar explanatory power than personal income. Nevertheless, consumer sentiment is the more powerful variable compared to news sentiment. Second, TV sentiment is more suited to nowcast private consumption than consumer sentiment. And, third, we show that both positive and negative Reuters sentiment matter for explaining stock returns. The sentiment effect is measurable and present over several months, and not only days, as previously assumed in the literature.

Zusammenfassung der Dissertation

Diese Dissertation befasst sich mit Sentiment in Zeitungs-, Fernseh-, und Finanzmarktnachrichten. Es wird getestet, ob und wie Sentiment Konsumenten- und Investorenverhalten erklären und vorhersagen kann. In jedem der empirischen Kapitel stellen wir einen neu- und einzigartigen Datensatz vor, um mit verschiedenen Modellierungsansätzen die erklärenden und vorhersagenden Eigenschaften von Mediensentiment zu untersuchen.

Das erste Kapitel umfasst einleitende Worte zur Thematik im allgemeinen sowie zum methodologischen Ansatz dieser Dissertation und es beschreibt die verwendeten Datensätze. Das zweite Kapitel testet die Modelle von Carroll et al (1994) und Sommer (2007) indem "News Sentiment" zu deren Konsumentenverhaltensgleichungen hinzugefügt wird. Es wird gezeigt, dass Autoregressive Moving Average (ARMA) Modelle sich am eignen, um den privaten Konsum mit makroökonomischen und Sentiment Variablen zu modellieren. Der eingeführte News Sentiment Index wurde mit Informationen von über 100,000 Zeitungsartikeln konstruiert und es wird die Vorgehensweise der Konstruktion des Indexes aufgezeigt. Das dritte Kapitel verwendet die "Nowcast" Methode von Kholodilin et al (2010) und Schmidt und Vosen (2011) um zu zeigen, dass Sentiment aus Fernsehnachrichten (TV Sentiment) eine valide Variable zum Vorhersagen bzw. "Nowcasten" des privaten Konsums ist. Der TV Sentiment Index wurde durch Information von über 10,000 Fernsehnachrichtensendungen konstruiert. Im letzten Kapitel wird die Aussagekraft von Sentiment aus Reuters Finanzmarktnachrichten auf Aktienkurse mit Vector Autoregressions (VAR) Modellen untersucht, indem wir insbesondere die Resultate von Tetlock (2007), Brown und Cliff (2005) sowie von Menkhoff und Rebitzky (2008) in Betracht ziehen. Für dieser Studie wurden über 3.6 Millionen Reuters Nachrichten analysiert.

Die Hauptresultate sind folgende: erstens, Sentiment aus Zeitungsnachrichten ist eine geeignete Variable, um sie zu Konsumentenverhaltensmodellen hinzuzufügen, wobei News Sentiment ähnliche erklärende Eigenschaften aufweist wie das Einkommen von Haushalten. Nichtsdestotrotz ist das Konsumentensentiment aussagekräftiger als News Sentiment. Zweitens, TV Sentiment is besser geeignet als Konsumentensentiment, um den privaten Konsum zu "nowcasten."

Und, drittens, zeigen wir, dass sowohl positives als auch negatives Reuters Sentiment aussagekräftig ist, um Aktienkursveränderungen zu erklären. Der Sentiment Effekt ist für mehrere Monate präsent und messbar, und nicht nur für ein paar Tage, wie in der bisherigen Literatur angenommen.

1 Introduction

"Everything is relative, and only that is absolute."

Auguste Comte (1798–1857)

1.1 Methodological Framework

In today's globalized world, we have become ever more dependent on the media. Hardly do people have to build their own opinions based on critical thinking and research if they watch, listen, or read the news. Behr and Iyengar (1985) show that public news influence public opinion. Mass opinion convergence has thus become more likely, as people tend to "take in" what the media tell them – directly or indirectly. Ultimately, if Behr and Iyengar's hypothesis holds true, what effect would that have on the economy and the financial markets? Can we establish a link between the psychological domain, i.e. media sentiment, and the economic and financial realm? Traditionally, psychology is considered a qualitative research area whereas economics and finance are considered to be quantitative. Key to solving this question is to discern between qualitative and quantitative research in order to establish a direct link between these so different areas. We therefore consider it plausible to quantify the qualitative realm of psychology, and, therefore, sentiment.

Qualitative research has its origin in the mid-twentieth century when psychologists sought to view their discipline more holistically. Understanding what it means to be human was central to their concern, opposing quantitative research, based on numerical facts. Characteristics such as love, hate, hope, the self, individuality, nature, creativity, and existence moved center-focus in qualitative research. Maykut and Morehouse (1994) even claim that qualitative research methods started to come into existence at the margins of acceptable science – from Sigmund Freud to Carl Rogers and Mary Ainsworth. Humanistic psychology is an approach to the human being by qualitative means, not by quantitative. A qualitative pioneer, Carl Rogers (1951) formulated his theory about a continually changing world of experience and perception each with its accompanying attributes. Individuals, Rogers claimed, are shaped by a fluid but consistent conceptual pattern of perceptions of characteristics as they interact with others and face different environments. In this context, Taylor and Bogdan (1998) have coined the relevant phrase: "The important reality is what people perceive it to be." Perception of quantitative domains, i.e. economics and finance, is thus a central aspect of this thesis. How can we draw a link then between perception and sentiment, qualitative characteristics, and quantitative domains, such as economics and finance?

Economic and financial performance can easily be expressed. One can consider GDP and population growth rates, asset prices and firm earnings along with numerous other variables. All of these variables are quantitatively educible. It is not in the interest of economic management to have ambiguous measures. Numbers are unambiguous. However, people do not think in numbers when they perceive anything. For example, our mind does not think '1' or '-1' when we see a car whether we like or dislike it. It would be easier though for other people to understand whether we find the car to be either '-1' or '1', as every person has probably a different understanding of the words 'nice' or 'ugly.' 'Nice' could be equivalent to '1' for some people, but only '0.8' or '0.9' for others. There lies the crucial difficulty if we want to generalize quantitatively about people's perceptions: they form a subjective and so mysterious field to investigate. Armstrong (1961) makes clear that whenever decisions are made, be it in politics, economics, or even the household, perceptions always play a role, yet they can vary greatly among the parties concerned. From this example, it seems important to understand the way that economic awareness is evoked through how news are written, reflecting sentiment about the present state of the economy and the financial markets.

When attempting to portray qualitative data quantitatively, we need to distinguish between different people's understandings of qualitative meanings, such as 'nice' or 'bad', and apply these to a defined quantitative pattern. In this thesis, we attempt to establish a relationship between the domains of economics, finance and psychology, or, put differently, between quantified sentiment portrayed through the media and hard economic facts. In order to achieve this, we introduce novel sentiment datasets gathered from various media outlets and with the aid of cutting-edge technology.

1.2 Data and Structure

This empirical thesis is structured as follows. The second chapter deals with explaining private consumption growth in the United States with a newly developed news sentiment index, and comparing the new index to an established consumer sentiment index, i.e. the University of Michigan Index of Consumer Sentiment (ICS). This exercise serves as an introduction to media sentiment research to the extent that the news sentiment index is constructed without the help of professional media sentiment providers, as is the case in chapters three and four. In the second chapter, we introduce a novel dataset with a news sentiment index that is constructed from a selection of over 100,000 newspaper articles from the economics section of two of the most read newspapers in the US. With the newly created news sentiment index, we construct private consumption behavior models with consumer sentiment data and other macroeconomic data, such as personal income, consumer and house prices, interest rates and unemployment figures. Chapter three takes the idea of chapter two further, and it introduces a novel sentiment variable gathered from TV news broadcasts in the US. This dataset was obtained from Mediatenor, a professional sentiment provider. Similar to chapter two, we implement TV sentiment into a private consumption behavior model in order to test its nowcasting ability. Last, but not least, chapter four considers the impact of sentiment on the financial markets, and in particular stock returns. The extensive dataset was obtained from Thomson Reuters, which comprises sentiment from almost four million Reuters news articles that relate to the examined stock index.

1.3 Contribution and Findings

The main contribution of this thesis is that we show with various approaches that sentiment is present in today's media, and that media sentiment has an influence on both consumers and financial market participants. We further prove that sentiment gathered from various sources can explain and predict both private consumption and financial market developments. More specifically, in the second chapter, we show with ARMA-models that news and consumer sentiment, when combined with other macroeconomic variables, achieve statistically significant results to explain changes in private consumption. Consumer sentiment, measured by the ICS, adds more explanatory power and statistical significance than news sentiment when tested individually and jointly. In this chapter, we conclude that news sentiment is a useful variable to add in consumer behavior models, especially when coupled with consumer sentiment and other select macroeconomic variables.

The third chapter extends the existing literature by introducing a new sentiment variable obtained from television news. As the average American watches over five hours of television a day, and television news have the greatest share of news sources among Americans, we consider sentiment from over 10,000 TV news broadcasts in the US in order to test it in a nowcasting framework. We show with ARMA-models that both consumer and TV sentiment, along with other variables, such as financial variables, add value in nowcasting private consumption. We make the case that TV sentiment has slightly more information content than consumer sentiment in order to nowcast private consumption.

In the fourth chapter, we examine the explanatory power of fundamental macroeconomic and behavioral factors with regards to stock returns of the Dow Jones Industrials Index. With a sentiment dataset from over 3.6 million Reuters news articles, we find significant correlations between Reuters sentiment and stock returns. We show with vector autoregression (VAR) models that Reuters sentiment can explain changes in stock returns better than macroeconomic factors. Considering positive and negative sections of Reuters sentiment, we find that both sentiments add value to the model. The main contribution of this chapter is that the news sentiment effect is measurable over months, and not only days, as eluded in the existing literature on news sentiment.

2 Explaining US Consumer Behavior with News Sentiment[1]

2.1 Introduction

2.1.1 Introduction and Summary

Many attempts have been made to explain private consumption growth in the United States because it makes up around 70% of its gross domestic product. Hence, the US are traditionally considered a consumer-oriented economy. Theoretically, as Hayashi (1982) noted, for example, consumer models were a function of hard facts, e.g. income, and soft facts, expectations about income and wealth, that attempted to explain private consumption. More recent studies in this domain, such as those of Doms and Morin (2004) and Sommer (2007), focused more on the idea of a possible influence of sentiment on the consumer. In their studies, they used the well-known Index of Consumer Sentiment (ICS) from the University of Michigan that was developed in the 1940s by George Katona, and then first released in 1952. Katona was also one of the first to start the debate on psychology and consumer economics. In one of his numerous works, Katona (1974) identified, in addition to traditional macroeconomic variables, such as personal income, the unemployment rate, interest rates, house prices, and others, that private consumption was also based on a subjective factor that he called "consumer sentiment." With rising sentiment, Katona noted, consumers increase their spending, and with decreasing sentiment, households trim their consumption.

[1] This chapter is based on Uhl (2011a).

In this chapter, we take the idea that sentiment is relevant for the consumer further by considering sentiment from the economics section of widely-read newspapers in the US. We attempt to show that sentiment among consumers can not only be identified by surveys, i.e. through the ICS as done in many studies, and thus by asking the consumers directly about their sentiment, but also by considering how newspapers portray news about the economy that potentially influence the consumer and therefore shape her sentiment. Hence, in this exercise, we want to show an innovative way of extracting sentiment from newspapers rather than using survey-based consumer sentiment. We call this kind of sentiment "news sentiment." Therefore, we introduce a novel data set with a news sentiment index that was constructed from over 100,000 newspaper articles from two top ten US newspapers by circulation. Briefly, we show how to construct such a news sentiment index with publicly available information.

By formulating ARMA-models, we show that soft facts like news and consumer sentiment (i.e. the ICS), both individually and combined with other variables, i.e. the hard facts, such as personal income, consumer and house prices, unemployment, and interest rates, achieve statistically significant results to explain changes in private consumption. We make three distinct findings. First, both consumer and news sentiment are statistically significant and they add explanatory power to private consumer behavior models. Second, we find that consumer sentiment is the more powerful sentiment variable. This finding is manifested in the ARMA-regression models analysis as well as in a standardized coefficients analysis. And, third, news sentiment performs as good as personal income in explaining private consumption, whereas consumer sentiment is more powerful in explaining private consumer behavior than personal income, inflation, the unemployment rate, interest rates and house prices combined. In general, we conclude that news sentiment is a useful variable to add to consumer behavior models, especially when coupled with consumer sentiment and other macroeconomic variables. The remainder of this section discusses the motivation of the paper and the related literature as well as the novel data set and its creation. Section 2.2 sets out the model and discusses the empirical results, while section 2.3 concludes.

2.1.2 Motivation and Literature Overview

According to the Permanent Income Hypothesis (PIH), first formulated by Friedman (1957), consumption patterns of consumers are not determined by current income but rather by their longer-term income expectations. Later, Hall (1978) contradicted this view, saying that consumption growth is unpredictable, following a random walk. Since then, many studies have emerged to tackle the issue of explaining and predicting changes in private consumption. The two contradicting views from Friedman (1957) and Hall (1978) brought the focus on expectations about longer-term income and wealth. What drives the longer-term income expectations of consumers? Hayashi (1982) formulated one solution to accounting for longer-term income expectations with the basic optimal consumption rule:

$$c_t = \alpha \left(A_t + H_t \right), \tag{2.1}$$

where c_t represents consumption at time t, and A_t is real nonhuman wealth. Real human wealth H_t is defined as the present discounted value of expected future real labor income:

$$H_t = \sum_{k=0}^{\infty} (1+\mu)^{-k} \; {}_t y_{t+k}, \tag{2.2}$$

where μ is the discount rate and ${}_t y_{t+k}$ refers to the household's expectation as of t of real, after-tax labor income at $t+k$. Hayashi (1982) points out that the rational expectations hypothesis[2] incorporates the idea that ${}_t y_{t+k} = E\left(y_{t+k} \mid I_t\right)$, where I_t is the set of information held by the household at t. Then, it is the information set that each household holds, out of which the consumption behavior is formed. How is this kind of information set made up? We want to focus on the questions how consumers make their consumption decisions and by which factors they are influenced. In order to do that, we consider more recent studies that have examined many possible variables and channels of influence that have an impact on the consumer.

Campbell and Mankiw (1989) extend the pure life-cycle / permanent-income hypothesis (PIH). As opposed to previous works, they distinguish between two kinds of consumers:

$$\Delta c_t^L = \varepsilon_t, \tag{2.3}$$

$$\Delta c_t^R = \Delta y_t^R, \tag{2.4}$$

where c_t^L refers to life-cycle consumers, c_t^R to rule-of-thumb consumers, ε_t to news received in period t about lifetime resources, and y_t to current income of private households. A crucial assumption in the Campbell-Mankiw frame-

[2] Also see Muth (1961).

work is that rule-of-thumb consumers receive a constant proportion λ of total income. Aggregate consumption is then given as follows by the combination of equations (2.3) and (2.4):

$$\Delta c_t = \lambda \triangle y_t + \varepsilon_t. \tag{2.5}$$

Thus, a combination between income – real facts – and news – emotional facts – that can possibly influence expectations of households needs to be taken into account more closely. Carroll et al (1994) examine the predictive power of consumer sentiment for future changes in consumption spending. They find that lagged consumer sentiment can partly explain current changes in household spending. Drawing on the study of Campbell and Mankiw (1989), Carroll et al (1994) are able to reject their hypothesis that lagged sentiment affects consumption growth only through the income channel, giving room for more variables that might affect consumer behavior. They claim that habit formation should be explored further to identify other channels that could possibly affect consumption growth, represented in the following form :

$$\triangle \log\, c_t = \alpha_0 + \sum_{i=1}^{N} \beta_i S_{t-i} + \gamma Z_{t-1} + v_t, \tag{2.6}$$

where $\triangle \log\, c_t$ refers to the logged private consumption growth rate, S_t to consumer sentiment (e.g. the ICS), Z_t is a vector of other variables, and v_t the error term. They leave room for speculation which other variables can be included in the vector Z_t. In another study, Acemoglu and Scott (1994) use UK data to show that confidence indicators outperform other macroeconomic variables that explain consumer behavior. In this light, they reject the Rational Expectations Permanent Income Hypothesis (REPIH) by Hall (1978) and conclude that the predictive ability of confidence indicators is inconsistent with forward-looking behavior. Lloyd (1999) finds that consumer sentiment surveys (including the ICS) perform better than professional forecasters when implemented in forecasting models for inflation and consumer expectations. In a more recent study, Carroll (2003) stresses the importance of news coverage of economic matters with respect to household's expectations, and ultimately their behavior resulting from these expectations.

Based on these initial findings that news as well as sentiment and expectations have a possible influence on consumers, we motivate this analysis. In general, we assume that the ordinary consumer is not a trained economist. The consumer obtains her information about economic conditions mainly through the news she reads. This, in turn, shapes her expectations and sentiment about future income and consumption of her household. We hypothesize that each article the consumer reads evokes a certain feeling, opinion, or emotion about the state of the subject, which can be either positive or negative. This feeling is what we call sentiment, or ultimately news sentiment, because we measure the sentiment in newspaper articles in order to get further clues about possible drivers of consumption behavior in the US. Doms and Morin (2004) examine the hypothesis that news media affect consumers' perceptions of the economy. They find that the tone and volume of economic reporting in news affect consumers. Further, they identify a short-lived effect of sentiment on consumer spending, lasting only a few months. Given their findings, we want to test whether a positive or negative tone in news reporting (i.e. news sentiment) drives consumption behavior. A few years later, Sommer (2007) formulated a benchmark model of habit formation in consumer preferences, which can explain two failures of the PIH. First, the sensitivity of aggregate consumption to predictable changes in income and to lagged sentiment, and, second, the persistence in consumption growth is higher than anticipated. Ang et al (2007) find that consumer sentiment surveys (e.g. the ICS) perform best in forecasting models of inflation as opposed to time-series, Phillips curve, and term structure forecasts. They obtain little evidence that combining forecasts produces superior forecasts to survey information alone. We thus take consumer surveys into account in explaining private consumption. Noteworthy to point out is that they hypothesize that one possibility for the better performance of survey forecasts is that the survey aggregate information is from many different sources, which are not captured by a single model. They claim that the superior information in median survey forecasts may be due to an effect that is similar to Bayesian model averaging. We take this idea of Bayesian model averaging by considering two widely read newspapers in our analysis, rather than focusing on a particular one. This gives us the advantage of not being dependent on a specific newspaper and its concomitant tone and reporting style. This also tackles the issue that Carroll et al (2010) raised, as they identify stickiness in aggregate expectations with important macroeconomic consequences, as people only occasionally pay attention to news reports.

Combined, these findings are key motivator to this study, since they hypothesize that *other variables* than solely income can affect consumer behavior. Breeden (1986), for example, found that interest rates as well as inflation are related to the expected growth rate of aggregate consumption. Additionally, Case et al (2003) find that changes in housing prices have substantial effects on private consumption. Thus, in line with the studies mentioned, we test whether these other variables influencing private consumption can be news and consumer sentiment, measured by the ICS, personal income, inflation, the unemployment rate, short-term interest rates, and the Case Shiller Home Price Index. We thus extend existing models with news sentiment and test whether this variable can add value to models that explain US private consumption.

2.1.3 Dataset and the Sentiment Classifier

We introduce a new dataset that quantifies news sentiment from the economics section of newspapers in the US from 1995 to 2009. A sentiment algorithm is used for the analysis of over 100,000 newspaper articles from the *Washington Post* (WP) and *USA Today* (UT). The sentiment algorithm distinguishes between positive and negative sentiment of newspaper articles in binary format, namely $\{-1\}$ for negative sentiment and $\{1\}$ for positive sentiment. The algorithm is based on a broad and complex database of positive and negative words and phrases. This sentiment algorithm is based on one of the most popular classifiers used in machine learning science: the Naive Bayes classifier. The sentiment algorithm java program was obtained from a free web-based sentiment algorithm provider and is tested for accuracy in this study.[3] According to Friedman et al (1997), this classifier learns the conditional probability of each attribute a_i, given its class label c, from training data. The sentiment algorithm was trained to distinguish between *positive* and *negative* sentiment from a pre-defined database of positive and negative words and phrases. The classification is then done by applying Bayes rule to compute the probability of c given the particular instance of $a_1, \ldots, a_i$, and then predicting the class with the highest posterior probability. This means that the computation is rendered

[3] See http://www.jane16.com, last accessed 20 September 2009.

feasible by making a strong independence assumption: all the attributes a_i are conditionally independent given the value of the class c. By independence, Friedman et al (1997) further note, probabilistic independence is meant, that is, a is independent of b given c whenever $Pr\left(a \mid b, c\right) = Pr\left(a \mid c\right)$ for all possible values of a, b, $and\, c$, whenever $Pr\left(c\right) > 0$. This means that in each article, every word and a combination of phrases is checked against the sentiment algorithm and classified as either positive or negative. The sentiment score is then obtained by applying Bayes' rule to the classifications that were obtained for each article individually, so that the output of either positive or negative is generated for each single article.

