# Explaining U.S. Consumer Behavior with News Sentiment

MATTHIAS W. UHL, KOF Swiss Economic Institute, ETH Zurich

9

We introduce a novel dataset with a news sentiment index that was constructed from a selection of over 300,000 newspaper articles from five of the top ten U.S. newspapers by circulation. By constructing ARMA models, we show that news and consumer sentiment, when combined with other macroeconomic variables, achieve statistically significant results to explain changes in private consumption. We make three distinct findings with respect to sentiment in consumption behavior models: first, both consumer and news sentiment add explanatory power and statistical significance to conventional consumer behavior models. Second, consumer sentiment, measured by the University of Michigan Index of Consumer Sentiment, adds more explanatory power and statistical significance than news sentiment when tested individually. Third, news sentiment is able to determine the signs of all coefficients in the model correctly, whereas consumer sentiment does not. In general, we conclude that news sentiment is a useful variable to add in consumer behavior models, especially when coupled with consumer sentiment and other macroeconomic variables. Tested individually, news sentiment is as good a proxy as personal income for explaining private consumption growth when tested individually.

Categories and Subject Descriptors: I 2.7 [**Artificial Intelligence**]: Natural language processing—*Speech recognition and synthesis*

General Terms: Algorithms, Economics, Human Factors

Additional Key Words and Phrases: News sentiment, consumer sentiment, private consumption

**ACM Reference Format:**
Uhl, M. W. 2011. Explaining U.S. consumer behavior with news sentiment. ACM Trans. Manag. Inform. Syst. 2, 2, Article 9 (June 2011), 18 pages.
DOI = 10.1145/1985347.1985350 http://doi.acm.org/10.1145/1985347.1985350

## 1. INTRODUCTION

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. If economists know how private consumption develops, they have a good understanding of how the overall economy is behaving. Drawing on the studies of Campbell and Mankiw [1989], Doms and Morin [2004], Sommer [2007], and Carroll et al. [2010], among others, we take their idea of possible influence of sentiment on the consumer further by suggesting that sentiment in newspapers influences consumption behavior. For this study, we introduce a novel dataset with a news sentiment index that was constructed from a selection of over 300,000 newspaper articles from five of the top ten U.S. newspapers by circulation. By constructing ARMA models, we show that news and consumer sentiment, when combined with other variables such as personal income and savings, consumer prices, unemployment, interest rates, as well as stock and exchange rate prices, achieve statistically significant results to explain

Author address: M. W. Uhl, KOF Swiss Economic Institute, ETH Zurich, Zurich, Switzerland; email: uhl@kof.ethz.ch.



DOI 10.1145/1985347.1985350 http://doi.acm.org/10.1145/1985347.1985350

changes in private consumption. We make three distinct findings with respect to sentiment in consumption behavior models: first, both consumer and news sentiment add explanatory power and statistical significance to conventional consumer behavior models. Second, consumer sentiment, measured by the University of Michigan Index of Consumer Sentiment (ICS), adds more explanatory power and statistical significance than news sentiment when tested individually. Third, news sentiment is able to determine the signs of all coefficients in the model correctly, whereas consumer sentiment does not. In general, we conclude that news sentiment is a useful variable to add in consumer behavior models, especially when coupled with consumer sentiment and other macroeconomic variables. Further, news sentiment performs as well as personal income in explaining private consumption growth when tested individually.

Section 1 discusses the motivation of the article and the related literature as well as the novel dataset. Section 2 sets out the model and discusses the empirical results, while Section 3 concludes.

### 1.1. 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

$$c_t = \alpha(A_t + H_t), \tag{1}$$

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

$$H_t = \sum_{k=0}^{\infty} (1+\mu)^{-k} \ {}_t y_{t+k}, \tag{2}$$

where $\mu$ 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 incorporates the idea that ${}_t y_{t+k} = E(y_{t+k}\,|I_t)$, 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.