Lewis (1998) discusses the Naive Bayes approach in historical context by concluding that the algorithm is experiencing a renaissance owing to its broad range of usability. In an empirical study, Rish (2001) concludes that the Naive Bayes classifier is very effective in practice, even though its probability estimates are in theory less accurate than other classifiers. Hand and Yu (2001) make the case for the Naive Bayes algorithm because of its intrinsic simplicity, which means low variance in the probability estimates and thus greater estimation accuracy. Last, but not least, Kotsiantis and Pintelas (2004) show that Naive Bayes is the most flexible learning method. Its accuracy can be boosted over most methods in less time for training, justifying its use in this study.

Visual Basic programs were written in order to ease the process of dealing with the large amount of data.[4] The sentiment algorithm scans each entire article and gives an output file with the respective sentiment score of each individual article. These scores were then aggregated on a quarterly basis.

[4] See Appendix 2.4.1 for more information on these programs.

Table 2.1: News Sentiment Sources and Summary Statistics

	Washington Post	USA Today
Mean	0.6028	0.6103
Median	0.6067	0.6112
Maximum	0.7055	0.7228
Minimum	0.5093	0.5099
Std. Dev.	0.0342	0.0510
Skewness	-0.1635	0.2091
Kurtosis	4.0330	2.5665
Sum	35.5628	36.0071
Sum Sq. Dev.	0.0679	0.1508
Observations	59	59
Number of available articles from the Economics section that were examined for sentiment*	74'206	28'832
Average Daily Circulation**	604'650	1,826,622

* Source: LexisNexis and own calculations.

** Source: Audit Bureau of Circulation Survey 31/3/2010, (http://abcas3.accessabc.com/ecirc/newstitlesearchus.asp), last accessed 9 June 2010.

Table 2.1 shows summary statistics and the average daily circulation of each newspaper along with how many newspaper articles were examined for sentiment of each newspaper.[5] In total, over 100,000 individual newspaper articles were examined for sentiment. The sentiment scores from both the WP and UT show similar properties: their means are both around 0.61, and the scores range from roughly 0.5 to 0.7. Given these initial findings, we create a news sentiment index from the sentiment scores of these two newspapers by simply taking the average between the two, as follows:

$$NS = \frac{(WP + UT)}{2}, \tag{2.7}$$

where NS refers to the newly created news sentiment index.

Quarterly US private consumption data as well as consumer price index data were obtained from the *U.S. Department of Commerce Bureau of Economic Analysis* database.[6] Unemployment rate figures, short-term interest rates (3-month USD LIBOR), and the Case Shiller Home Price Index data were obtained from Thomson Reuters Datastream. The Case Shiller Home prices were adjusted for inflation from the consumer price index data.[7] Real personal income data were downloaded from the *Bureau of Economic Analysis* website.[8] The University of Michigan Index of Consumer Sentiment (ICS) data were downloaded from the University of Michigan and Thomson Reuters

[5] The examined newspapers were selected by availability from the top list of daily average circulation. The articles were downloaded from the *LexisNexis* database.

[6] See http://www.bea.gov/, last accessed 15 April 2010.

[7] Specifically, we use the S&P/Case-Shiller Home Price Index – 10-City Composite. See also S&P/Case-Shiller Home Price Indices – Index Methodology for more detailed information, available at http://www.standardandpoors.com/indices/articles/en/us/?articleType=PDF&assetID=1245190941370, last accessed 7 October 2011.

[8] See http://www.bea.gov/, last accessed 24 June 2010.

public access website.[9] The ICS is constructed from answers to five questions relating to current economic conditions of consumers as well as consumer expectations.[10] Table 2.2 shows the summary statistics of all variables used for the analysis.

2.2 Empirical Analysis

2.2.1 Modeling

In their analysis, Doms and Morin (2004) draw a chart that sets news coverage, household's sentiment and expectations as well as private consumption in relation to each other. In fig. 2.1, we have amended this information flow chart and added news sentiment, postulating that the reader, i.e. the consumer, is influenced by the news and the sentiment portrayed through news that she reads, which form her expectations and sentiment, ultimately driving changes in future consumption. In our analysis, we want to examine whether news sentiment contains information that is not captured by the ICS. Hence, we examine whether news sentiment can be used as a complement to the ICS.

In the information flow chart in fig. 2.1, consumers receive news about the economic activity, which shape their expectations and sentiment, ultimately driving future consumption behavior. Doms and Morin (2004) have identified the ICS to be a good variable to add to consumption growth models. In this paper, we want to test further which factors explain changes in private

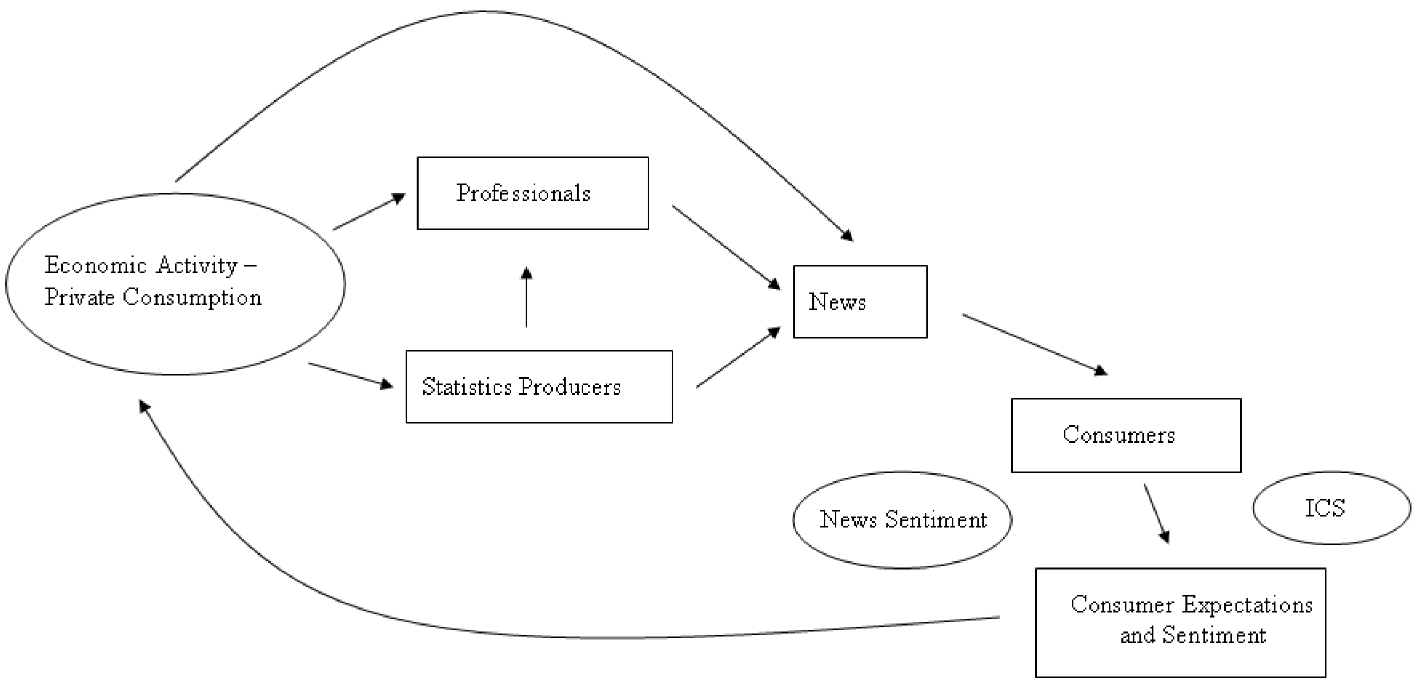
[9] See http://www.sca.isr.umich.edu/, last accessed 8 June 2010.

[10] A detailed description of the calculation of the index and the individual questions can be found on the homepage of the surveys of consumer from the University of Michigan and Thomson Reuters. See *Index Calculations,* http://www.sca.isr.umich.edu/documents.php?c=i, last accessed 8 June 2010.

Table 2.2: Summary Statistics

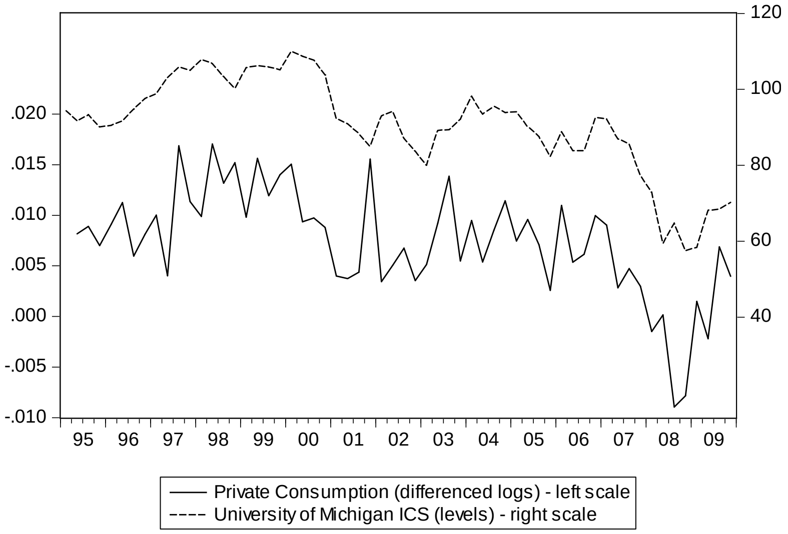
	Private Consumption (differenced logs)	University of Michigan - Index of Consumer Sentiment	News Sentiment Index	Real Personal Income (differenced logs)	Inflation - Consumer Price Index (differenced logs)	Unemployment Rate	Short-term Interest Rates (USD LIBOR 3 month) (differences)	Real Case Shiller Home Price Index (differenced logs)
Mean	0.007397	90.61019	0.607808	0.011546	0.006133	5.330508	-0.105354	0.012203
Median	0.008123	92.14476	0.60695	0.012204	0.006749	5.1	-0.04	0.018002
Maximum	0.017059	110.0553	0.67975	0.032123	0.015281	10	1.41625	0.060435
Minimum	-0.008948	57.50786	0.5355	-0.023221	-0.02438	3.9	-2.795	-0.088908
Std. Dev.	0.005314	13.05964	0.03118	0.008466	0.005624	1.273098	0.636905	0.02993
Skewness	-0.66882	-0.795807	0.178785	-1.077307	-2.868377	2.085946	-1.693685	-1.241053
Kurtosis	4.089205	3.255172	2.847865	6.675169	16.17973	7.754398	8.391231	4.753045
Jarque-Bera	7.315135	6.38761	0.371211	44.61684	507.9302	98.35543	99.65997	22.70028
Probability	0.025795	0.041016	0.830601	0	0	0	0	0.000012
Sum	0.436409	5346.001	35.86065	0.681205	0.361836	314.5	-6.2159	0.720006
Sum Sq. Dev.	0.001638	9892.15	0.056386	0.004157	0.001835	94.00508	23.52755	0.051958
Observations	59	59	59	59	59	59	59	59

Figure 2.1: Information Flow to Consumers



Note: ICS = the University of Michigan Index of Consumer Sentiment

Figure 2.2: Time-series graph of private consumption and the University of Michigan Index of Consumer Sentiment

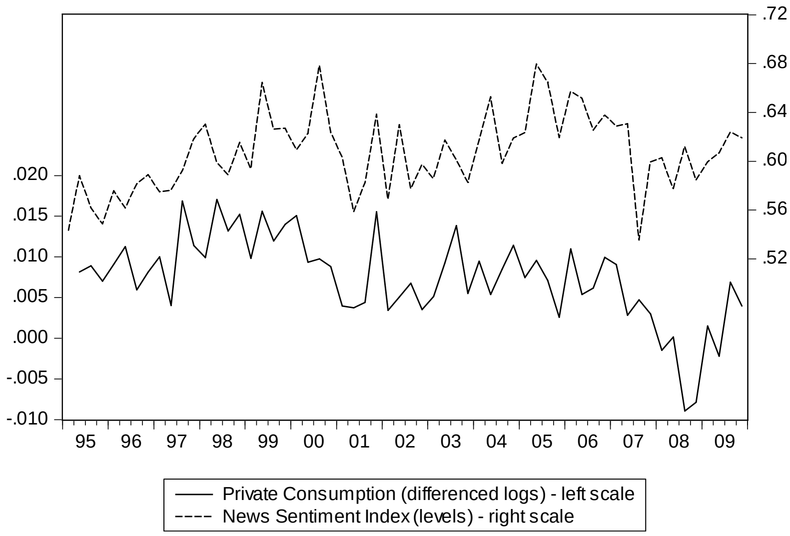


consumption. We therefore add news sentiment and the ICS to a base model in order to examine if these sentiment variables add explanatory power to private consumption behavior models.

In fig. 2.2, we show private consumption growth plotted against the ICS. A co-movement of the ICS and private consumption is apparent, especially during the crisis in 2008/09 when the decline of both consumption and consumer sentiment was most pronounced. The years preceding the recent financial crisis show steady growth rates in private consumption as well as in consumer sentiment. During the dot.com crisis, the drop in consumer sentiment was not as pronounced as in the more recent one, but the level of sentiment in the years succeeding the dot.com crisis never reached the peak of the year 2000 again in the time horizon examined here.

The news sentiment index also shows some co-movement with private consumption growth rates in fig. 2.3. As the ICS, the news sentiment index experienced its height in 2000 and the subsequent crash when the dot.com bubble

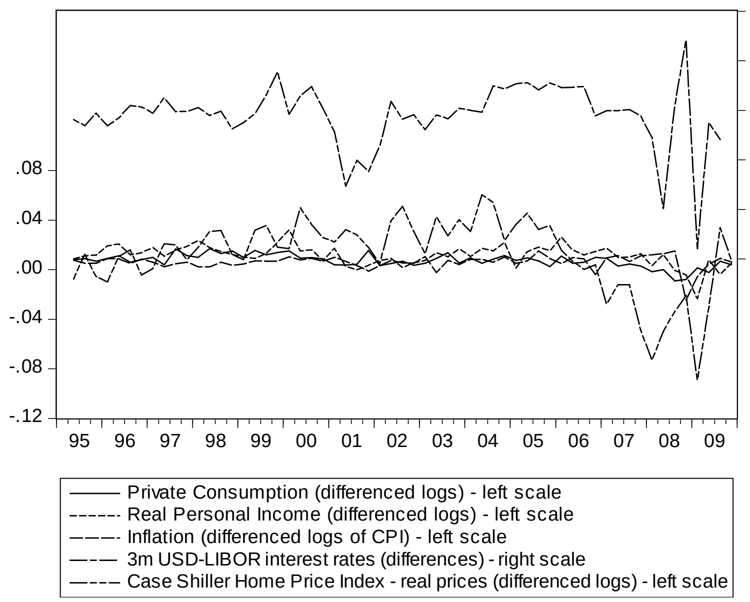
Figure 2.3: Time-series graph of private consumption and the news sentiment index



burst. However, the down periods in news sentiment are more pronounced than in the ICS, and in this light, the series appear to show greater volatility. This is partly due to the fact that we deliberately did not apply a filter to the series in order to consider the plain effect of the series and to avoid the accusation of data fitting. The news sentiment index ranges between 0.5 and 0.7, indicating that news are – in the period examined – positively biased, since the scale is from $\{-1\}$ to $\{1\}$. The identification of a positive bias in news sentiment is consistent with the phenomena that Baron (2006) identifies in news media reporting, although this is contrary to the general belief that "bad news sell."

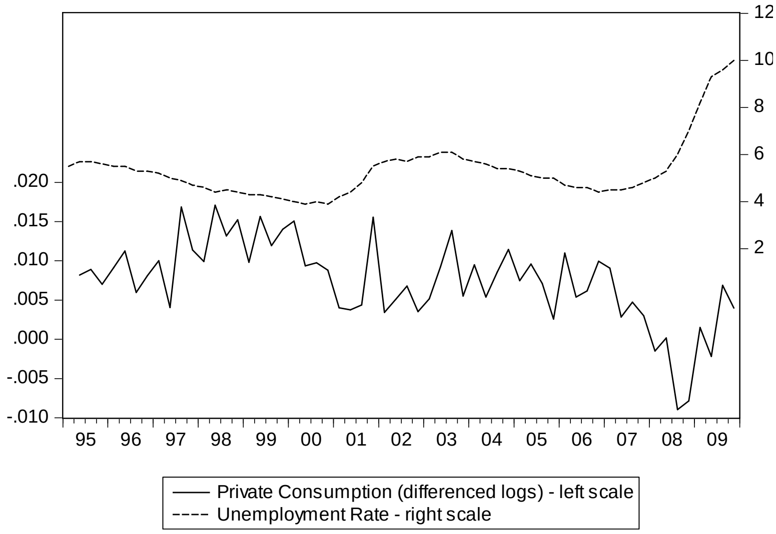
Although the two sentiment variables, namely the ICS and the news sentiment index, are central to this study, we account for other possible factors that can influence consumer behavior in our model as well, depicted by the vector Z_t, such as income, interest rates, unemployment, inflation and housing prices to proxy for changes in personal wealth.

Figure 2.4: Time-series graph of private consumption and other macroeconomic variables



In fig. 2.4, we plot personal income, inflation, real house prices as well as interest rates against private consumption growth. Real personal income, consumer prices and private consumption co-move nicely with each other. Again, for all these variables the drop in growth rates is most pronounced during the recent financial crisis when growth rates for all variables turned negative. Given that the recent financial crisis hailed from a housing market debacle in the US, this crisis is very pronounced in the Case Shiller Home Price Index, which experienced negative growth rates from as early as 2007. The low in the housing market occurred after the collapse of Lehman Brothers, between the end of 2008 and the beginning of 2009. Both the dot.com crisis of 2000 as well as the recent financial crisis of 2008/2009 are visible in the time-series of short-term interest rates. During both crises, interest rates turned negative, a sign of monetary expansion and macroeconomic turmoil. The unemployment rate is also a good proxy for the health of an economy. If the unemployment

Figure 2.5: Time-series graph of private consumption and the unemployment rate



rate is low, the economy should be in good state, while if it is high, the economy is probably in a crisis. We thus expect an inverse relationship between private consumption growth and the unemployment rate. Fig. 2.5 shows this relationship nicely. During both the dot.com crisis and the financial crisis, the unemployment rate surged. When considering the magnitude of the increase in the unemployment rate during both crises, we see that the recent financial crisis was much more severe for the real economy than the dot.com crisis in 2000. In a nutshell, the graphical interpretation shows that the variables taken into consideration are suited to explain private consumption growth, so that we can test whether these variables have similar explanatory power statistically than they have graphically.

We construct a model that is based on simple autoregressive and moving average models. As in Ang et al (2007), we use the Schwarz criterion (BIC) to determine the order of the autoregression (AR) and moving average (MA) processes. Table 2.3 shows the results of various ARMA(v,w)-models with

Table 2.3: Schwarz Information Criterion Test Results

ARMA-Lag Structure	Schwarz Information Criterion (BIC)	Inverted MA Roots*
ARMA(2,1)	-7.742469	<1
ARMA(1,2)	-7.762936	<1
ARMA(2,2)	-7.677903	<1
ARMA(3,1)	-7.791473	>1
ARMA(3,2)	-7.578827	<1
ARMA(3,3)	-7.507185	<1
ARMA(2,3)	-7.606932	<1
ARMA(1,3)	-7.745416	<1

*Note: Inverted Roots of MA process have to be smaller than 1 so that the process is stationary and invertible.

This table shows various Schwarz Information Criteria test results in order to determine the best ARMA-structure of the base model with private consumption as dependent variable and real personal income, inflation, the unemployment rate, short-term interest rates, and the Case Shiller Home Price Index (real prices) as independent variables. The most suitable ARMA-structure has the lowest BIC and was found to be of the order ARMA(1,2), as indicated.

private consumption as dependent variable and the vector Z_t as regressor, as follows:[11]

$$\triangle \log c_t = k + \alpha_v \triangle \log c_{t-v} + \beta Z_t + \theta_w \varepsilon_{t-w} + \varepsilon_t, \qquad (2.8)$$

where $\triangle \log c_t$ refers to logged private consumption growth, k is the constant term, Z_t to a vector of independent variables, such as real personal income, inflation, real house prices, unemployment, and short-term interest rates, while ε_t represents the error term.

Given the higher order of the ARMA-processes, we need to consider the inverted MA-roots, which need to be less than 1, so that the process is stationary and invertible. If the MA-roots are equal to or greater than 1, the results obtained are not reliable. Given that the ARMA(1,2)-model has the lowest BIC among all the other models, we take this model as our base model. This is

[11] We tested the various ARMA(v,w)-models with $v = \{1 \ldots 3\}$and $w = \{1 \ldots 3\}$.

in line with Sommer's (2007) findings, as he applies an ARMA(1,2) structure to modeling private consumption as well. Hence, according to Sommer (2007) and Carroll et al (2010), we formulate the model with an ARMA(1,2) process as follows:

$$\triangle \log\, c_t = k + \alpha_1 \triangle \log\, c_{t-1} + \beta Z_t + \sum_{i=1}^{2} \theta_i \varepsilon_{t-i} + \varepsilon_t, \tag{2.9}$$

Further, we include the sentiment variables in the vector S_t and combine it with the above regression equation (2.9), so that

$$\triangle \log\, c_t = k + \alpha_1 \triangle \log\, c_{t-1} + \beta Z_t + \gamma S_t + \sum_{i=1}^{2} \theta_i \varepsilon_{t-i} + \varepsilon_t, \tag{2.10}$$

where S_t refers to the news sentiment index and/or the ICS. According to Hendry (1993), we examine the base model with all variables, and, subsequently, we drop insignificant variables from the vector Z_t to narrow down the model to the most significant variables that can explain private consumption growth. We further examine the model with S_t, so that we can specifically identify the value of the ICS and the news sentiment index jointly and individually.

In order to account for possible non-stationarity in the data, we test equation (2.10) with all available variables for unit roots. Table 2.4 shows the various tests as well as the results. We tested for both a common unit root and individual unit root processes. We can reject the null hypothesis of a common unit root process with a 90% confidence level, while we can reject the null hypotheses of individual unit root processes with a 99% confidence level.

Table 2.4: Group unit root tests

Group unit root test

Series: Private Consumption (differenced logs), University of Michigan Index of Consumer Sentiment (levels), News Sentiment Index (levels), Real Personal Income (differenced logs), Inflation (differenced logs of Consumer Price Index), Unemployment Rate, Short-term interest rates (differences of 3-month USD LIBOR), Case Shiller Home Price Index - real values (differenced logs)

Sample: 1995Q1 2009Q4

Method	Statistic	Probability**	Cross-sections	Obs
Null: Unit root (assumes common unit root process)				
Levin, Lin & Chu (2002) test	-1.6109	0.0536	8	456
Null: Unit root (assumes individual unit root process)				
Im, Pesaran and Shin (2003) W-stat	-6.47118	0.0000	8	456
ADF - Fisher Chi-square [according to Dickey & Fuller (1979)]	92.664	0.0000	8	456
PP - Fisher Chi-square [according to Phillips & Perron (1988)]	111.916	0.0000	8	467

** Probabilities for Fisher tests are computed using an asymptotic Chi-square distribution. All other tests assume asymptotic normality.

This table shows the various unit root tests for the base model. We can reject all null hypotheses of unit roots in the data.

2.2.2 ARMA-Regression Results

We examine private consumption growth with our formulated model. In all regressions, we test for serial correlation with the Breusch-Godfrey Serial Correlation Lagrange Multiplier tests, according to Godfrey (1978) and Breusch and Pagan (1979). According to these tests, we find no significant serial correlation in any of the regressions.

As outlined in equation (2.9), our model has an ARMA(1,2) structure. The results of the regressions run without S_t are presented in table 2.5. In regression (1), we include all variables available in the vector Z_t. Real personal income, interest rates and house prices are statistically significant, whereas inflation and the unemployment rate do not show statistical significance. The coefficients of personal income and house prices have positive signs, which means that with higher income and rising house prices, consumers tend to increase their consumption. This is in line with Case et al (2003) who show that rising house prices lead to higher private consumption, so that we expect a positive coefficient sign. With falling unemployment rates, the consumption should go up. In regression (2), we drop the unemployment rate from the vector Z_t. We still obtain robust and statistically significant coefficients for all variables except inflation. Hence, we assume that inflation is not as relevant for consumers as their personal income, interest rates, and house prices. Interest rates have a negative coefficient sign and are also statistically significant. This alludes to the borrowing behavior of US consumers who finance part of their consumption on loans. If interest rates fall, then they can consume more because the interest payments decrease, thus freeing up more capital of the household. In regression (3), we exclude inflation. In the first three regressions, the adjusted

Table 2.5: Regression Coefficient Estimates without Sentiment Variables

	Regression Coefficient Estimates of ARMA(1,2) models (standard errors in parentheses beneath coefficients)			
	Dependent Variable: Private Consumption (log differenced)			
	(1)	(2)	(3)	(4)
Independent Variables				
Real Personal Income (log differenced)	0.229385***	0.225284***	0.186883**	0.133287*
	(0.073439)	(0.071067)	(0.073625)	(0.073140)
Inflation - Consumer Price Index (log differenced)	-0.139651	-0.135471		
	(0.096907)	(0.097248)		
Unemployment Rate	-0.000370			
	(0.000928)			
Short-Term Interest Rates (3-month USD LIBOR) (differenced)	-0.003043**	-0.002955**	-0.002645**	
	(0.001300)	(0.001273)	(0.001152)	
Case Shiller Home Price Index - real prices (log differenced)	0.054678**	0.049821*	0.051015*	
	(0.025583)	(0.025080)	(0.025523)	
Constant	0.006469	0.004486***	0.003992**	0.005694***
	(0.005087)	(0.001468)	(0.001578)	(0.002076)
AR(1)	0.711584**	0.777332***	0.828695***	0.797431***
	(0.274794)	(0.132543)	(0.128206)	(0.142527)
MA(1)	-0.501508*	-0.552902***	-0.577252***	-0.587716***
	(0.284662)	(0.186784)	(0.202148)	(0.157087)
MA(2)	0.161265	0.172979	0.08831	0.245821*
	(0.163471)	(0.163176)	(0.172712)	(0.145313)
R-squared	0.542002	0.540841	0.523634	0.45684
Adjusted R-squared	0.467227	0.476559	0.46759	0.415847
N (after adjustments)	58	58	58	58
Schwarz criterion	-7.788195	-7.855671	-7.888888	-7.897687
Root Mean Squared Error (RMSE)	0.003596	0.0036	0.003667	0.003916
Theil Inequality Coefficient	0.20703	0.207574	0.211219	0.226155

*,**,*** denote statistical significance at the 10%-, 5%-, and 1%-level, respectively

Note: All models calculated with heteroskedasticity consistent coefficient covariance and standard errors according to White (1980).

R-squared value was around 0.46. This suggests that the three statistically significant variables personal income, interest rates and house prices add the most explanatory power to the private consumption regression, whereas the unemployment rate and inflation do not seem to matter too much in this framework. With regression (4), we show that real personal income has the highest explanatory power among the macroeconomic variables examined that can explain private consumption. The adjusted R-squared value is quite high at 0.41. In a next step, we add the sentiment variables to equation (2.9) with the vector S_t as outlined in equation (3.2).

Table 2.6 shows the results.[12] In regression (5), almost all independent variables are statistically significant, except inflation and house prices, while both sentiment variables are highly statistically significant. The news sentiment index, the ICS, and personal income have positive coefficient signs. This means that higher sentiment and higher income explain rising consumption growth. The coefficient of the unemployment rate has the wrong sign, as we would expect that the sign is negative because higher unemployment should lead to less consumption. The coefficient of the Case Shiller Index is not statistically significant but it has the correct coefficient sign. The adjusted R-squared value is much higher compared to regression (1) that does not include sentiment: 0.68. Regressions (6) and (7) include the sentiment variables individually. In these regressions, personal income and interest rates are highly statistically significant. The Case Shiller Index has the correct coefficient sign and is statistically significant in the regressions with news sentiment (6) and

[12] Note that we have also tested the optimal ARMA-structure according to the Schwarz Information Criterion (BIC) Test. As in the previous regressions without the sentiment variables S_t, we confirm the optimal structure to be of the order ARMA(1,2). See Appendix 2.4.2 for the test results.

Table 2.6: Regression Coefficient Estimates with Sentiment Variables

Regression Coefficient Estimates of ARMA(1,2) models
(standard errors in parentheses beneath coefficients)

	Dependent Variable: Private Consumption (log differenced)								
	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Independent Variables									
News Sentiment Index (level)	0.040368***	0.042231**		0.041522***	0.038750**		0.027444**	0.032305*	
	(0.014543)	(0.018007)		(0.014713)	(0.018535)		(0.011338)	(0.018723)	
University of Michigan Index of Consumer Sentiment (level)	0.000313***		0.000314***	0.000287***		0.000294***	0.000313***		0.000311***
	(0.0000324)		(0.0000523)	(0.0000244)		(0.000045)	(0.0000232)		(0.0000457)
Real Personal Income (log differenced)	0.210054***	0.223363***	0.193235***	0.111430**	0.156963**	0.111906*			
	(0.065816)	(0.0070126)	(0.059252)	(0.054556)	(0.059256)	(0.057340)			
Inflation - Consumer Price Index (log differenced)	-0.117491	-0.150478	-0.117660						
	(0.070611)	(0.097046)	(0.079460)						
Unemployment Rate	0.000767***	-0.000296	0.000865*						
	(0.000213)	(0.000710)	(0.000512)						
Short-Term Interest Rates (3-month USD LIBOR) (difference	-0.002305***	-0.003356***	-0.002402**	-0.002233***	-0.002137**	-0.002011***			
	(0.000814)	(0.001197)	(0.000922)	(0.000604)	(0.000807)	(0.000654)			
Case Shiller Home Price Index - real prices (log differenced)	0.002407	0.052453**	0.016547						
	(0.016175)	(0.026028)	(0.019370)						
Constant	-0.051715***	-0.019781*	-0.027599***	-0.045628***	-0.018793*	-0.020703***	-0.037941***	-0.012673	-0.020812***
	(0.010652)	(0.011669)	(0.006717)	(0.009593)	(0.011093)	(0.003988)	(0.007932)	(0.010952)	(0.004221)
AR(1)	0.688765***	0.815143***	0.425284	0.770694***	0.865619***	0.409504	0.693802***	0.837459***	0.097748
	(0.232569)	(0.139189)	(1.069318)	(0.126010)	(0.113669)	(1.040858)	(0.167858)	(0.110104)	(2.070965)
MA(1)	-0.946796***	-0.590288***	-0.561613	-0.898558***	-0.567582***	-0.437218	-0.937085***	-0.59034***	-0.218277
	(0.316116)	(0.206101)	(1.039560)	(0.171594)	(0.165379)	(1.008484)	(0.207201)	(0.141943)	(2.022967)
MA(2)	-0.051557	0.103609	0.172615	-0.074562	0.13972	0.132102	-0.034715	0.206638	0.078435
	(0.207833)	(0.175446)	(0.176179)	(0.163957)	(0.163656)	(0.162756)	(0.201897)	(0.142503)	(0.280370)
R-squared	0.740973	0.584069	0.678825	0.690112	0.539364	0.632889	0.648	0.456235	0.586521
Adjusted R-squared	0.685861	0.506082	0.618605	0.646728	0.485171	0.5897	0.614153	0.415197	0.555315
N (after adjustments)	58	58	58	58	58	58	58	58	58
Schwarz criterion	-8.218114	-7.814533	-8.073066	-8.248858	-7.922467	-8.149412	-8.261451	-7.896574	-8.170484
Root Mean Squared Error (RMSE)	0.002704	0.003426	0.003011	0.002958	0.003606	0.003219	0.003152	0.003918	0.003416
Theil Inequality Coefficient	0.152427	0.196299	0.170269	0.167118	0.206506	0.18259	0.178819	0.226286	0.194809

*,**,*** denote statistical significance at the 10%-, 5%-, and 1%-level, respectively

Note: All models calculated with heteroskedasticity consistent coefficient covariance and standard errors according to White (1980).

without sentiment (1). However, house prices are not statistically significant when consumer sentiment is included (7). Neither the inflation rate is statistically significant in regressions (6) and (7) nor is the unemployment rate. This finding is in line with our previous regressions without sentiment. Hence, we get a first impression of how the sentiment variables behave in a consumption behavior framework. The ICS appears more dominant in the models whereas the news sentiment variable seems to be less dominant. This is manifested in the observation that there are less significant variables in regression (7), which includes the ICS, than in regression (6), which includes news sentiment. The adjusted R-squared value of regression (5) with both sentiments is the highest (0.68), whereas that of regression (7) with the ICS is higher (0.61) than that of regression (6), which includes news sentiment (0.50). Regression (1) without a sentiment variable has an adjusted R-squared value of 0.46. Therefore, we can claim that both news and consumer sentiment add value to the model when explaining private consumption growth, although with different explanatory power.

In regressions (8) to (10), we follow the same pattern as in the regressions before, but this time we exclude the unemployment rate, inflation and the Case Shiller Home Price Index because of statistical insignificance and wrong expected coefficient signs in some of the previous regressions. Thus, regression (8) includes both sentiment variables, which are highly statistically significant. Comparing regressions (9) and (10), the difference between the ICS and news sentiment becomes apparent, as the regression that includes the ICS shows a much higher adjusted R-squared value (0.59) and thus higher statistical significance. Therefore, we note that both real personal income and interest rates can explain private consumption with both sentiment variables, individually and jointly. Last but not least, we consider the sentiment variables jointly and individually without any other variables from the vector Z_t in regressions (11) to (13). We confirm what we have anticipated thus far, namely that the ICS is better suited to explain private consumption than news sentiment. However, news sentiment adds explanatory power to the model, when we compare the adjusted R-squares of the regressions. Regression (11) has an adjusted R-squared value of 0.61, whereas the regressions (12) and (13) that include the sentiment variables individually have values of 0.41 and 0.55, respectively. When comparing these values with the general regressions without the sentiment variables (cf. table 2.5), we find that news sentiment has similar explanatory power than personal income (both have adjusted R-squared values of 0.41), while the ICS has more explanatory power than all five variables of vector Z_t combined.

2.2.3 Standardized Coefficient Analysis

We further test the relative economic importance of the variables examined by a standardization of the coefficients, according to Bring (1994). Therefore, we standardize the dependent variable, namely private consumption denoted as y_t, and the independent variables, such as the variables in the sentiment vector S_t and the other macroeconomic variables from the vector Z_t that are denoted as x_t, as follows:

$$x_t^*,\, y_t^* = \frac{x_t,\, y_t - \bar{x},\, \bar{y}}{\sigma_{x,y}}, \tag{2.11}$$

where x_t^* and y_t^* are the new standardized variables, x_t and y_t the initial variables, $\bar{x}$ and $\bar{y}$ the means of the initial variables, and σ_x and σ_y the standard deviations of the variables. The advantage of such a standardization is that one can compare the coefficients of the regression with respect to size. Given that the variables are measured in different units, especially the sentiment variables, we compare the coefficients by considering their standard deviations. Thus, a one standard deviation change in the independent variable shows the relative change in the standard deviation of the dependent variable, and thus its relative impact on the dependent variables compared to other independent variables.

We run the same regressions as outlined in equations (3.2) and (2.9) with the new standardized values of the variables in order to get a clearer understanding about the relative importance of each independent variable on private consumption. Table 2.7 shows the results.

In regression (14), we estimate the model with both sentiment variables and all macroeconomic variables of the vector Z_t. The picture with regards to the relative importance of the coefficients as in the previous regressions is confirmed. The ICS is by far the most dominant variable in the model, followed by personal income, interest rates, and news sentiment. Interestingly, house prices do not tend to matter as much as in the previous regressions. The high adjusted R-squared of 0.73 shows that both sentiment scores add significant explanatory power to the model. In regressions (15) to (17), we exclude the sentiment variables individually and jointly. Regression (15) includes only the news sentiment index. With a statistically significant coefficient of 0.24 news sentiment has less relative importance than personal income, interest rates and house prices, but more than inflation and the unemployment rate,

Table 2.7: Standardized Coefficient Estimates

	Standardized Coefficient Estimates of ARMA(1,2) models (standard errors in parentheses beneath coefficients)			
	Dependent Variable: Private Consumption (standardized differenced logs)			
	(14)	(15)	(16)	(17)
Independent Variables				
News Sentiment Index (standardized levels)	0.225016***	0.247799**		
	(0.070413)	(0.105657)		
University of Michigan Index of Consumer Sentiment (standardized levels)	0.758032***		0.770667***	
	(0.120720)		(0.128521)	
Real Personal Income (standardized differenced logs)	0.339656***	0.355880***	0.307844***	0.365445***
	(0.110062)	(0.111708)	(0.094401)	(0.117002)
Inflation - Consumer Price Index (standardized differenced logs)	-0.167948**	-0.159177	-0.124509	-0.147803
	(0.075161)	(0.102684)	(0.084085)	(0.102556)
Unemployment Rate (standardized)	0.152472	-0.070777	0.207177*	-0.088571
	(0.095907)	(0.170205)	(0.122585)	(0.222339)
Short-Term Interest Rates (3-month USD LIBOR) (standardized differences)	-0.358343***	-0.402152***	-0.287897**	-0.364767**
	(0.103172)	(0.143400)	(0.110523)	(0.155810)
Case Shiller Home Price Index - real prices (standardized differenced logs)	0.051629	0.295271**	0.093171	0.307964**
	(0.135121)	(0.146623)	(0.109119)	(0.144084)
Constant	-0.031201	-0.081932	-0.001839	-0.022241
	(0.090996)	(0.257478)	(0.081144)	(0.218457)
AR(1)	0.849429***	0.815008***	0.425573	0.711541**
	(0.132609)	(0.139124)	(1.068861)	(0.274948)
MA(1)	-1.145633***	-0.590292***	-0.561896	-0.501465*
	(0.250135)	(0.206031)	(1.039108)	(0.284764)
MA(2)	-0.049106	0.103711	0.172661	0.161269
	(0.190161)	(0.175413)	(0.176044)	(0.163440)
R-squared	0.783969	0.584069	0.678825	0.542002
Adjusted R-squared	0.738005	0.506082	0.618606	0.467227
N (after adjustments)	58	58	58	58

*,**,*** denote statistical significance at the 10%-, 5%-, and 1%-level, respectively

Note:

1) All models calculated with heteroskedasticity consistent coefficient covariance and standard errors according to White (1980).

2) The standardized coefficients were calculated by subtracting the mean from each variable and then dividing it by its standard deviation. See text for further explanation and

while it adds explanatory power to the overall model, manifested in a higher adjusted R-squared value. Regression (16) clearly shows that the ICS is the most dominant variable over all other variables, as it has by far the highest standardized coefficient of 0.77. The coefficients of personal income, interest rates and house prices are lower than in the previous regression with news sentiment, confirming the dominance of the ICS over news sentiment in the model. The adjusted R-squared value is also significantly higher (0.61) than in regression (15) with news sentiment, for which it is 0.50. Regression (17) does not include the sentiment variables, but only the variables from vector Z_t. In that regression, taking the size of the coefficient into account, we see that personal income and inflation are equally important for private consumers, while changes in house prices also play a greater role for consumer behavior.