For example, 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

$$\Delta c_t^L = \varepsilon_t, \tag{3}$$

$$\Delta c_t^R = \Delta y_t^R, \tag{4}$$

where $c_t^L$ refers to life-cycle consumers, $c_t^R$ to rule-of-thumb consumers, $\varepsilon_t$ to news received in period $t$ about lifetime resources, and $y_t^R$ to current income of private households. A crucial assumption in the Campbell-Mankiw framework is that rule-of-thumb consumers receive a constant proportion $\lambda$ of total income. Aggregate consumption is then given as follows by the combination of Eqs. (3) and (4).

$$\Delta c_t = \lambda \Delta y_t + \varepsilon_t \tag{5}$$

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 form

$$\Delta \log c_t = \alpha_0 + \sum_{i=1}^{N} \beta_i S_{t-i} + \gamma Z_{t-1} + v_t, \tag{6}$$

where $\Delta \log c_t$ refers to differenced logs of private consumption, $S_t$ to consumer sentiment and expectations (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 U.K. 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 households' expectations, and ultimately the 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 start our analysis. In general, we assume in this study 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 U.S. Doms and Morin [2004] examine the hypothesis that news media affects 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, 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 a basket of 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, thus avoiding subjectivity bias. Further, Carroll [2009] examined the budget constraint, which should require consumption to adjust fully to any permanent shock to income. He shows that an increase in permanent income reduces the ratio of assets to permanent income, which results in a temporary increase of precautionary saving. Carroll et al. [2010] 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 variables other than 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. Thus, in line with the studies mentioned, we test whether these other variables influencing private consumption can not only be news sentiment, but also consumer sentiment and expectations, measured by the University of Michigan Index of Consumer Sentiment (ICS), personal income and savings, inflation, the unemployment rate, short-term interest rates, a general stock index such as the S&P 500, and the EURUSD exchange rate, in order to account for changes in wealth effects. We thus extend existing models with news sentiment and test whether this novel variable can add value to models that explain U.S. private consumption.

### 1.2. Dataset

In this study, we introduce a new variable that quantifies news sentiment from the economics section of various newspapers in the U.S. from 1995 to 2009 on a quarterly basis. A sentiment algorithm is used for the analysis of over 300,000 newspaper articles from the Washington Post (WP), USA Today (UT), the Houston Chronicle (HC), the New York Times (NYT), and the Wall Street Journal (WSJ). Table I shows the average daily circulation of each newspaper and how many newspaper articles were examined for sentiment of each newspaper.[1] A news sentiment index was then created from the two newspapers that performed best in the models and that were the most comprehensive graphically: the WP and 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.[2] Visual Basic programs were written in order to ease the process of dealing with the large amount of data.[3] The sentiment algorithm scans each article (headline plus full body) and gives an output file with the respective

[1]The examined newspapers were selected from the top ten list of daily average circulation according to availability in the *LexisNexis* database.
[2]See Appendix A.1 for more information on the sentiment classifier.
[3]See Appendix A.2 for more information on these programs.

Table I. Newspaper Statistics

| | Number of articles examined for sentiment (1995–2009) | Average Daily Circulation* |
|---|---|---|
| *USA Today* | 28,832 | 1,826,622 |
| *Washington Post* | 74,206 | 604,650 |
| *Houston Chronicle* | 30,919 | 494,131 |
| *New York Times* | 114,454 | 951,063 |
| *Wall Street Journal—Abstracts* | 74,420 | 2,092,523 |
| Total | 322,831 | |

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

sentiment rating of each individual article. The article ratings were then aggregated on a quarterly basis.

Quarterly U.S. private consumption data as well as consumer price index data were obtained from the U.S. Department of Commerce Bureau of Economic Analysis database.[4] Unemployment rate figures, short-term interest rates (3-month USD LIBOR), S&P 500 Stock Index, as well as EURUSD exchange rate data were obtained from Thomson Reuters Datastream. Personal income data were downloaded from the Bureau of Economic Analysis Web site.[5] Personal savings data were obtained from the ALFRED database.[6] The University of Michigan Index of Consumer Sentiment (ICS) data were downloaded from the University of Michigan and Thomson Reuters public access Web site.[7] The ICS is constructed from answers to five questions relating to current economic conditions of consumers as well as consumer expectations.[8]

## 2. EMPIRICAL ANALYSIS

Doms and Morin [2004] draw a chart that sets news coverage, households' sentiment and expectations, as well as private consumption in relation to each other. In Figure 1, we have amended this information flow chart and added news sentiment, postulating that the reader, that is, the consumer, is influenced by the news and sentiment portrayed through news that she reads which forms her expectations and sentiment, ultimately driving changes in future consumption.