2.2.4 Summary of Results

We thus summarize three key findings. First, both consumer and news sentiment are statistically significant and they add explanatory power to private consumer behavior models. Second, it appears that consumer sentiment, measured by the ICS, performs markedly better than the news sentiment index in consumption behavior models. Hence, we find that consumer sentiment is the more powerful sentiment variable. This finding is manifested in the ARMA-regression models analysis as well as in the standardized coefficients analysis. And, third, news sentiment performs as good as personal income in explaining private consumption, whereas consumer sentiment is more powerful in explaining private consumer behavior than personal income, inflation, the unemployment rate, interest rates and house prices combined. In general, we conclude that news sentiment is a useful variable to add to consumer behavior models, especially when coupled with consumer sentiment and other macroeconomic variables.

2.3 Conclusion

The Index of Consumer Sentiment from the University of Michigan has been examined widely in the literature, and, in a broader context, it has been shown that consumer sentiment can partly explain the consumption behavior of private households. Nevertheless, it has been argued that other variables are also important in the consumption equation, such as personal income, interest rates, or house prices, for example. However, a study that considers a sentiment index generated from newspapers articles and its statistical ability to explain private consumption in the US has not yet been undertaken. In order to contribute to the literature in this respect, we introduce a novel data set and procedure by creating a news sentiment index that accounts for positive and negative sentiment from over 100,000 newspaper articles of the economics section of two of the most widely-read newspapers in the US from 1995 to 2009. This novel dataset is constructed by utilizing a new text mining approach. We briefly elaborate on the process of constructing such a news sentiment index from publicly available news sources.

Further, we examine empirically the connection and impact of this news sentiment index on consumer behavior. In accordance with previous findings of Carroll et al (1994) and Sommer (2007), among others, we formulate consumption behavior models that incorporate consumer and news sentiment among other macroeconomic variables, such as personal income, inflation, house prices, the unemployment rate, and interest rates. Following Doms and Morin (2004), we add news sentiment to their information flow chart (c.f. fig. 2.1) and examine that relationship empirically. We construct ARMA(1,2)-models, which show that both consumer and news sentiment add explanatory power to consumption behavior models. We find that an increase in both sentiment variables explain an increase in private consumption. Further, we show with a standardized coefficient analysis that consumer sentiment (i.e. the University of Michigan Index of Consumer Sentiment) has more explanatory power than news sentiment in our models. Last, but not least, we find that news sentiment performs as good as personal income in explaining private consumption, whereas consumer sentiment is the more powerful variable than personal income, inflation, the unemployment rate, interest rates and house prices combined when explaining private consumer behavior. News sentiment is a valid variable to add in consumption behavior models, as it captures a different set of information than consumer sentiment and other macroeconomics variables, such as personal income or house prices.

This first analysis of sentiment in the media and its explanatory power for private consumption leaves room for future research in order to specify and improve the methods for explaining private consumption that are based on news and consumer sentiment as well as other variables that affect the ordinary household in their consumption behavior. One possible source to examine sentiment in the media is sentiment in TV news shows that relate to economic matters. This research question is considered precisely in the next chapter.

2.4 Appendix

2.4.1 Visual Basic Programs

The Visual Basic Programs were written in order to handle the vast amount of articles and process them into a suitable format for the java program that features the sentiment algorithm. When downloading the articles from the *LexisNexis* database, the articles of one day are summarized in one text file. The first program was written to cut the articles into separate text files in order to format them for the java program that runs the sentiment analysis. The output log-file from the java program was then formatted and coded into $\{-1\}$ for negative sentiment and $\{1\}$ for positive sentiment. Neutral values are not coded by the algorithm in order to avoid ambiguity. The daily sentiment data were then aggregated to quarterly values.

2.4.2 Schwarz Information Criterion (BIC) Tests

ARMA-Lag Structure	Schwarz Info Criterion (BIC)	Inverted MA Roots*
ARMA(1,1)	-8.158792	<1
ARMA(2,1)	-8.203223	1
ARMA(1,2)	-8.231254	<1
ARMA(2,2)	-8.066476	<1
ARMA(3,1)	-8.123197	1
ARMA(3,2)	-8.251037	>1
ARMA(3,3)	-8.035030	<1
ARMA(2,3)	-8.008461	<1
ARMA(1,3)	-8.181111	<1

*Note: Inverted Roots of MA process have to be smaller than 1 so that the process is stationary and invertible.

This table shows various Schwarz Information Criteria test results in order to determine the best ARMA-structure of the extended model with private consumption as dependent variable and news sentiment, the ICS, real personal income, inflation, the unemployment rate, short-term interest rates, and the Case Shiller Home Price Index (real prices) as independent variables. The most suitable ARMA-structure has the lowest BIC and is found to be of the order ARMA(1,2), as indicated.

3 Nowcasting Private Consumption with TV Sentiment[13]

3.1 Introduction

3.1.1 Contextual Setting and Summary

As laid out in the previous chapter, sentiment in newspapers is a valid explanatory variable in private consumption behavior models. Although the newspaper is a widely accepted medium, the television industry has become the most influential in the last decade. The United States of America are one of the more prominent cases how television has influenced society. According to some industry polls done in the past years, most Americans obtain their information about news through television.[14] According to a recent survey of Nielsen (2010), the average American watches over five hours of television per day. Given these developments, we want to address the question whether watching TV news influences the watchers. Does the way of reporting and the content of TV news shows have an impact on private households and their behavior? In a study on the impact of how news stories are portrayed, Maier (2005) finds that the hard facts are not necessarily most important to the reader, but rather how these facts are presented. This component of how the news are presented on TV is what we call TV sentiment in this study. Given these findings, we examine sentiment in TV news broadcasts to identify a possible link between TV sentiment and the consumption behavior of households. We do this with a dataset that contains sentiment from four of

[13] This chapter is based on Uhl (2012).
[14] See Pew (2004) and Harris Interactive (2007).

the most widely watched TV news broadcasts in the US. We first perform principal components analyses of both the widely-used University of Michigan Index of Consumer Sentiment (ICS) and the new variables of TV sentiment. In an empirical exercise, we show the usefulness of financial variables and the principal components of the sentiment variables of the ICS and TV sentiment in a nowcasting environment. We find that TV sentiment is a valid variable to add in nowcasting private consumption, as it shows higher nowcasting power than the ICS, especially when coupled with financial variables, such as stock returns and interest rates.

This chapter is structured as follows: this section gives an overview of the literature and the data, section 3.2 lays out the model, section 3.3 discusses the empirical results, while section 3.4 concludes.

3.1.2 Related Literature

Today's public obtains most information about world and economic affairs from the media, as we are constantly confronted with a news flow. But does this influence us – as consumers – in a way that we adjust our consumption behavior according to the news we read, hear or watch, in particular with regards to the economic development and state of the economy? Is it possible that consumers adjust their consumption behavior, when they hear news about the rising unemployment rate and that many people are being fired? Possibly, they might be more cautious given these news because they fear being laid off as well, thus spending less and saving more for potential bad times. Is it possible that the tone of the economic news reporting matters? For example, if journalists constantly report bad news about the economy, does this influence the consumer? Does such a phenomenon exist and if so, how can we measure it?

Recent studies have confirmed this phenomenon by showing that news have an impact on consumers, and that there is an influence on the consumer through sentiment. For example, Carroll (2003) finds that news coverage as well as volume of economic topics in news are relevant to the consumer. Given his findings, we can infer that people do pay attention to news. So, if news matter, does the style of reporting matter as well? In a theoretical study, Sims (2003) shows that the tone and volume of news matter to the ordinary people when they form their opinion of the state of the overall economy. This influence goes beyond the pure information content of a particular state of the economy, i.e. the hard facts do not always matter. Thus, according to Sims (2003), it is also important how news are being portrayed, i.e. sentiment in the news influences the consumers as well. Doms and Morin (2004) consider this issue empirically by examining whether the news media influences perceptions of consumers. They come to the conclusion that there is an effect of sentiment on household spending behavior because the tone and reporting in news affect consumers. Therefore, we want to consider the tone and reporting, i.e. the sentiment, of news more closely in order to make out a possible impact of sentiment on consumers. However, we do not only want to consider sentiment in relation to consumer behavior, but also other factors that might drive consumer behavior, such as stock prices and interest rates in order to account for wealth changes, as laid out in Ludwig and Slok (2002).

Other studies dealing with the impact of news on the public examine whether there is a bias in the media, which might ultimately reflect on and influence the public. Mullainathan and Shleifer (2005) identify theoretically that there are biases in economic and political news and that these are slanted towards the customers of the media outlet. Therefore, when examining sentiment in the media, we need to consider various channels or news shows on television in order to capture the most general picture possible. In a later study, Baron (2006) confirms that the news media plays an essential role in society and identifies issues in media bias. Recently, Gentzkow and Shapiro (2010) construct an index of media slant. They find that those newspapers with specific political views are more likely to be read by readers with similar views. The concept, or, the existence of media bias, is important to this study because it shows that there is a subjective component inherent in the news that goes beyond sheer hard facts.

In the studies discussed above, the media samples have mainly been taken from the print media. Only a few studies have dealt so far with audiovisual media outlets and their impact on the consumer. For example, Strömberg (2004) identifies large and highly significant effects of radio on voting behavior, while TV news broadcasts have only been looked at more recently. DellaVigna and Kaplan (2007) consider Fox News in cable markets and its impact on voting behavior in the US, taking media bias into account. Their results suggest that Fox News have a significant impact on viewers to vote Republican. Their findings suggest that a subjective component might be involved in TV reporting that influences the voters. On another note, the influence of TV has been examined on stock investors by Meschke and Kim (2011). In their study, they investigate CEO interviews, while documenting significant positive abnormal returns accompanied by abnormally high trading volume. They find evidence that enthusiastic individual investors are prone to trading more based on CNBC interviews, confirming that there might be a sentiment factor that influences people to act in a particular way. Therefore, in this analysis, we want to examine whether sentiment in TV news broadcasts have an impact on consumers, while at the same time controlling for wealth effects, such as stock prices and interest rates that are readily available. According to a study of Ang *et al* (2007), consumer sentiment surveys perform best in forecasting models. Given this finding, we want to use the ICS as a proxy to measure consumer sentiment.[15] Curtin (2007) shows the top sources of information on the economy among households.[16] In this survey, the most common source for information gathering in the US is television. Hence, we have a possible link between consumer and TV sentiment. Both variables might be useful to explain private consumption. In recent exercises, Kholodilin *et al* (2010) and Schmidt and Vosen (2011) both use Google Trends results to nowcast US private consumption in a real-time framework. We adapt an approach similar to Kholodilin *et al* (2010) who perform a principal component analysis of their various indicators. In a nowcasting environment, they show that Google Trends results are useful in nowcasting private consumption. In our analysis, we perform nowcasts with the variables at hand in order to test the nowcasting power of TV sentiment.

[15] See Curtin (1982) and Curtin (2007) for a discussion of the University of Michigan Indices.

[16] See table 4 in Curtin (2007): Sources of Information on Official Rates of Unemployment, Consumer Prices, and Gross Domestic Product.

3.1.3 Data

The monthly TV sentiment dataset is from MediaTenor, a professional news sentiment provider. The sentiment data were compiled exclusively from US TV news shows on the US economy. Contrary to other approaches and studies, the sentiment was coded by humans, not by a machine or pre-defined automatic algorithm.[17] Tagged topics range broadly and contain possible links to the development and the state of the economy.[18] Important to note is that employees from MediaTenor are highly trained to adhere to a very specific pre-defined sentiment rating and coding table, in order to avoid subjectivity bias. The dataset is constructed from sentiment from four of the most widely watched news broadcasts in the US: ABC World News Tonight, CBS Evening News, FOX News, and NBC Nightly News. In total, statements in over 10,000 TV news shows were coded for sentiment from January 2005 to December 2009. The summary statistics of the individual sentiments from the four TV news shows are shown in table 3.1.

In order to better understand what drives the widely used ICS, we disentangle the index into its five components in order to get further clues what might drive consumer sentiment. The ICS is constructed from answers to five questions relating to current economic conditions of consumers as well as consumer expectations.[19] The five components of the ICS are made up of

[17] See *Human Analysis vs. Software* for an evaluation of MediaTenor's approach vs. machine-based approaches, available at http://www.mediatenor.com/mca_brain_vs_software.php, last accessed 1 March 2011.

[18] For a more detailed description of the methodology that MediaTenor uses, go to http://www.mediatenor.com/mca_methodology.php, last accessed 1 March 2011.

[19] A detailed description of the calculation of the index and the individual questions can be found on the homepage of the surveys of consumer from the University of Michigan and Thomson Reuters. See *Index Calculations,*

Table 3.1: Summary Statistics of TV Sentiment Sources

	ABC World News Tonight	CBS Evening News	FOX News	NBC Nightly News
Mean	-0.4812	-0.4950	-0.4320	-0.5296
Median	-0.6610	-0.6220	-0.5326	-0.6000
Maximum	1.0000	0.5000	0.5714	0.7333
Minimum	-1.0000	-1.0000	-1.0000	-1.0000
Std. Dev.	0.4585	0.3716	0.3862	0.3702
Skewness	1.3849	0.9813	0.7632	1.2689
Kurtosis	4.7122	3.2176	2.7707	4.4263
Jarque-Bera	26.5090	9.7474	5.9559	21.1867
Probability	0.0000	0.0076	0.0509	0.0000
Sum	-28.871	-29.701	-25.922	-31.778
Sum Sq. Dev.	12.40349	8.146239	8.797797	8.084831
Observations (2005M01 - 2009M12)	60	60	60	60

Source: MediaTenor

questions that consumers are asked with regards to their current buying conditions, their business conditions in 12 months and 5 years, the current conditions of their personal finances and their expected conditions of personal finances. Table 3.2 shows the summary statistics of the individual components. Monthly private consumption data were obtained from the ALFRED database.[20] As in Kholodilin *et al* (2010) who argued for real-time vintages of private consumption, we also use monthly unrevised real-time data. This applies for the ICS data as well, as the publication lag might be a crucial factor, and in order to obtain meaningful results that we can compare to the other data, we use unrevised "real-time" data as opposed to the final revised data. The ICS data were downloaded from Thomson Reuters Datastream. Breeden (1986) shows, for example, that interest rates have a potential impact on private consumption growth, so that we include short-term as well as long-term interest rates in our analysis. In order to account for changes in wealth effects, we include stock prices of the S&P 500 stock index.[21] The financial variables data were obtained from Thomson Reuters Datastream.

3.2 Modeling

3.2.1 Principal Components Analysis

In accordance with Kholodilin *et al* (2010), which base the construction of their Google indicator on the factor model of Stock and Watson (1999, 2002),

http://www.sca.isr.umich.edu/documents.php?c=i, last accessed 20 February 2011.

[20] See ArchivaL Federal Reserve Economic Data. Available at http://alfred.stlouisfed.org/, last accessed 15 September 2010.

[21] See, for example, Ludwig and Slok (2002) for a detailed discussion of the effect of changes in wealth on private consumption.

Table 3.2: Summary Statistics of the University of Michigan Index of Consumer Sentiment Components

	Buying Conditions	Business Conditions in 1 year	Business Conditions in 5 years	Personal Finance - Current Conditions	Personal Finance - Expected Conditions
Mean	139.1000	79.0167	83.6000	96.4000	117.2833
Median	149.0000	81.5000	83.5000	103.0000	118.0000
Maximum	172.0000	118.0000	107.0000	123.0000	133.0000
Minimum	88.0000	31.0000	59.0000	58.0000	96.0000
Std. Dev.	24.5893	24.2337	10.9949	22.0609	8.6516
Skewness	-0.4590	-0.2996	-0.1839	-0.4588	-0.3628
Kurtosis	1.7734	1.9862	2.3901	1.6428	2.6683
Jarque-Bera	5.8687	3.4673	1.2681	6.7105	1.5915
Probability	0.0532	0.1766	0.5304	0.0349	0.4512
Sum	8'346.000	4'741.000	5'016.000	5'784.000	7'037.000
Sum Sq. Dev.	35673.4	34648.98	7132.4	28714.4	4416.183
Observations (2005M01 - 2009M12)	60	60	60	60	60

we apply a principal components analysis[22] to the two sentiment indicators that we are going to implement: TV sentiment and the ICS. Contrary to Kholodilin *et al* (2010), we have less principal components, so it makes sense to consider all principal components in the nowcast later on. Table 3.3 shows the eigenvalues and eigenvectors of the principal components of TV sentiment. The first principal component covers 66% of the total variance, whereas the first three components cover 92%. In the Eigenvectors section, we can see that CBS and NBC have values over 0.50, so that these are possibly contributing the most to principal component 1. ABC is in the second principal component, whereas FOX News in the third. Given its small share in the eigenvalues analysis, principal component 4 is of minor importance. It is interesting to note that the various sentiment from the TV news shows contribute differently to the various components, rather than being "grouped" in one principal component. Hence, there must be differences in the style of reporting among the different news shows, which contribute differently to the total variance of the components. Fig. 3.1 shows the graph of the four principal components of TV sentiment. We note that the first principal component graph traces private consumption quite well, with the trough being at the end of 2008/beginning of 2009, the recent financial crisis. The other principal components do not show a distinct pattern.

Table 3.4 shows the principal components analysis of the ICS. The first three principal components make up over 97% of the total variance. In the first principal component, business conditions in one year as well as personal finance expectations appear the most relevant. Therefore, we can say that the

[22] See also Gabriel (1971) for a more detailed discussion of how the principal components analysis is done here.

Table 3.3: Principal Components Analysis of TV Sentiment

Eigenvalues: (Sum = 4, Average = 1)

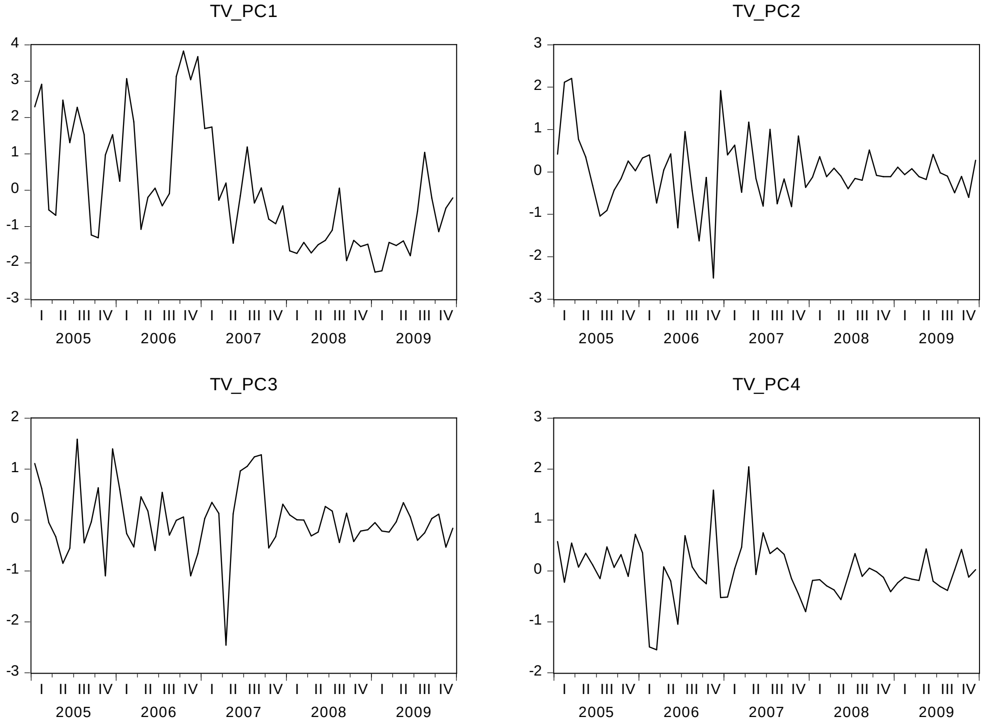
Number	Value	Difference	Proportion	Cumulative Value	Cumulative Proportion
1	2.6414	2.0330	0.6603	2.6414	0.6603
2	0.6083	0.1782	0.1521	3.2497	0.8124
3	0.4301	0.1100	0.1075	3.6799	0.92
4	0.3201	---	0.0800	4.0000	1

Eigenvectors (loadings):

Variables	PC1	PC2	PC3	PC4
ABC	0.4419	0.8824	0.1414	-0.0779
CBS	0.5290	-0.3032	-0.1855	-0.7706
FOX	0.5028	-0.3425	0.7333	0.3034
NBC	0.5216	-0.1099	-0.6386	0.5550

This table shows the principal components analysis of TV Sentiment. PC1 – PC4 stand for principal components 1 – 4.

Figure 3.1: Principal Components of TV Sentiment



first component is about future conditions, or expectations of consumers in the near future, i.e. in one year. The second principal component is about the current conditions, as the components current business conditions and current assessment of personal finances are the most relevant here. The third component aims at longer-term expectations, i.e. the variable business conditions in five years. Given its low eigenvalue, the fourth component appears to be of minor importance. However, it appears that it contains information on both current and expected personal finances, while the fifth component has information on the business conditions. Hence, we note that expectations about future conditions in one year explain the greatest share of the variance, while current conditions are second. Fig. 3.2 shows the graphs of the principal components of the ICS. The graph of the first principal component shows the recent financial crisis quite nicely, whereas the other components do not immediately show a discernible picture. We test the predictive accuracy of these components with nowcasts in the next step.

3.2.2 Nowcasting and Evaluation

As in Kholodilin *et al* (2010), our sample size is rather of limited size because the TV sentiment data is only available for five full years. We thus apply a parsimonious ARMA(1,2)-model of the following form:

$$\triangle \log c_t = k + \alpha_1 \triangle \log c_{t-1} + \beta x_t + \sum_{i=1}^{2} \theta_i \varepsilon_{t-i} + \varepsilon_t, \tag{3.1}$$

where $\triangle \log c_t$ refers to logged private consumption growth, x_t is an exogenous variable that can be a financial variable or a principal component of TV or the ICS, whileε_t is the error term. The ARMA(1,2)-model is chosen because we have shown in the last chapter that this is the most suitable model for private consumption, which is also in accordance with the findings of Sommer (2007) and Carroll *et al* (2010). For the exogenous variable x_t, it can be either one of the three financial variables, one of the principal components of both TV and consumer sentiment, and a combination of these variables. We thus have 33 models that we run. Thus, the benchmark model without any of the indicators is an ARMA(1,2)-model:

$$\triangle \log c_t = k + \alpha_1 \triangle \log c_{t-1} + \sum_{i=1}^{2} \theta_i \varepsilon_{t-i} + \varepsilon_t. \tag{3.2}$$

Based on equations (3.1) and (3.2) we perform one-step ahead nowcasts with a similar methodology as in Kholodilin *et al* (2010). We use real-time vintages of private consumption, as these flash estimates are the most relevant

Table 3.4: Principal Components Analysis of the University of Michigan Index of Consumer Sentiment

Eigenvalues: (Sum = 5, Average = 1)

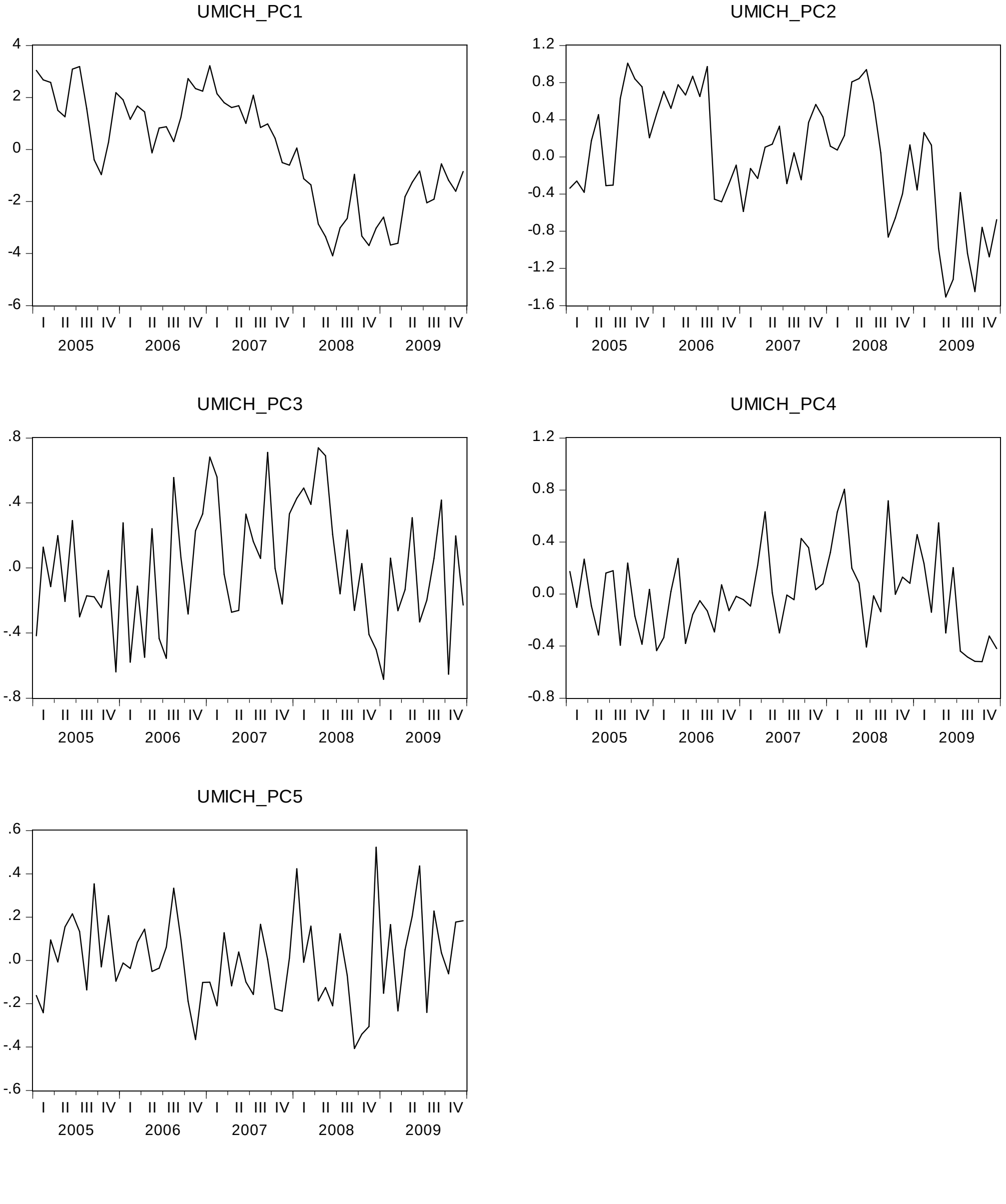
Number	Value	Difference	Proportion	Cumulative Value	Cumulative Proportion
1	4.3160	3.9118	0.8632	4.3160	0.8632
2	0.4042	0.2665	0.0808	4.7202	0.944
3	0.1377	0.0363	0.0275	4.8579	0.9716
4	0.1013	0.0606	0.0203	4.9592	0.9918
5	0.0408	---	0.0082	5.0000	1

Eigenvectors (loadings):

Variables	PC1	PC2	PC3	PC4	PC5
Buying Conditions	0.4517	0.4758	-0.0215	-0.2709	0.7041
Business Conditions in 1 year	0.4613	-0.2085	-0.0933	-0.7353	-0.4408
Business Conditions in 5 years	0.4365	-0.5164	0.6777	0.2281	0.1773
Personal Finance - Current Conditions	0.4380	0.5904	0.1701	0.3984	-0.5215
Personal Finance - Expected Conditions	0.4480	-0.3390	-0.7089	0.4186	0.0810

This table shows the principal components analysis of the ICS. PC1 – PC5 stand for principal components 1 – 5.

Figure 3.2: Principal Components of the ICS



to economists and professional forecasters as well as policy makers. Kholodilin *et al* (2010) further make the case that the data are revised up to 23 months after the first flash estimate. This would in turn, they argue, shrink the already small sample, which applies here, too. We estimate the models, from which we then perform one-step ahead nowcasts for the period 2008M01 until 2009M12. Once the nowcast of the first month (2008M01) is done, the model is re-estimated with actual values of 2008M01. Then, it proceeds with a nowcast for the next month, i.e. 2008M02. This procedure is repeated until we reach the end of the nowcast period in 2009M12.[23] The evaluations of the nowcasts are based on each nowcast month. The Root Mean Squared Error (RMSE) is used to evaluate the accuracy of the various nowcasts:

$$\sqrt{\sum_{t=T+1}^{T+h} \left(\hat{y}_t - y_t\right)^2 / h}, \tag{3.3}$$

where the actual and forecasted value in period t is y_t and $\hat{y}_t$, respectively. In order to compare the accuracy of the nowcasts, we apply the Theil Inequality Coefficient as in Theil (1958):

$$\frac{\sqrt{\sum_{t=T+1}^{T+h} \left(\hat{y}_t - y_t\right)^2 / h}}{\sqrt{\sum_{t=T+1}^{T+h} \hat{y}_t^2 / h} + \sqrt{\sum_{t=T+1}^{T+h} y_t^2 / h}} \tag{3.4}$$

The Theil Inequality Coefficient is always between zero and one, where zero indicates a perfect fit of the model.

[23] See Appendix 3.5.1 for a detailed explanation of how the nowcasts are done.

3.3 Empirical Results

Table 3.5 shows the results of the nowcasts with the RMSEs and Theil Inequality Coefficients. We want to compare all 33 models and their nowcast accuracy. The first column shows the RMSE of the various individual nowcasts with the financial variables as well as the principal components of TV sentiment and the ICS as well as a combination of them. Column 2 shows the RMSE of the benchmark ARMA(1,2)-model, which is 0.003586. Hence, we compare this number with the RMSEs of all other models in order to see which models perform better and which perform worse, i.e. which variables and principal components beat an ARMA(1,2)-model. Given that we have taken the study of Kholodilin *et al* (2010) as a reference, we want to show first how their benchmark model performs in our framework as opposed to the benchmark model of the form ARMA(1,2) that we employ in this study. Their benchmark model is a simple AR(1)-process as follows:

$$\triangle \log c_t = k + \alpha_1 \triangle \log c_{t-1} + \varepsilon_t. \tag{3.5}$$

The RMSE of their benchmark model is 0.00393, which is significantly higher than the RMSE of our benchmark model. This justifies our findings from chapter 2, suggesting that using an ARMA(1,2)-process in consumption behavior models is superior to an AR(1)-structure. Then, we take the financial variables S&P 500 stock returns as well as long- and short-term interest rates and apply these to the ARMA(1,2) model. For all three variables we achieve a lower RMSE than for our base model. The RMSE of the model with stock returns is significantly lower than that of the other two, thus suggesting that stock returns are superior to interest rates in nowcasting private

Table 3.5: Nowcasts with Root Mean Squared Errors (RMSE) of the models and Theil Inequality Coefficients

	One-step-ahead Nowcasts (2008M1 to 2009M12)		
	RMSE (1)	RMSE Benchmark-model ARMA(1,2) (2)	Theil Inequality Coefficient (3)
ARMA(1,2)	--	0.003586	0.757995
AR(1) - benchmark model of Kholodilin et al (2010)	0.00393	0.003586	0.847014
1) S&P 500 stock index (differenced logs)	0.003071	0.003586	0.531546
2) Long-term interest rates (differenced 10-year US Treasury rates)	0.0035	0.003586	0.747883
3) Short-term interest rates (differenced 3-month USD LIBOR)	0.003581	0.003586	0.747304
Variables 1) - 3) combined	0.003097	0.003586	0.521553
Principal Component 1 of UMICH ICS	0.003425	0.003586	0.644185
Principal Component 2 of UMICH ICS	0.003659	0.003586	0.76601
Principal Component 3 of UMICH ICS	0.003584	0.003586	0.761438
Principal Component 4 of UMICH ICS	0.00355	0.003586	0.74863
Principal Component 5 of UMICH ICS	0.003593	0.003586	0.754855
Principal Component 1 of TV Sentiment	0.003374	0.003586	0.654928
Principal Component 2 of TV Sentiment	0.003591	0.003586	0.757997
Principal Component 3 of TV Sentiment	0.003594	0.003586	0.762037.
Principal Component 4 of TV Sentiment	0.003537	0.003586	0.745298
Principal Components 1 & 2 of TV Sentiment	0.003308	0.003586	0.649755
Principal Components 1 & 3 of TV Sentiment	0.003379	0.003586	0.658408
Principal Components 1 & 4 of TV Sentiment	0.003371	0.003586	0.654318
Principal Components 2 & 3 of TV Sentiment	0.003602	0.003586	0.762141
Principal Components 2 & 4 of TV Sentiment	0.003547	0.003586	0.746664
Principal Components 3 & 4 of TV Sentiment	0.003544	0.003586	0.749454
Principal Components 1 & 2 of UMICH ICS	0.003412	0.003586	0.636161
Principal Components 1 & 3 of UMICH ICS	0.003402	0.003586	0.618899
Principal Components 1 & 4 of UMICH ICS	0.003025	0.003586	0.480916
Principal Components 1 & 5 of UMICH ICS	0.003236	0.003586	0.550416
Principal Components 2 & 3 of UMICH ICS	0.003664	0.003586	0.764112
Principal Components 2 & 4 of UMICH ICS	0.003558	0.003586	0.728842
Principal Components 2 & 5 of UMICH ICS	0.003661	0.003586	0.766773
Principal Components 3 & 4 of UMICH ICS	0.003548	0.003586	0.750667
Principal Components 3 & 5 of UMICH ICS	0.003591	0.003586	0.758139
Principal Components 4 & 5 of UMICH ICS	0.003557	0.003586	0.747333
Principal Components 1 of TV Sentiment and UMICH ICS	0.003307	0.003586	0.628981
Variables 1) - 3), Principal Components 1 & 2 of TV Sentiment	0.002821	0.003586	0.456257
Variables 1) - 3), Principal Components 1 & 4 of UMICH ICS	0.003064	0.003586	0.538144
Variables 1) - 3), Principal Components 1 of TV Sentiment and UMICH ICS	0.002697	0.003586	0.404195

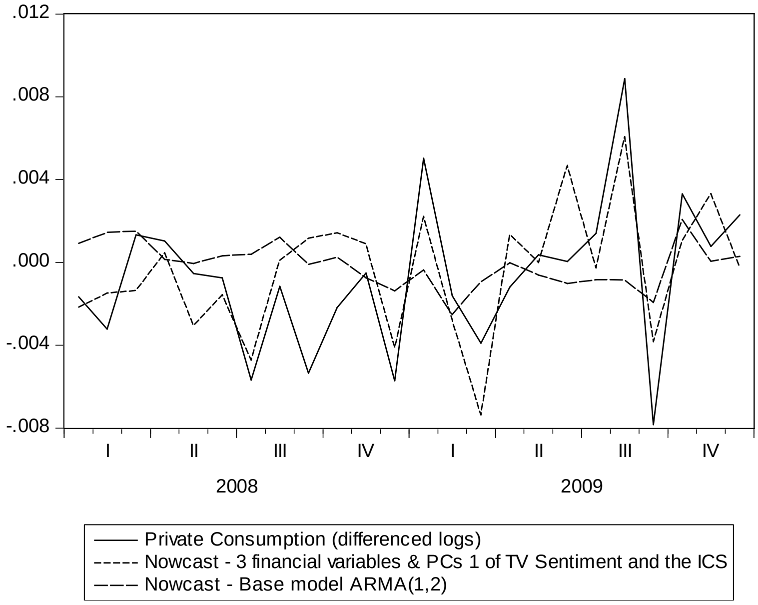
(1) This column shows the Root Mean Squared Error of the respective forecasts.
(2) This column shows the Root Mean Squared Error of the benchmark ARMA(1,2) model.
(3) This column shows the Theil Inequality Coefficient, indicating the goodness of fit of the respective forecast.

consumption. This becomes evident when considering the much lower Theil Inequality Coefficient. When we combine all three variables, we get an even lower Theil Inequality Coefficient. Hence, combining the three financial variables is efficient in this nowcasting exercise because the different variables contain a different set of information. The stock price index possibly account for changes in wealth effects among households, whereas the interest rates possibly capture lending and financing conditions.

We then turn to the individual principal components of both the ICS and TV sentiment. The nowcasts of both the first principal components of the ICS and TV sentiment achieve the lowest RMSEs and Theil Inequality Coefficients. We therefore assume that the most relevant information for nowcasting private consumption of both sentiment variables is captured in the first principal components. When comparing the RMSEs of the first principal components of TV sentiment and the ICS, we note that the RMSE of TV sentiment is lower than that of the ICS. However, the Theil Inequality Coefficient is slightly lower for the first principal component of the ICS. Given that both the RMSEs and the Theil Inequality Coefficients are pretty close together, but markedly lower than our base model, we can say that both sentiment variables add value in nowcasting private consumption.

Next, we combine various principal components of both TV sentiment and the ICS to pairs in order to test whether this adds value to the nowcasts. The lowest RMSE and Theil Inequality Coefficient achieve the first and second principal components of TV sentiment. This makes sense, as these two components make up over 80% of the total variance. A more interesting result is that the first and fourth principal components of the ICS achieve by far the lowest RMSE and Theil Inequality Coefficient score. As described earlier in

Figure 3.3: Graph of nowcasts of base and best model and private consumption



the principal components analysis section, the first component of the ICS aims at future conditions and expectations in one year, and the fourth at current and expected finances. Therefore, it is most important to the consumer how the near-term future expectations for the business conditions are, but it is also important how ones own personal finances are, both the current and the expected situation.

In a last step, we combine the three financial variables and the principal components of TV sentiment and the ICS, which performed the best in the previous nowcasts. Clearly, the nowcast that includes the financial variables and the first principal components of both the ICS and TV sentiment achieves the lowest RMSE and Theil Inequality Coefficient. When comparing the evaluation statistics of the nowcasts that include the financial variables and the best principal component pairs of the ICS and TV sentiment individually, we note that the nowcast with the first and second principal components of TV sentiment perform better than the nowcast with the first and fourth principal component of the ICS. Fig. 3.3 depicts private consumption, the nowcast from the base model, and the best nowcast with the financial variables and the first principal components of TV sentiment and the ICS. Even though the best nowcast tracks actual private consumption quite well, it becomes obvious that in especially times of crises, when the growth rates are more volatile, the nowcast is not entirely able to capture the large spikes, as in the third quarter of 2009. The base model is almost not able to track the volatility in the growth rates.

Therefore, we summarize the following findings: first, we show that an ARMA(1,2) structure as base model is superior to an AR(1) structure as suggested by Kholodilin *et al* (2010). Second, we find that stock returns of the broad-based stock index S&P 500 is suited for better nowcasts than interest rates. This is also in line with Ludwig and Slok (2002). Third, the first principal components of TV sentiment and the ICS are pretty much equally suited for nowcasting private consumption. Fourth, when combining the best pairs of the principal components of the ICS and TV sentiment, the first and fourth principal components of the ICS perform markedly better than the first and second principal components of TV sentiment. Last, but not least, when combining the financial variables with the best pairs of the principal components of the ICS and TV sentiment, TV sentiment adds more value to the nowcast than the ICS. Nevertheless, when combining the financial variables and the first principal components of both TV sentiment and the ICS, the best nowcasting result is achieved.

3.4 Conclusion

In this chapter, we show with principal components analyses of the ICS and TV sentiment the nowcasting ability of these and other financial variables. Given that television news have the greatest share of news sources among the American population, we extend the existing literature of nowcasting private consumption by introducing a new sentiment variable that was created from sentiment from four of the most widely watched TV news broadcasts in the US. The sentiment was gathered from over 10,000 TV news broadcasts from January 2005 to December 2009. The principal components analysis of the ICS shows that future conditions and expectations of business conditions and personal finances explain the greatest share of the variance in the ICS. For TV sentiment, CBS and NBC news shows explain the greatest variance, with ABC and FOX news having inferior explanatory power. We further confirm the findings of the previous chapter that an ARMA(1,2) structure is the superior base model than an AR(1)-model, as suggested in Kholodilin *et al* (2010). In the nowcast evaluation, we find that stock returns perform markedly better than interest rates. Further, we find that TV sentiment adds great value in nowcasting private consumption. This is rooted in the fact that the first principal component of TV sentiment achieves a lower RMSE in the nowcast than the first principal component of the ICS. When adding the financial variables to the best principal components pairs of TV sentiment and the ICS, TV sentiment also adds more value than the ICS. Nevertheless, when combining all three financial variables and the first principal components of both the ICS and TV sentiment, we achieve the best nowcast of private consumption. Hence, not only do financial variables and the widely used ICS add power in nowcasting private consumption, but also the newly introduced variable TV sentiment.