In our analysis, we want to examine two channels of influence, depicted by the numbers 1 and 2 in Figure 1. In channel 1, the influence channel is as laid out in Doms and Morin [2004]. Consumers receive news about the economic activity which shapes their expectations and sentiment, ultimately driving future consumption behavior. The second channel of influence that we consider focuses on the influence of sentiment in the news that gets to the consumer, ultimately driving consumption behavior. In the succeeding models, we want to test possible channels of influence in various ways by combining news sentiment and consumer expectations and sentiment as well as considering these individually in our model. Although these two independent variables are central to this study, we account for other possible factors than can influence consumer behavior in our model as well, such as income and savings, inflation, unemployment, and financial market variables to proxy for changes in personal wealth.

[4]http://www.bea.gov/ (accessed 4/10).
[5]http://www.bea.gov/ (accessed 6/10).
[6]ArchivaL Federal Reserve Economic Data. http://alfred.stlouisfed.org/ (accessed 9/10).
[7]http://www.sca.isr.umich.edu/ (accessed 6/10).
[8]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. Index Calculations, http://www.sca.isr.umich.edu/documents.php?c=i (accessed 6/10).

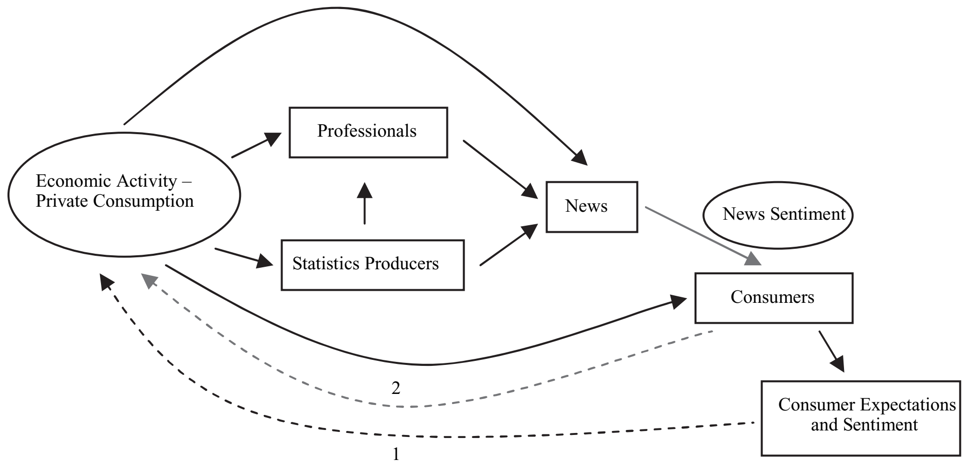


Fig. 1. Information Flows to consumers.

We select the models according to the Box and Jenkins [1970] model selection approaches. We first test the variables for unit roots with Augmented Dickey-Fuller (ADF) tests according to Dickey and Fuller [1979], before analyzing them graphically.[9] We find that all variables have a unit root on the level, except news sentiment. To exclude the possibility of spurious regression results and stationarity as Granger and Newbold [1974] noted, we use log differenced values for the dependent variable private consumption and log differenced values for all other independent variables except for the ICS and news sentiment. For the news sentiment index and the ICS we use level data, as Sommer [2007] finds that the level of sentiment in consumption behavior models is more relevant than growth rates, since the correlation is then much higher between these variables.

In Figure 2, we show private consumption growth plotted against the ICS. A comovement 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 comovement with private consumption growth rates in Figure 3. As the ICS, the news sentiment index experienced its height in 2000 and the subsequent crash when the dot.com bubble burst. However, the down periods in news sentiment are more pronounced than in the ICS, and in this light, the series appears 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 sells.”

[9]See Appendix A.3 for the exact formulation of the ADF test.