3.5 Appendix

3.5.1 Nowcasts

The nowcasts are one-step ahead forecasts with coefficients from the ARMA(1,2)-regressions. The nowcasts follow the static method, which means that after each step when a nowcast is done for $t+1$, the actual values of the variables are taken, when proceeding with the next step nowcast at $t+2$. In this chapter, the following ARMA(1,2) model is taken as follows:

$$y_t = x_t^{'}\beta + u_t,$$

$$u_t = \rho_1 u_{t-1} + \epsilon_t,$$

where y_t is private consumption and x_t refers to the independent variables, while

$$\epsilon_t = \phi_1\epsilon_{t-1} + \phi_2\epsilon_{t-2} + \varepsilon_t.$$

β is a vector of unknown parameters, and b are estimates of the unknown parameters β. The model is estimated with data up to $t = S - 1$. The fitted residuals are defined as $e_t = y_t - x_t'b$. Given that the values of x_t are available, the nowcasts for $t = S$ are as follows:

$$\hat{y}_S = x_S^{'}b + \hat{\rho}_1 e_{S-1},$$

where the residuals $\hat{u}_t = \hat{y}_t - x_t^{'}b$ are formed from the nowcast values of y_t. The MA-error terms are estimates of the pre-sample ϵ_{S-1}, $\epsilon_{S-1}, \ldots, \epsilon_{S-q}$ using the recursion:

$$\hat{\epsilon}_t = \hat{u}_t - \hat{\theta}_1 \hat{\epsilon}_{t-1} - \ldots - \hat{\theta}_q \hat{\epsilon_{t-q}},$$

given that the estimates of the q lagged innovations are available, so that we obtain

$$\hat{y}_S = \hat{\phi}_1 \epsilon_{S-1} + \hat{\phi}_2 \epsilon_{S-2},$$

for $t = 1, \ldots, S-1$, where S is the beginning of the nowcast period. See Pindyck and Rubinfeld (1998) as well as the user guide of Eviews 7.0 as sources and for further information on the nowcasting procedure.

4 Reuters Sentiment and Stock Returns[24]

4.1 Introduction

4.1.1 Related Literature

The Efficient Market Hypothesis (EMH), introduced by Fama (1970), has been questioned widely on the grounds of psychological phenomena occurring in financial markets. Financial economists and psychologists alike have devoted time to research that relates sentiment among investors to financial market returns, such as Shiller (1981) who notes that financial markets display excess volatility and overreaction to new information. Summers (1986) then posed the question whether the stock market rationally reflects fundamental values and came to the conclusion that most tests of market efficiency have had little power to solidify the EMH. Thus, one should not conclude erroneously that market prices represent rational assessments of fundamental valuations based on the grounds that many studies have found that the EMH cannot be rejected. De Bondt and Thaler (1985) and Cutler *et al* (1989) identified a link between news and stock prices. Since then, studies have centered around the potential influence that the media has on investor behavior. DeLong *et al* (1990) were among the first to find that investors are subject to news. Barberis *et al* (1998) show that news can cause both over- and underreaction to stock prices by formulating a parsimonious model of investor sentiment. They claim that news are incorporated only slowly into stock prices. Hence, there are two main points to consider when dealing with the relationship between sentiment and stock returns: first, how to measure sentiment is crucial, as there are

[24] This chapter is based on Uhl (2011b).

numerous ways of measuring it, and, second, the time-frame is important to consider when examining over- and underreaction to news.

With regards to the first point, for example, Klibanoff *et al* (1998) show that country-specific news reported on the front page of the *New York Times* affect the pricing of closed-end country funds. Huberman and Regev (2001) find that an article in the *Financial Times* on a biochemical firm made prices of that company soar. Antweiler and Frank (2004) consider the influence of Internet stock message boards.[25] Baker and Wurgler (2007) argue that the key nowadays for researchers is to find out how to measure investor sentiment and quantify its effects. Owing to the quest for more accuracy in explaining financial market returns from a behavioral point of view, studies have been aiming towards the quantification of sentiment recently. Tetlock (2007) is one of the first to quantitatively measure the interactions between the media and the stock market using daily content from a *Wall Street Journal* column. High media pessimism, he finds, predicts falling stock market prices followed by a reversion to fundamentals. Unusually high or low pessimism predicts high trading volume as well.[26] In a follow-up to Tetlock's (2007) study, Tetlock *et al* (2008) use a simple quantitative measure of language to predict individual firms' accounting earnings and stock returns. Linguistic media content, they conclude, captures aspects of firms' fundamentals that are otherwise hard to quantify, which are quickly incorporated into stock prices. With regards to measuring sentiment, Tetlock (2007) uses the General Inquirer (GI), a quan-

[25] See also Cao and Wei (2005), Edmans et al (2007), Hirshleifer (2001), Hirshleifer and Shumway (2003), Kamstra et al (2003), Maier (2005), Mullainathan and Shleifer (2005), and Yuan et al (2006), among others.

[26] Note: low pessimism is not equal to optimism, as Tetlock (2007) only considers negative words – a measure only for pessimism – in his analysis.

titative content analysis program.[27] As explained in the appendix in Tetlock (2007), the GI has one major shortcoming: it is only able to distinguish between positive and negative words, or sentiment categories, but not between context, while Tetlock (2007) only considers negative sentiment.

Thus, we introduce and test a dataset that measures sentiment quantitatively in a systematic way, while trying to avoid subjectivity bias. With the growing importance of the media in the past decades, the obvious publicly available information are news, as De Bondt and Thaler (1985) as well as Cutler *et al* (1989) noted as early as a few decades ago. Given that the financial markets have mainly two major news providers, i.e. Bloomberg and Reuters, we consider one of the two – Reuters – providers from which we have quantified sentiment data of all Reuters news pieces from January 2003 to December 2010. With *Reuters sentiment,* we mean a (positive or negative) feeling, opinion, or emotion evoked among a reader while reading a certain Reuters news article. As opposed to Tetlock's (2007) dataset, the sentiment classifier used in this study is able to account for both individual words and context in the sentiment analysis through a newly developed sentiment algorithm by Thomson Reuters. Tetlock's (2007) study and findings thus serve as motivation, as we identify the need to not only consider the predictive power of negative sentiment on stock returns, but also of positive sentiment.

With regards to the second point central to this study, we consider the frequency of the sentiment models. Tetlock (2007) uses daily data, while he finds a sentiment effect that lasts for several days. The studies by Tetlock *et al* (2008) and Tetlock (2011) also consider daily news sentiment data. Other

[27] See *The General Inquirer Home Page*, available at http://www.wjh.harvard.edu/~inquirer/, last accessed 23 November 2010.

studies consider longer time-frames. However, these studies do not look at news sentiment, but rather at investor sentiment, such as Brown and Cliff (2005). They examine monthly investor sentiment (gathered from a direct survey measure among investors) and show that asset prices revert to intrinsic values, whereas over longer horizons, high sentiment leads to overvaluation. They further find that sentiment is significantly related to long-run stock returns, even up to several years. This is especially present in large-capitalization growth stocks. They also use the Dow Jones Industrials stock index, which we do as well, as we want to test this long-run relationship between sentiment and stock returns. Menkhoff and Rebitzky (2008) consider investor sentiment in the US-dollar market and find that investor sentiment is connected to exchange rate returns at longer horizons, even more than two years. Livnat and Petrovits (2009) examine whether stock price reactions to earnings surprises and accruals vary systematically with the level of investor sentiment on a monthly basis.

To our knowledge, there are no other studies that consider the kind of news sentiment, i.e. Reuters sentiment, that we use in this study with regards to the time-frame and the Dow Jones Industrials Stock Index. Therefore, this study is the first to consider longer-time frames on the impact of news sentiment, i.e. Reuters sentiment, on stock returns. Using sentiment of Reuters news, we build Vector Autoregression (VAR) models. We find that both negative and positive Reuters sentiment have an impact on stock returns, although there are differences between the two sentiments. We also find that the news sentiment effect is measurable over months, and not only days. The remainder of this section describes the dataset. Section 4.2 discusses the empirical results, and section 4.3 concludes.

4.1.2 Dataset

We analyze both positive and negative sentiment in relation to stock returns. The sentiment scores are not only obtained through simply coding positive and negative words according to a database. Owing to new technological advance in text mining, Thomson Reuters is able to undertake a sentiment analysis that takes the context into account. For example, the sentiment algorithm is able to distinguish between negative words and negations of positive words. "Good" would be categorized as positive in the sentiment analysis, but "not good" would be classified as negative. This has not been possible so far in textual mining programs that are based on a pre-defined databases of positive and negative words only. Thus, Reuters allows us to contribute to the literature with a more precise methodological approach as opposed to earlier studies.

Based on this dataset, we introduce the concept of measuring sentiment in Reuters news articles quantitatively in order to explain stock returns. Every Reuters news article is coded with positive $\{1\}$, neutral $\{0\}$, or negative $\{-1\}$ sentiment. In the past, most solutions have come from the text mining industry that caters to the financial markets industry, in which news texts can be scanned in great quantities and a short amount of time for sentiment with specific sentiment algorithms. Thomson Reuters is one of the few providers of sentiment classified news.[28] The dataset at hand consists of high-frequency (tick data) sentiment rated Thomson Reuters news pieces, classified from a wide list of topics for the US market.[29] For this study, we filter all Reuters

[28] See Thomson Reuters News Analytics, http://thomsonreuters.com/products_services/financial/financial_products/quantitative_research_trading/news_analytics, last accessed 7 September 2010.

[29] The topics range from financial market to economic and political news, categorized into topic codes. See *Reuters Codes – A quick guide*, available at

news items for sentiment from the Equities topic codes section.[30] Then, we extract both positive and negative sentiment values in order to form two independent time-series in order to aggregate the tick sentiment scores to monthly values. In total, over 3.6 million Reuters news items were coded for sentiment from January 2003 to December 2010.

Monthly price return data for the Dow Jones Industrials stock index were obtained from Thomson Reuters Datastream. The corresponding monthly volume data for the Dow Jones stock index are from MasterData.[31] To capture the real macroeconomic development, we use a time series of the Conference Board Leading Economic Indicators Index (LEI). This index consists of a combination of leading indices, such as production, employment, monetary, and consumer data for the US.[32] The advantage over using many different indicators is that one variable is easier to handle in our subsequent model than multiple variables. Given that we attempt to explain stock returns with non-conventional measures – inconsistent with the EMH – such as sentiment, we need to include fundamental facts that are consistent with the EMH to capture all possible channels of influence on the stock index, and to compare the fundamental to the behavioral. The Conference Board Leading Economic Indicators Index appears the most suited for "summarizing" macroeconomic factors in one variable. Monthly data were obtained from Datastream.

---

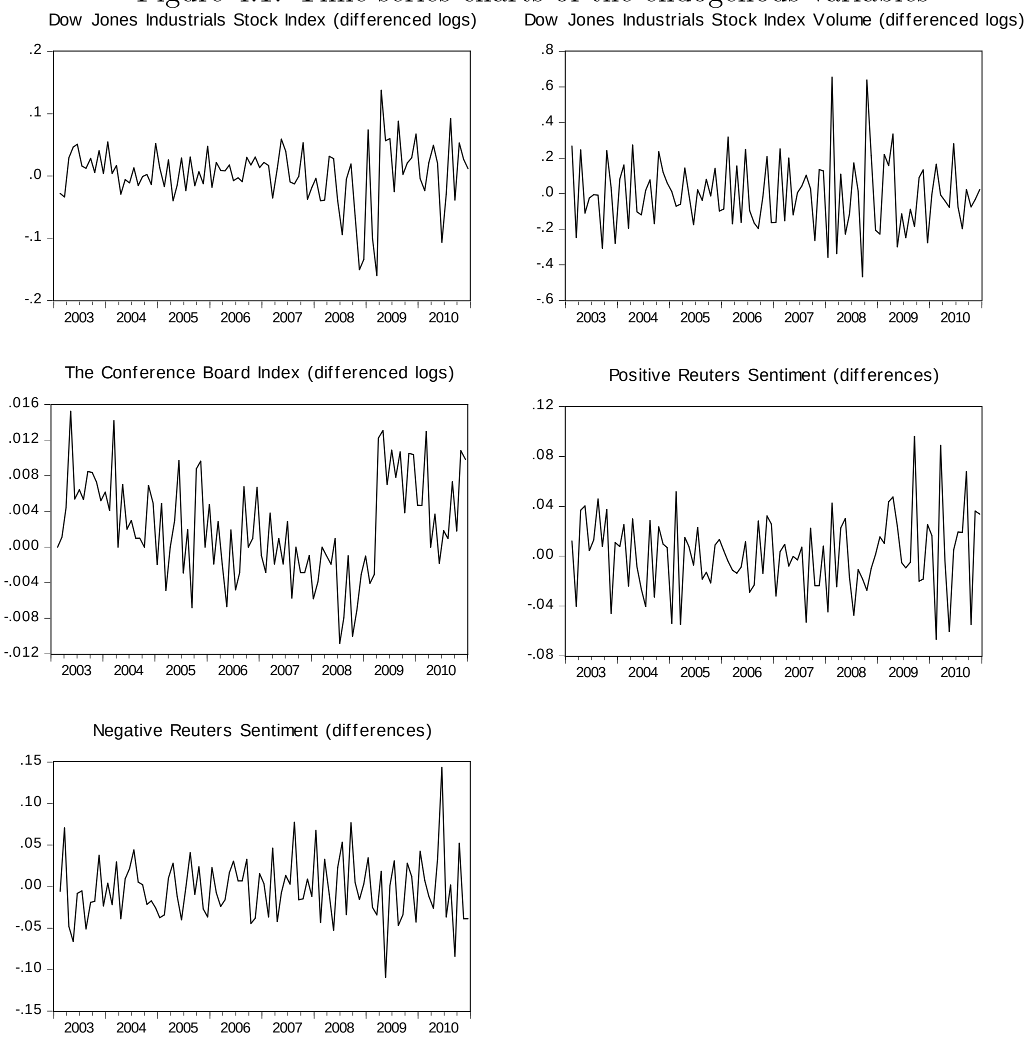
https://customers.reuters.com/training/trainingCRMdata/promo_content/ReutersCodes.pdf, last accessed 9 December 2010.

[30] We filter for "U" in the product code section, and for "DIV, MRG, RES, RESF, RCH, STX" in the topic code section. These codes mean that we filter for news related to dividends, ownership changes, broker research, corporate results, results forecasts and stock markets for North American companies.

[31] See www.masterdatacsv.com, last accessed 15 October 2010.

[32] See *Global Business Cycles Indicators* for more detailed information at http://www.conference-board.org/economics/bci, last accessed 7 December 2010.

Figure 4.1: Time-series charts of the endogenous variables



To get a first understanding of the data, we look at the variables graphically in fig. 4.1. In order to account for non-stationarity in the data, we take the differenced logs of the Dow Jones stock index prices and volume as well as of the LEI. We simply difference both positive and negative sentiment, as we do not need to logarithmize these time-series, as they are range-bound between -1 and 1 by definition. All five variables appear stationary. Yet, we can make out the recent financial crisis in 2008/2009 and the subsequent rise in stock prices and macroeconomic conditions, as depicted in the LEI. We need to undertake further empirical tests to find out whether a combination of fundamental data, i.e. the LEI, and behavioral data, i.e. Reuters sentiment, can explain changes in stock prices.

4.2 Empirical Analysis

4.2.1 The Model and Empirical Results

We construct a Vector AutoRegression model (VAR) in order to tackle possible endogeneity issues in the data. Before formulating the model, we first test for the optimal lag length with various optimal endogenous lag length tests, as we want to identify a possible lead-lag relationship between the endogenous variables. Table 4.1 shows the results.

The five tests that were carried out to identify the optimal lag structure of the model make a fairly clear case for three lags. Three out of the five tests indicate that an endogenous lag structure of three is optimal. Only the Schwarz information and Hannan-Quinn information criteria show an optimal lag structure of 0 and 1, respectively. Therefore, we formulate the VAR-model

Table 4.1: Optimal Endogenous Lags for the VAR-model

Vector Autoregression Lag Order Selection Criteria

Endogenous variables: Dow Jones Industrial Stock Index (differenced logs), Dow Jones Industrial Stock Index Volume (differenced logs), The Conference Board Leading Economic Indicator Index (differenced logs), Positive Reuters Sentiment (differences), Negative Reuters Sentiment (differences)
Exogenous variables: Constant
Sample: 2003M01 2010M12

Lag	LogL	LR	FPE	AIC	SC	HQ
0	860.4344	NA	1.98E-15	-19.66516	-19.52344*	-19.60809
1	905.6922	84.27309	1.25E-15	-20.13085	-19.28054	-19.78846*
2	940.8195	61.37185	9.93E-16	-20.36367	-18.80476	-19.73594
3	968.6976	45.50224*	9.42e-16*	-20.42983*	-18.16233	-19.51678
4	989.9269	32.20993	1.06E-15	-20.34315	-17.36705	-19.14477
5	1014.303	34.18317	1.12E-15	-20.32881	-16.64412	-18.8451
6	1032.419	23.3214	1.41E-15	-20.17056	-15.77727	-18.40152
7	1043.618	13.12988	2.14E-15	-19.85329	-14.75141	-17.79892
8	1064.612	22.20027	2.7E-15	-19.76119	-13.95072	-17.42149

* indicates lag order selected by the criterion
LR: sequential modified LR test statistic (each test at 5% level)
FPE: Final prediction error
AIC: Akaike information criterion
SC: Schwarz information criterion
HQ: Hannan-Quinn information criterion

The first column of the table shows the lag order of the VAR-model. The next column shows the Log Likelihood of the system (LogL). The next five columns show the relevant lag order selection criteria tests, as explained in the table. The LR, FPE, and AIC tests indicate an optimal lag length of 3, where the SC and HQ suggest a lag length of 0 and 1, respectively.

allowing up to three endogenous lags. We thus construct the model in order to discern more closely how suited Reuters sentiment is for explaining stock returns as follows:

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + A_3 y_{t-3} + BD_t + u_t, \tag{4.1}$$

where y_t is a vector of k observable endogenous variables, D_t is a deterministic variable containing a constant, and u_t is a k- dimensional unobservable zero mean white noise process. The endogenous variables in this model are Dow Jones Industrials stock index, positive and negative Reuters sentiment, Dow Jones Industrials stock index volume, and the Conference Board Index.

First, we perform several robustness checks in order to show that the VAR model is well-suited for this kind of analysis. In table 4.2, we perform lag exclusion Wald tests that test each lag in the VAR individually, as in Lütkepohl (1991). For each lag, we report the χ^2 (Wald) statistic for the joint significance of the endogenous variables at each lag from one to three. The statistics are reported for each endogenous variable and also for all endogenous variables jointly. The test shows that all three lags are highly statistically significant for the joint test. For most of the lags, the test statistics are statistically significant, justifying the lag structure of the VAR-model.