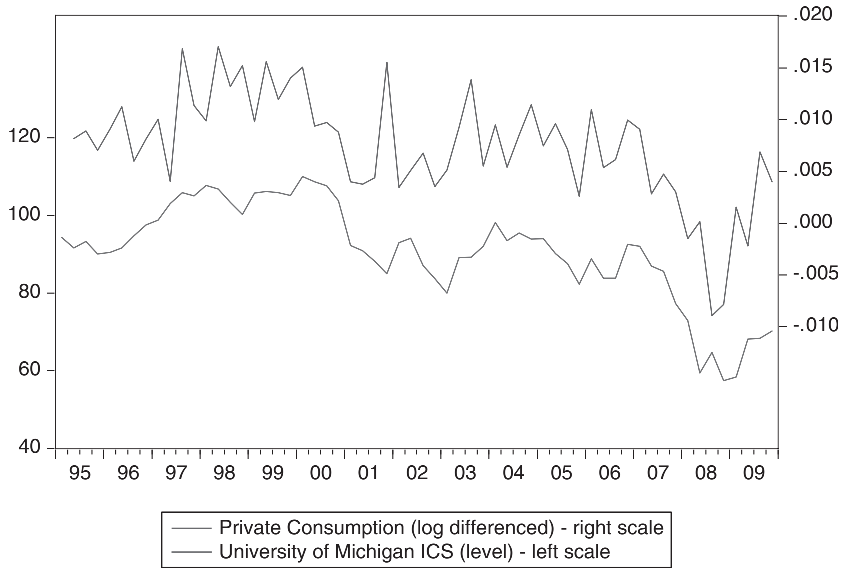


Fig. 2. Time-series chart of private consumption and University of Michigan ICS.

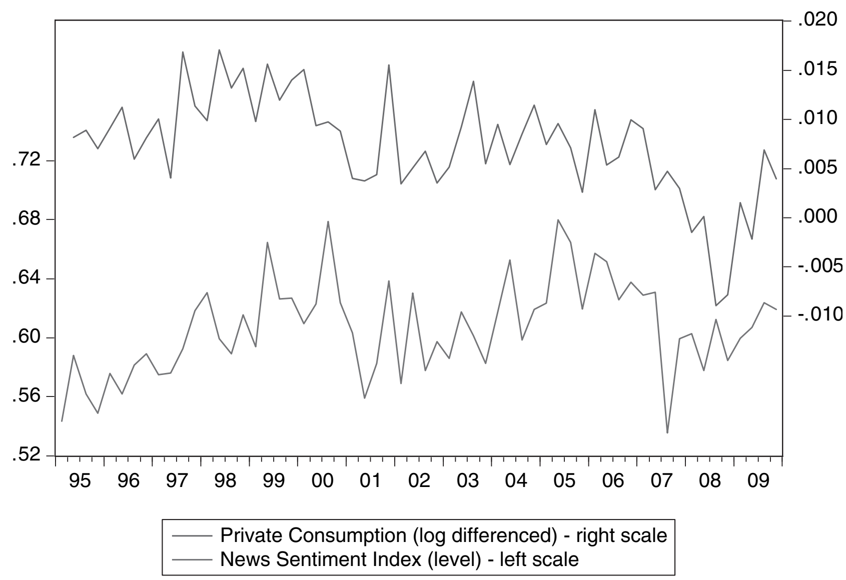


Fig. 3. Time-series chart of private consumption and news sentiment index.

In Figure 4, we plot personal income and savings growth as well as inflation against private consumption growth. Personal income, inflation, and consumption comove 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 three variables turned negative.

Conceptionally, we would expect that if prices fall consumption would go up. It is difficult to make this out in the graph, so that we need to consider this later in our regression model. Personal savings growth rates paint a slightly different picture.