Second, we perform tests for non-normality, as shown in table 4.3. We perform two different tests for non-normality to achieve higher confidence of the robustness of the model. The first test for non-normality is modeled according to Doornik & Hansen (1994), while the second is modeled according to Lütkepohl (1993). With these tests, the residual vector is transformed such that its components are independent. We then check the compatibility of the third and

Table 4.2: Lag Exclusion Wald Tests

VAR Lag Exclusion Wald Tests

Sample: 2003M01 2010M12

Chi-squared test statistics for lag exclusion:
Numbers in [ ] are p-values

	Dow Jones Industrial Stock Index (differenced logs)	Positive Reuters Sentiment (differences)	Negative Reuters Sentiment (differences)	Dow Jones Industrial Stock Index Volume (differenced logs)	The Conference Board Leading Economic Indicator Index (differenced logs)	Joint
Lag 1	27.24876***	12.30064**	18.1703***	32.27614***	10.37978*	108.7613***
	[5.10e-05]	[0.030893]	[0.002740]	[5.24e-06]	[0.065162]	[2.02e-12]
Lag 2	20.32065***	13.79196**	6.006862	20.63985***	14.17151**	70.06319***
	[0.001088]	[0.016986]	[0.305552]	[0.000947]	[0.014556]	[3.77e-06]
Lag 3	17.206***	9.890945*	6.375122	4.84292	10.00275*	48.46066***
	[0.004125]	[0.078385]	[0.271409]	[0.435350]	[0.075157]	[0.003283]
df	5	5	5	5	5	25

*, **, *** indicate statistical significance at the 10%-, 5%-, and 1%-level, respectively

This table shows Chi-squared test statistics for lag exclusion of the VAR model. The exclusion is rejected when the test statistic coefficient is statistically significant. The last column shows the joint test that includes all endogenous variables, justifying the lag order of three of the VAR-model.

Table 4.3: Tests for Non-normality

Tests for Nonnormality	
H0: Sample has nonnormal distribution	
Reference: Doornik & Hansen (1994)	
joint test statistic:	70.8709***
p-value:	0
degrees of freedom:	10
skewness only:	26.8373***
p-value:	0.0001
kurtosis only:	44.0336***
p-value:	0
Reference: Lütkepohl (1993)	
joint test statistic:	46.5621***
p-value:	0
degrees of freedom:	10
skewness only:	21.4681***
p-value:	0.0007
kurtosis only:	25.0939***
p-value:	0.0001

*** indicates statistical significance at the 1%-level

This table shows various tests for non-normality in the data. According to all test results, we can reject the null hypothesis that the sample has a non-normal

fourth moments with those of a normal distribution. We test the Null hypothesis that the sample has a non-normal distribution. The difference between these two methods is that Lütkepohl (1993) used a Cholesky decomposition of the residual covariance matrix. Both test statistics reject the Null hypothesis with a 1%-level of statistical significance that the sample has a non-normal distribution.

Third, we test for serial correlation and heteroskedasticity in the model. Table 4.4 shows two relevant tests that are carried out. We perform a VAR residual Portmanteau test for autocorrelation with the the multivariate Box-Pierce/Ljung-Box Q-statistics for residual serial correlation up to the lag order of 8. This test is modeled in accordance with Lütkepohl (1991). According to the test statistics, we note that the null hypothesis of no residual autocorrelation up to lag 8 cannot be rejected. The second test shown in table 4.4 that we perform is the ARCH-LM test that reports the multivariate LM test statistics for residual serial correlation. This test is specified as in Engle (1982). For all five variables that we have in the model, depicted by $u1 \dots u5$ in the table, we cannot reject the null hypothesis of no autoregressive conditional heteroskedasticity. This bodes well for our model, as we cannot reject the null hypothesis of serial correlation and heteroskedasticity in the model.

We empirically test equation (4.1) to obtain further clues whether Reuters sentiment as well as other variables can account for changes in stock returns. Table 4.5 shows the results of the VAR estimation, allowing for up to three lags, as specified by the various tests above. We note that positive Reuters sentiment is statistically significant at lag one, while negative Reuters sentiment is statistically significant at lag three. Therefore, there are possibly differences in the information content of positive and negative Reuters sentiment. Further, Dow Jones volume is statistically significant at lag two, whereas the Conference Board Index is not statistically significant at any of the lags. What can we draw from these observations? DeLong *et al* (1990) identify two sets of traders in the market: professional arbitrageurs and unsophisticated traders (i.e. noise traders). The prevailing risk in the market, they find, is created by the unpredictability of the noise traders. Professional arbitrageurs respond

Table 4.4: Tests for Autocorrelation and Heteroskedasticity

Vector Autoregression Residual Portmanteau Tests for Autocorrelations
H0: No residual autocorrelations up to lag h

Lags	Q-Stat	Prob.	Adj Q-Stat	Prob.	df**
1	4.634807	NA*	4.685739	NA*	NA*
2	17.17037	NA*	17.49987	NA*	NA*
3	26.99108	NA*	27.65161	NA*	NA*
4	47.75162	0.4014	49.35581	0.3406	46
5	70.94641	0.4795	73.88364	0.3842	71
6	101.7126	0.3255	106.7963	0.212	96
7	119.5008	0.5215	126.0495	0.3583	121
8	133.7872	0.7569	141.6964	0.5852	146

* The test is valid only for lags larger than the VAR lag order. In our model, we have a VAR lag order of 3.
** df is degrees of freedom for (approximate) chi-square distribution

ARCH-LM TEST with 16 lags
H0: No autoregressive conditional heteroskedasticity

variable	teststat	p-value(χ^2)	F stat	p-value(F)
u1	12.8775	0.6817	0.969	0.5006
u2	15	0.5104	1.187	0.305
u3	17	0.3647	1.4029	0.1722
u4	5	0.9963	0.3268	0.992
u5	18	0.2995	1.5202	0.1232

to the behavior of noise traders rather than acting on fundamentals. In doing so, professional arbitrageurs consider pseudo signals, such as volume and price patterns, but also news. Taking DeLong *et al*'s (1990) model into account, our findings can be interpreted as follows: since the professional traders consider pseudo signals, such as volume and news, they generate stock price fluctuations.

However, they interpret various variables differently, especially with regards to the timing. For instance, positive Reuters sentiment seems to matter the most after one month, while negative Reuters sentiment appears the most relevant with a lag of three months. Volume is also an indicator of subsequent stock returns, while it is most significant with a lag of two months. Fundamental data, as measured by the Conference Board Leading Economic Indicator, does not seem to have a great influence on a monthly basis. It is likely that news about fundamental macroeconomic data are immediately incorporated into stock prices, while Reuters sentiment appears to be more persistent.

The frequency of our model is also important when we consider the model of Barberis *et al* (1998) who find that news are incorporated only slowly into stock prices. Given the monthly frequency framework, we show that sentiment from Reuters news is incorporated only slowly into stock prices, as traders react to sentiment with a time lag of up to three months, while it depends on the type of sentiment how fast it is incorporated into stock prices. Negative sentiment takes longer to "digest" while positive sentiment is incorporated faster. This is in line with the common belief that "bad news sell," while positive news do not seem to matter that much on a longer time horizon.

To analyze the dynamic interactions between the endogenous variables of the VAR process, we draw on the impulse response analysis. The impulse

Table 4.5: Vector Autoregression Coefficient Estimates

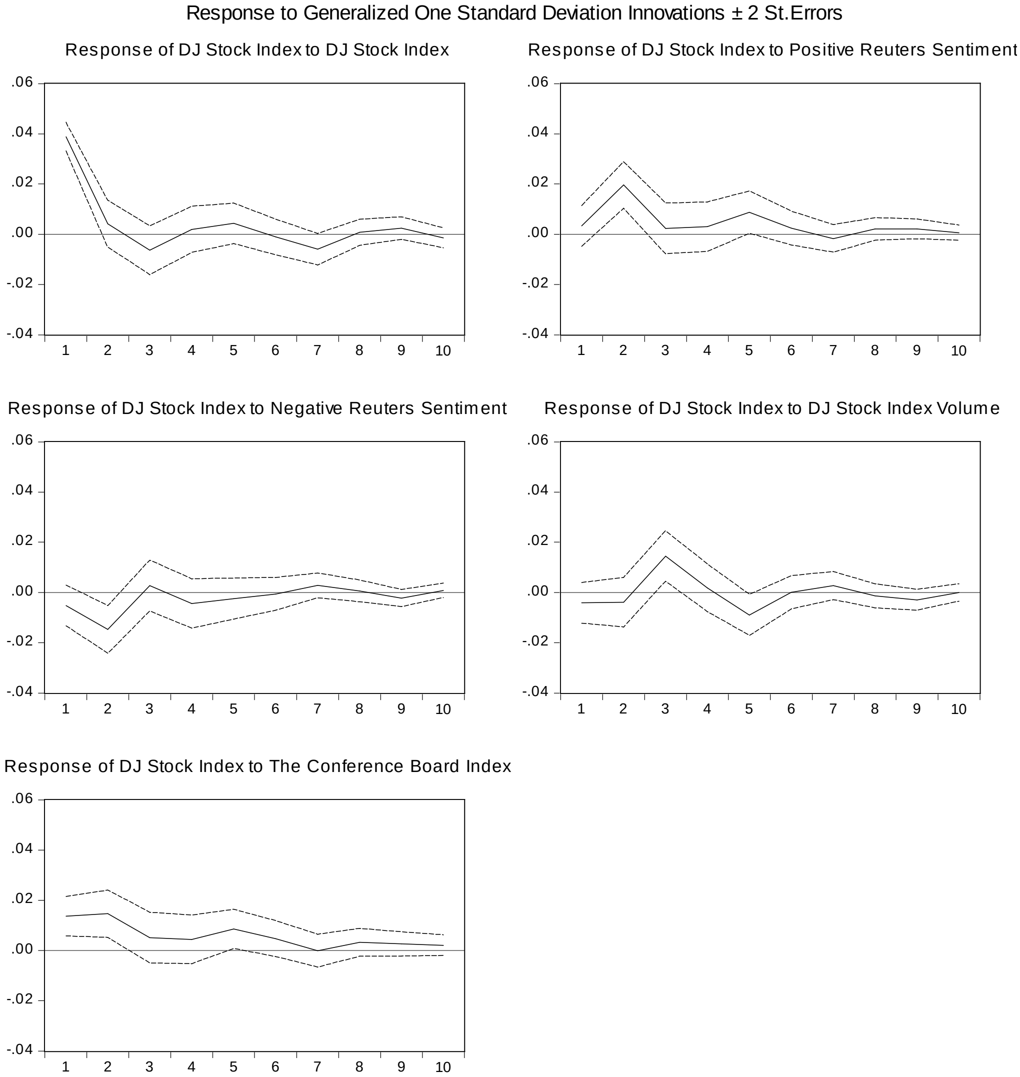
Vector Autoregression Estimates

Coefficients of the VAR model, t-statistics in [ ] below coefficients
Sample (adjusted): 2003M05 2010M12

	Dow Jones Industrial Stock Index (differenced logs)	Positive Reuters Sentiment (differences)	Negative Reuters Sentiment (differences)	Dow Jones Industrial Stock Index Volume (differenced logs)	The Conference Board Leading Economic Indicator Index
Dow Jones Industrial Stock Index (differenced logs) (-1)	-0.026706	-0.043994	-0.1151	-1.004787**	0.006737
	[-0.23606]	[-0.52958]	[-1.14369]	[-2.13456]	[0.48785]
Dow Jones Industrial Stock Index (differenced logs) (-2)	-0.2508**	-0.277006***	0.146297	0.685554	-0.029429**
	[-2.17819]	[-3.27631]	[1.42830]	[1.43096]	[-2.09390]
Dow Jones Industrial Stock Index (differenced logs) (-3)	0.331017***	-0.112241	0.119753	-0.651486	0.004221
	[2.85185]	[-1.31690]	[1.15979]	[-1.34896]	[0.29793]
Positive Reuters Sentiment (differences) (-1)	0.508469**	-0.467022***	-0.028607	-0.797366	0.018753
	[2.51365]	[-3.14415]	[-0.15898]	[-0.94736]	[0.75949]
Positive Reuters Sentiment (differences) (-2)	0.130718	-0.09974	-0.077865	1.128908	0.02824
	[0.62336]	[-0.64774]	[-0.41741]	[1.29384]	[1.10326]
Positive Reuters Sentiment (differences) (-3)	0.115493	0.244731*	-0.069082	0.022856	0.0468**
	[0.59737]	[1.72386]	[-0.40167]	[0.02841]	[1.98307]
Negative Reuters Sentiment (differences) (-1)	-0.183135	-0.105546	-0.340575***	1.377173**	-0.011779
	[-1.23741]	[-0.97121]	[-2.58685]	[2.23640]	[-0.65200]
Negative Reuters Sentiment (differences) (-2)	0.223575	0.06733	-0.189578	-0.34082	0.013454
	[1.39030]	[0.57019]	[-1.32522]	[-0.50937]	[0.68543]
Negative Reuters Sentiment (differences) (-3)	-0.304507**	0.213922**	-0.013291	0.672918	0.021399
	[-2.08758]	[1.99722]	[-0.10243]	[1.10873]	[1.20183]
Dow Jones Industrial Stock Index Volume (differenced logs) (-1)	-0.043327	-0.011047	-0.029407	-0.397572***	0.004753
	[-1.52500]	[-0.52951]	[-1.16353]	[-3.36318]	[1.37058]
Dow Jones Industrial Stock Index Volume (differenced logs) (-2)	0.072058**	-0.033387	-0.003014	-0.376673***	0.000857
	[2.38435]	[-1.50450]	[-0.11212]	[-2.99552]	[0.23235]
Dow Jones Industrial Stock Index Volume (differenced logs) (-2)	0.035296	-0.027906	0.012274	-0.076817	-0.002478
	[1.28346]	[-1.38190]	[0.50172]	[-0.67133]	[-0.73823]
The Conference Board Leading Economic Indicator Index (differenced logs) (-1)	1.418229	1.424169*	-2.373959**	2.83503	0.174477
	[1.29082]	[1.76525]	[-2.42889]	[0.62015]	[1.30097]
The Conference Board Leading Economic Indicator Index (differenced logs) (-2)	1.075218	1.054724	-1.141515	-3.627065	0.340013***
	[1.00312]	[1.34005]	[-1.19716]	[-0.81326]	[2.59873]
The Conference Board Leading Economic Indicator Index (differenced logs) (-3)	-0.408501	0.412525	1.503371	-2.014655	0.179943
	[-0.38083]	[0.52374]	[1.57551]	[-0.45140]	[1.37431]
Constant	-0.002735	-0.002448	0.000416	0.007823	0.000755
	[-0.56252]	[-0.68562]	[0.09617]	[0.38668]	[1.27279]
R-squared	0.450392	0.297527	0.284339	0.431786	0.425044
Adj. R-squared	0.341916	0.158881	0.14309	0.319639	0.311565
Log likelihood of the system	1023.555				

Note: *,**,*** indicate statistical significance at the 10%-,5%-, and 1%-level, respectively.

Figure 4.2: Impulse Responses from VAR-model



responses are modelled according to the generalized impulse response method according to Pesaran and Shin (1998).

Fig. 4.2 shows the results of the impulse responses based on the VAR model in equation (4.1). We focus on the responses of the Dow Jones Industrials Index stock returns to the endogenous variables. The response of stock returns to Reuters positive and negative sentiment are both greatest after two months. Negative sentiment shocks stock returns negatively, while positive sentiment shocks stock returns positively. This finding makes sense, as one would expect that declines in stock returns are followed by negative sentiment, while price gains are followed by positive sentiment in Reuters news. The responses of stock returns to the Conference Board Index are similar for months one and two after the shock, although the response from the Conference Board Index is not as pronounced as the one from the sentiment variables. After two months, the shocks gradually disappear. Volume of the Dow Jones index does not have an immediate effect on stock returns. Hong and Stein (1999) show theoretically that prices overreact in the long run. Our findings conform with their assumption that stock returns are prone to overreact in longer time frames. In this study, we show that both positive and negative Reuters sentiment can cause such an overreaction of stock prices, while the effect is measurable for months after the initial shock. On the one hand, this finding contradicts Tetlock (2007) who showed that the sentiment effect is only present for a maximum of a few days, claiming that sentiment is incorporated quickly into stock prices. On the other hand, our findings are in line with Livnat and Petrovits (2009) who account for a post-earnings announcement drift among investor sentiment, but not news sentiment.[33] Hence, our findings suggest a news sentiment drift. On a longer time-frame, our findings do not contradict with the studies of Brown and Cliff (2005) as well as Menkhoff and Rebitzky (2008) who show that investor sentiment is connected with both stock and exchange rate returns for months and even years. While we do not find evidence of a yearly impact of Reuters sentiment on stock returns, we find evidence of a relationship between Reuters sentiment and stock returns for several months.

[33] See also Chan (2003) for evidence of a post-news drift.

In order to get a clearer understanding about how much of the variation in stock returns is driven by sentiment as well as fundamentals, we draw on the Forecast Error Variance Decomposition (FEVD). The FEVD separates the variation in an endogenous variable into the component shocks to the VAR. The FEVD provides information about the relative importance of each random innovation in affecting the variables in the VAR. Denoting the ij-th element of the orthogonalized impulse response coefficient matrix ψ_n, the variance of the forecast error $y_{k,T+h} - y_{k,T+h|\mathrm{T}}$ is

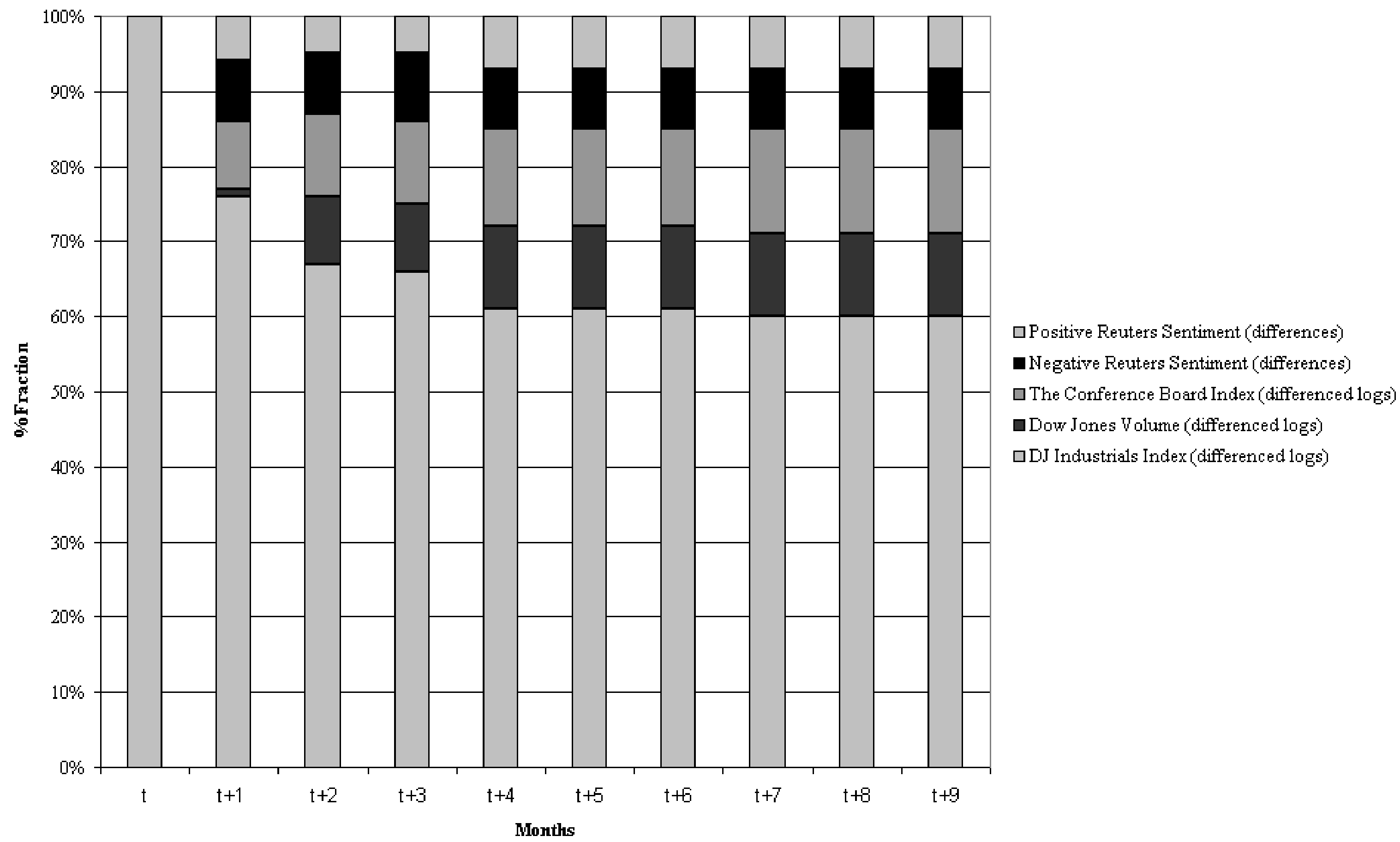
$$\sigma_k^2(h) = \sum_{n=0}^{h-1} \left(\psi_{k1,n}^2 + \cdots + \psi_{kK,n}^2\right) = \sum_{j=1}^{K} \left(\psi_{kj,0}^2 + \cdots + \psi_{kj,h-1}^2\right). \tag{4.2}$$

The term $\left(\psi_{kj,0}^2 + \cdots + \psi_{kj,h-1}^2\right)$ is considered as the contribution of variable j to the h-step forecast error variance of variable k. Dividing the above terms by $\sigma_k^2(h)$ gives us the contribution of variable j to the h-step forecast error variance of variable k in percent,

$$\omega_{kj}(h) = \left(\psi_{kj,0}^2 + \cdots + \psi_{kj,h-1}^2\right) / \sigma_k^2(h).$$

The FEVD of Dow Jones Stock index returns is depicted in fig. 4.3. The impact of the economic factors, in the form of the Conference Board Index, makes up around 10–13% of the variance of the forecast error of stock returns. However, the largest share have positive and negative Reuters sentiment, making up around 15–20% of the variance of the forecast error of stock returns. Dow Jones Industrials volume only attributes to about 10% of the variation in stock returns. This is in line with our empirical results from the VAR and

Figure 4.3: Forecast Error Variance Decomposition of Dow Jones Industrials Stock Index



the impulse response functions, speaking for Reuters sentiment as a relevant variable to influence stock returns.