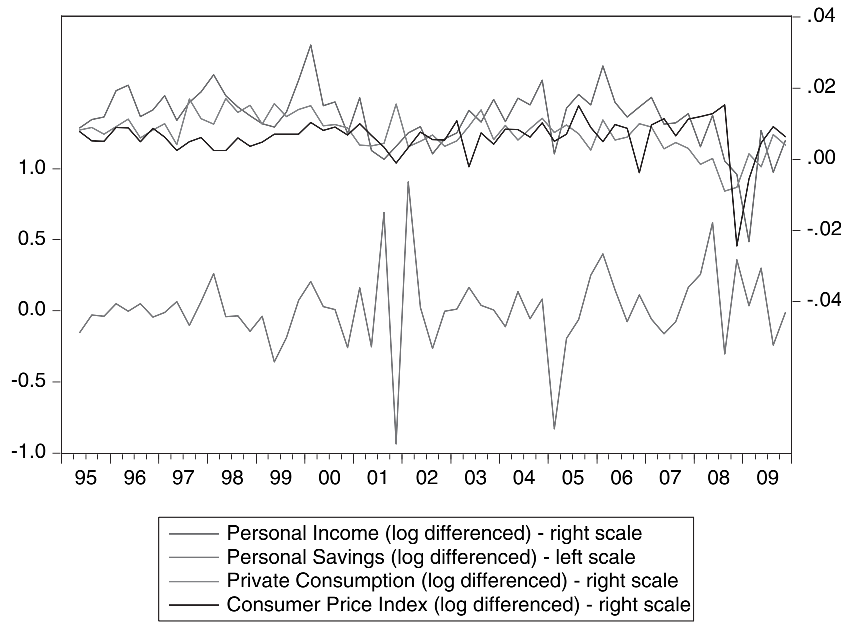


Fig. 4. Time-series chart of private consumption, personal income and savings, and consumer price index.

Growth rates of personal savings experienced the highest volatility during the dot.com crisis. In general, we would expect to have a negative correlation between savings and consumption because when people save less they can spend more. This hypothesis is well pronounced during the economic upswing of 2004/05 when savings rates declined and during the financial crisis of 2008/09 when savings rates increased.

Figure 5 shows private consumption growth plotted against the EURUSD exchange rate and S&P 500 stock index returns. The stock index shows a comovement with private consumption, while the drop and subsequent recovery of the stock index is most pronounced during the financial crisis. The EURUSD exchange rate, however, does not clearly show similarities to private consumption graphically. The unemployment rate is also a good proxy of the health of an economy. If the unemployment 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. Figure 6 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. Figure 7 makes the same case. Short-term interest rates showed a much higher volatility during the Lehman crisis than during the dot.com crisis. 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 base 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 II shows the

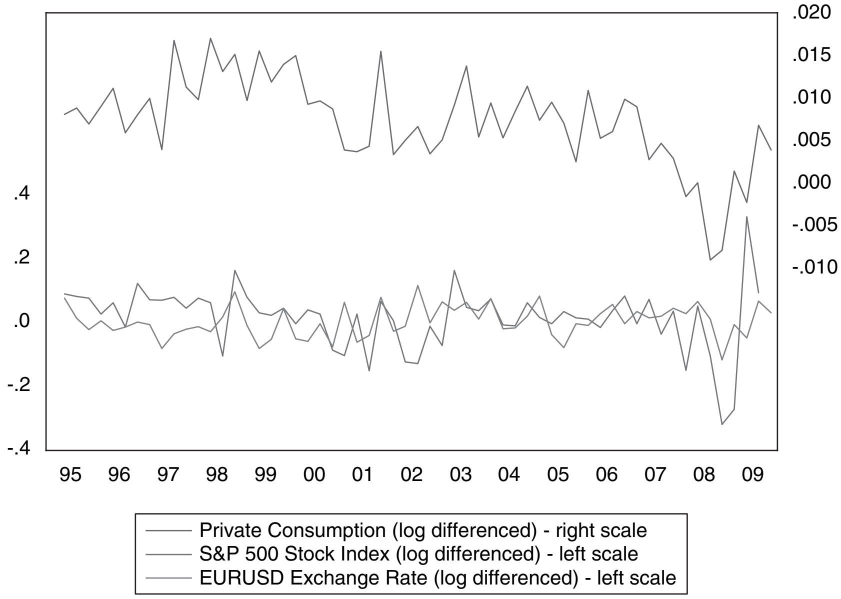


Fig. 5. Time-series chart of private consumption, S&P 500 stock index, and EURUSD exchange rate.