4.2.2 Summary of Results

According to the various tests and analyses that we have undertaken, we stress four major findings. First, positive and negative Reuters sentiment have an impact on stock returns, although the impact is different, especially with regards to time and content. Negative Reuters sentiment appears to have a higher influence on stock returns than positive Reuters sentiment, while negative Reuters sentiment is more persistent than positive Reuters sentiment. Second, volume has a measurable influence on stock returns with a lag of two months. Third, fundamental factors, such as the Conference Board Leading Economic Indicator, do not have a measurable effect on stock returns. The reason for this is possibly that fundamental news are incorporated quickly into stock prices, at most within one month. Hence, we are not able to measure a statistically significant effect in our monthly model. And, last but not least, the main contribution of this paper lies in the identification of a longer-lasting news sentiment effect that is present over months, and not only days.

4.3 Conclusion

Based on recent findings in the literature on the impact of news and investor sentiment on stock returns, we examine specifically two research questions in this chapter. First, is not only negative news sentiment relevant for determining stock price fluctuations, but does also positive news sentiment have an impact on stock returns? Second, while the news sentiment effect has only been considered in shorter time-frames, such as daily time intervals in Tetlock (2007), we consider, based on the literature about investor sentiment, whether there is a longer-lasting effect of news sentiment on stock returns. In order to test this, we use a monthly dataset with sentiment values that were aggregated from over 3.6 million Reuters news articles. We call the news sentiment measure employed in this study Reuters sentiment. Furthermore, we not only consider negative news sentiment, as Tetlock (2007) does, but also positive news sentiment.

We show with vector autoregression models, impulse response functions and a forecast error variance decomposition analysis that positive and negative Reuters sentiment have an impact on stock returns, while differences exist between the sentiment measures with regards to time and content. Negative Reuters sentiment appears to be more persistent than positive Reuters sentiment. Fundamental factors, such as the Conference Board Leading Economic Indicator, have also a significant effect on stock returns in a monthly framework. We conclude that while fundamental news matter, there is evidence of a news sentiment drift. Hence, the main contribution of this paper is that there is a longer-lasting news sentiment effect present in Reuters news that can be measured over several months.

References

[1] Acemoglu, D. and A. Scott (1994) Consumer Confidence and Rational Expectations: Are Agents' Beliefs Consistent with the Theory? *The Economic Journal*, 104, 1–19.

[2] Akaike, H. (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control,* AC–19, 716–723.

[3] Ang, A., G. Bekaert, and M. Wei (2007) Do macro variables, asset markets, or surveys forecast inflation better? *Journal of Monetary Economics*, 54, 1163–1212.

[4] Antweiler, W. and M.Z. Frank (2004) Is all that talk just noise? The information content of internet stock message boards. *Journal of Finance,* 59, 1259–1293.

[5] Armstrong, D.M. (1961) *Perception and the Physical World*, London.

[6] Baker, M. and J. Wurgler (2007) Investor sentiment and the cross-section of stock returns. *Journal of Finance,* 61, 1645–1680.

[7] Barberis, N., A. Shleifer, and R. Vishny (1998) A model of investor sentiment. *Journal of Financial Economics,* 49, 307–343.

[8] Baron, D.P. (2006) Persistent Media Bias. *Journal of Public Economics*, 90, 1–36.

[9] Behr, R. L., and S. Iyengar (1985) Television News, Real-World Cues, and Changes in the Public Agenda. *Public Opinion Quarterly*, 49, 38–57.

[10] Bouakez, H., and R. Nooman (2007) Why does private consumption rise after a government spending shock? *Canadian Journal of Economics*, 40, 954–979.

[11] Box, G., and G. Jenkins (1970) *Time Series Analysis Forecasting and Control*, Holden-Day, San Francisco.

[12] Breeden, D.T. (1986) Consumption, Production, Inflation and Interest Rates. *Journal of Financial Economics,* 16, 3–39.

[13] Breitung, J., R. Brüggemann, and H. Lütkepohl (2004) Structural vector autoregressive modelling and impulse responses. *In* H. Lütkepohl and M. Krätzig (eds.), *Applied Time Series Econometrics*. Cambridge: Cambridge University Press.

[14] Breusch, T. S., and A. R. Pagan (1979) A Simple Test for Heteroskedasticity and Random Coefficient Variation. *Econometrica,* 48, 1287–1294.

[15] Bring, J. (1994) How to Standardize Regression Coefficients. *The American Statistican*, 48, 209–213.

[16] Brown, G.W. and M.T. Cliff (2005) Investor Sentiment and Asset Valuation. *The Journal of Business*, 78 (2), 405–440.

[17] Campbell, J. and G. Mankiw (1989) Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence, in O. Blanchard and S. Fischer (eds.*) NBER macroeconomics annual 1989*, MIT Press, Cambridge, MA.

[18] Cao, M. and J. Wei (2005) Stock market returns: A note on temperature anomaly. *Journal of Banking and Finance,* 29, 1559–1573.

[19] Carroll, C. (2003) Macroeconomic Expectations of Households and Professional Forecasters. *The Quarterly Journal of Economics*, 118, 269–298.

[20] Carroll, C. (2009) Precautionary Saving and the Marginal Propensity to Consume Out of Permanent Income. *Journal of Monetary Economics*, 56, 780–790.

[21] Carroll, C., J. Fuhrer, and D. Wilcox (1994) Does Consumer Sentiment Forecast Household Spending? If So, Why? *American Economic Review,* 84, 1397–1408.

[22] Carroll, C., J. Slacalek, and M. Sommer (2010) International Evidence On Sticky Consumption Growth. *Review of Economic and Statistics.*

[23] Case, K.E., J.M. Quigley, and R.J. Shiller (2003) Home-buyers, Housing and the Macroeconomy. *Berkeley Working Papers*, ejournals.

[24] Chan, W.S. (2003) Stock Price Reaction to News and No-news: Drift and Reversal after Headlines. *Journal of Financial Economics,* 70, 223–260.

[25] Curtin, R.T. (1982) Indicators of Consumer Behavior: The University of Michigan Surveys of Consumers. *Public Opinion Quarterly*, 46, 340–352.

[26] Curtin R. (2007) What US Consumers Know About Economic Conditions. *Second OECD Workshop on "Measuring and Fostering the Progress of Societies,"* Istanbul, June 2007.

[27] Cutler, D.M., J.M. Poterba, and L.H. Summers (1989) What moves stock prices? *Journal of Portfolio Management,* 15, 4–12.

[28] De Bondt, W.F.M. and R. Thaler (1985) Does the Stock Market Overreact? *Journal of Finance,* 40, 793–805.

[29] DellaVigna, S., Kaplan, E. (2007) The Fox News Effect: Media Bias and Voting. *The Quarterly Journal of Economics,* 122, 1187–1234.

[30] DeLong, J.B., A. Shleifer, L.H. Summers and R. J. Waldmann (1990) Noise trader risk in financial markets. *Journal of Political Economy,* 98, 703– 738.

[31] Dickey, D.A. and W.A. Fuller (1979) Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association,* 74, 427–431.

[32] Doms, M. and N. Morin (2004) Consumer Sentiment, the Economy, and the News Media. *FRBSF Working Paper.*

[33] Doornik, J. A. and H. Hansen (1994) A practical test of multivariate normality. *Unpublished paper*, Oxford University, Nuffield College.

[34] Edmans, A., D. Garcia, and O. Norli (2007) Sports Sentiment and Stock Returns. *Journal of Finance,* 62, 1967–1998.

[35] Efron, B. and Tibshirani, R. J. (1993) *An Introduction to the Bootstrap.* New York: Chapman & Hall.

[36] Engle, R. F. (1982) Autoregressive conditional heteroskedasticity, with estimates of the variance of United Kingdoms inflations. *Econometrica*, 50, 987–1007.

[37] Fama, E. (1970) Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance,* 25, 383–417.

[38] Fang, L., and J. Peress (2009) Media Coverage and the Cross-Section of Stock Returns. *Journal of Finance,* 64, 2023–2052.

[39] Friedman, M. (1957) *A Theory of the Consumption Function*, Princeton University Press, Princeton, NJ.

[40] Friedman, N., D. Geiger, and M. Goldszmidt (1997) Bayesian network classifiers. *Machine Learning,* 29, 131–163.

[41] Gabriel, K. R. (1971) The biplot-graphic display of matrices with application to principal components analysis. *Biometrika*, 58, 453–467.

[42] Gentzkow, M., and J. Shapiro (2010) What drives media slant? Evidence from U.S. daily newspapers. *Econometrica,* 78, 35–71.

[43] Godfrey, L. G. (1978) Testing for Multiplicative Heteroskedasticity. *Journal of Econometrics,* 8, 227–236.

[44] Granger, C.W.J. and P. Newbold (1974) Spurious Regressions in Econometrics. *Journal of Econometrics,* 2, 111–120.

[45] Hafer, R.W. and R.G. Sheehan (1989) The sensitivity of VAR forecasts to alternative lag structures. *International Journal of Forecasting*, 5, 399–408.

[46] Hall, R.E. (1978) Stochastic Implications of the Life-Cycle/Permanent Income Hypothesis: Theory and Evidence. *Journal of Political Economy,* 96, 971–987.

[47] Hand, D.J., and K. Yu (2001) Idiot's Bayes – not so stupid after all? *International Statistical Review,* 69, 385–399.

[48] Hansen, P.R. (2005) A test for superior predictive ability, *Journal of Business and Economic Statistics*, 23(4), 365–380.

[49] Harris Interactive (2007) TV Network News Top Source of News and Information Today. *The Harris Poll*, 52, June 11, 2007.

[50] Hayashi, F. (1982) The Permanent Income Hypothesis: Estimation and Testing by Instrumental Variables. *The Journal of Political Economy,* 90, 895–916.

[51] Hendry, D.F. (1993) *Econometrics: Alchemy or Science? Essays in Econometric Methodology*, Blackwell Publishers, Oxford.

[52] Hirshleifer, D. (2001) Investor psychology and asset pricing. *Journal of Finance,* 91, 342–346.

[53] Hirshleifer, D. and T. Shumway (2003) Good day sunshine: Stock returns and the weather. *Journal of Finance,* 58, 1009–1032.

[54] Hong, H. and J.C. Stein (1999) A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets. *Journal of Finance,* 54, 2143–2184.

[55] Huberman, G. and T. Regev (2001) Contagious speculation and a cure for cancer: A nonevent that made stock prices soar. *Journal of Finance,* 56, 387–396.

[56] Im, K.S., M.H. Pesaran, and Y. Shin (2003) Testing for Unit Roots in Heterogeneous Panels. *Journal of Econometrics*, 115, 53–74.

[57] Johansen, S. (1991) Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica,* 59, 1551–1580.

[58] Johansen, S. (1995) *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, Oxford: Oxford University Press.

[59] Kahneman, D. and A. Tversky (1981) The Framing of Decisions and the Psychology of Choice. *Science,* 211, 453–458.

[60] Kamstra, M., L. Kramer, and M. Levi (2003) Winter blues: a SAD stock market cycle. *American Economic Review,* 93, 324–343.

[61] Katona, G. (1974) Psychology and Consumer Economics. *Journal of Consumer Research*, 1, 1–8.

[62] Klibanoff, P., O. Lamont, and T.A. Wizman (1998) Investor Reaction to salient news in closed-end country funds. *Journal of Finance,* 53, 673–699.

[63] Kholodilin, K. A., M. Podstawski, and B. Siliverstovs (2010) Do Google Searches Help in Nowcasting Private Consumption? A Real-Time Evidence for the US. *KOF Working Papers No. 256*, April, Zürich.

[64] Kotsiantis, S. B., and P. E. Pintelas (2004) Increasing the Classification Accuracy of Simple Bayesian Classifier. *Lecture Notes in Artificial Intelligence,* 3192, 198–207.

[65] Levin, A., C.F. Lin, and C. Chu (2002) Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties. *Journal of Econometrics*, 108, 1–24.

[66] Lewis, D. (1998) Naive (Bayes) at forty: The independence assumption in information retrieval, in *ECML'98: Tenth European Conference On Machine Learning.*

[67] Livnat, J. and C. Petrovits (2009) Investor Sentiment, Post-Earnings Announcement Drift, and Accruals. *SSRN Working Paper* available at http://ssrn.com/abstract=1262757.

[68] Lloyd, B. T. Jr. (1999) Survey Measures of Expected U.S. Inflation. *The Journal of Economic Perspectives,* 13, 125–144.

[69] Ludwig, A. and T. Slok (2002) The Impact of Changes in Stock Prices and House Prices on Consumption in OECD Countries. *IMF Working Paper*, 02/1.

[70] Lütkepohl, H. (1991) *Introduction to Multiple Time Series Analysis.* Berlin: Springer Verlag.

[71] Lütkepohl, H. (1993) *Introduction to Multiple Time Series Analysis.* 2nd Edition, Berlin: Springer Verlag.

[72] Maier, S.R. (2005) Accuracy Matters: A Cross-Market Assessment of Newspaper Error and Credibility. *Journalism & Mass Communication Quarterly,* 82, 533–551.

[73] Maykut, P., and R. Morehouse (1994) *Beginning qualitative research: a philosophic and practical guide*, London.

[74] Meschke, F. and Y.H. Kim (2011) CEO Interviews on CNBC. *SSRN Working Paper.*

[75] Menkhoff, L. and R.R. Rebitzky (2008) Investor sentiment in the US-Dollar: Longer-term, non-linear orientation on PPP. *Journal of Empirical Finance*, 15, 455–467.

[76] Mullainathan, S. and A. Shleifer (2005) The Market for News. *American Economic Review,* 95, 1031–1053.

[77] Muth, J.F. (1961) Rational Expectations and the Theory of Price Movements. *Econometrica*, 29, 315–335.

[78] Nielsen (2010) How People Watch: A Global Nielsen Consumer Report, August 2010.

[79] Pesaran, M. H., and Y. Shin (1998) Impulse Response Analysis in Linear Multivariate Models. *Economics Letters*, 58, 17–29.

[80] Pew Report (2004) News Audiences Increasingly Politicized. *The Pew Research Center for the People & the Press.* Available at http://people-press.org/report/215/news-audiences-increasingly-politicized, last accessed 20 January 2011.

[81] Phillips, P.C.B., and P. Perron (1988) Testing for a Unit Root in Time Series Regression. *Biometrika*, 75, 335–346.

[82] Pindyck, R. S., and D. L. Rubinfeld (1998) *Econometric Methods*, 4th Edition, New York: McGraw-Hill.

[83] Rish, I. (2001) An empirical study of the naive Bayes classifier. *IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence*, available at http://www.cc.gatech.edu/home/isbell/classes/reading/papers/Rish.pdf.

[84] Rogers, C. (1951) *Client-centered therapy, its current practice, implications, and theory*. London.

[85] Schmidt, T., and S. Vosen (2011) Forecasting Private Consumption: Survey-based Indicators vs. Google Trends. *Journal of Forecasting*, 30 (6), 565–578.

[86] Sharpe, W.F. (1994) The Sharpe Ratio. *Journal of Portfolio Management,* 21, 49–58.

[87] Shiller, R. J. (1981) Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? *American Economic Review,* 71, 421–436.

[88] Sims, C. (2003) Implications of Rational Inattention. *Journal of Monetary Economics*, 50, 665–690.

[89] Sommer, M. (2007) Habit Formation and Aggregate Consumption Dynamics. *The B.E. Journal of Macroeconomics – Advances*, 7, Article 21.

[90] Stock, J., and M. Watson (1999) Forecasting Inflation. *Journal of Monetary Economics*, 44, 293–335.

[91] Stock, J., and M. Watson (2002) Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics*, 20, 147–162.

[92] Strömberg, D. (2004) Radio's impact on public spending. *The Quarterly Journal of Economics,* 119, 189–221.

[93] Summers, L.H. (1986) Does the Stock Market Rationally Reflect Fundamental Values? *Journal of Finance,* 41, 591–601.

[94] Taylor, S., and R. Bogdan (1998) *Introduction to qualitative research methods: a guidebook and resource*, New York.

[95] Tetlock, P.C. (2007) Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *Journal of Finance,* 62, 1139–1168.

[96] Tetlock, P.C. (2011) All the news that's fit to reprint: Do investors react to stale information? *Review of Financial Studies*, forthcoming.

[97] Tetlock, P.C., M. Saar-Tsechansky, and S. Macskassy (2008) More than words: Quantifying language to measure firm's fundamentals. *Journal of Finance,* 63, 1437–1467.

[98] Theil, H. (1958) *Economic Forecasts and Policy.* Amsterdam: North Holland.

[99] Uhl, M.W. (2011a) Explaining U.S. Consumer Behavior with News Sentiment. *ACM Trans. Manag. Inform. Syst.*, 2, 2, Article 9 (June 2011), 1–18.

[100] Uhl, M. W. (2011b) Reuters Sentiment and Stock Returns. *KOF Working Papers No. 288*, September, Zurich.

[101] Uhl, M. W. (2012) And Action: TV Sentiment and the US Consumer. *Applied Economics Letters*, 19:11, 1029–1034.

[102] Yuan, K. L. Zheng, and Q. Zhu (2006) Are investors moonstruck? Lunar phases and stock returns. *Journal of Empirical Finance,* 13, 1–23.

[103] White, H. (1980) A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica,* 48, 817–838.

Curriculum Vitae

Personal Details

Name	Adi Matthias William Uhl
Date of birth	04 April 1981
Place of birth	Ridgewood, NJ, U.S.A.

Education

- Doctoral Studies at the ETH Zurich, Switzerland (06/2009–12/2011)
- M.Sc. in Economic and Social History at the University of Oxford, U.K. (10/2006–09/2007)
- BA in History and Social Sciences & BA in International Business Administration at the American University of Paris, France (08/2003–05/2006)
- EAP Reciprocity Student at the University of California at Berkeley, CA, U.S.A. (08/2004–12/2004)

Professional Experience

- Researcher, International Business Cycles, KOF Swiss Economic Institute, ETH Zurich, Switzerland (05/2009–12/2011)
- Sales & Trading, Foreign Exchange & Money Markets, UBS Investment Bank, Singapore & Switzerland (01/2008–04/2009)
- Commercial Banker, Deutsche Bank AG, Germany (08/2001–06/2003)

KOF Dissertation Series

KOF

ETH Zürich
KOF Konjunkturforschungsstelle
WEH D 4
Weinbergstrasse 35
8092 Zürich

Telefon +41 44 632 42 39
Fax +41 44 632 12 18
www.kof.ethz.ch
kof@kof.ethz.ch