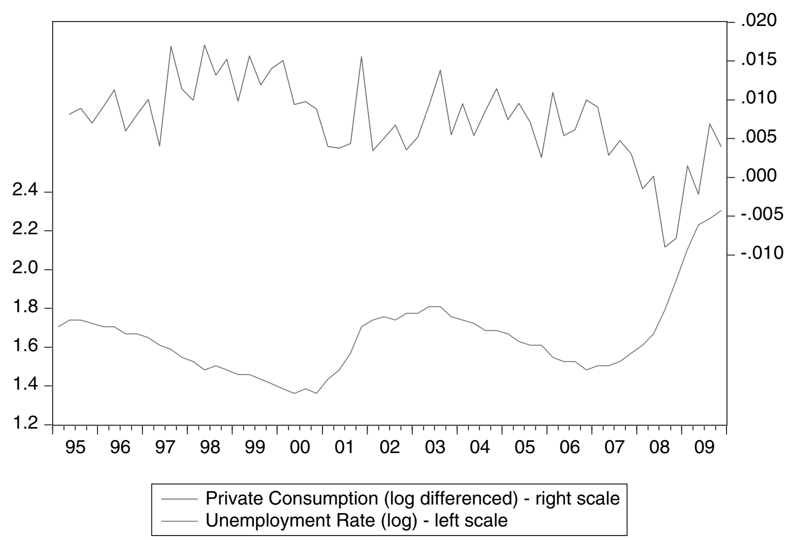


Fig. 6. Time-series chart of private consumption and unemployment.

results of various ARMA(x,y) models with private consumption as dependent variable and all other available variables as independent variables.[10]

Given the high 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. Taking this restriction into account, we can only choose between an ARMA(2,1) and an

[10]We tested the various ARMA(x,y) models with x = {1...4} and y = {1...4}.

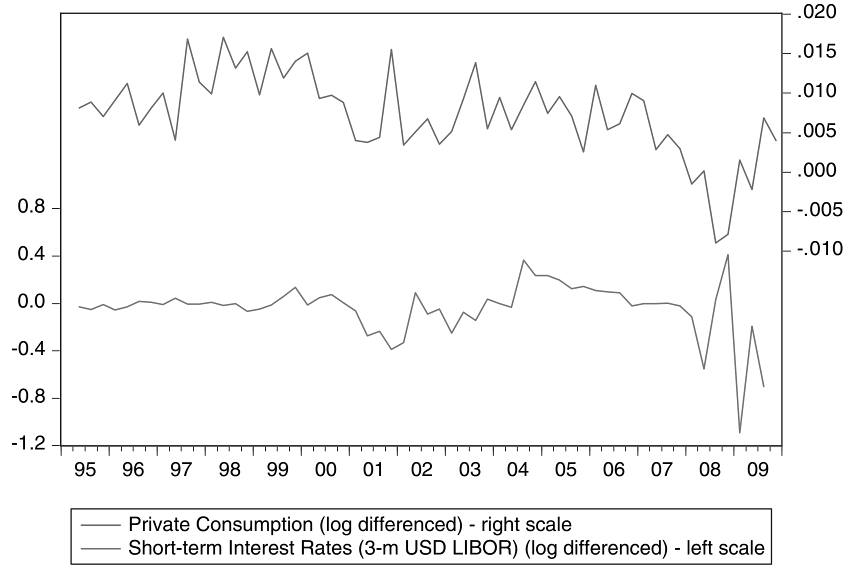


Fig. 7. Time-series chart of private consumption and short-term interest rates.

Table II.

| ARMA-Lag Structure | Schwarz Info Criterion (BIC) | Inverted MA Roots* |
|---|---|---|
| ARMA(1,1) | −8.695444 | 1 |
| ARMA(2,1) | −8.425664 | <1 |
| ARMA(1,2) | −8.452986 | <1 |
| ARMA(2,2) | −8.538988 | 1 |
| ARMA(3,1) | −8.657786 | >1 |
| ARMA(3,2) | −8.587306 | >1 |
| ARMA(3,3) | −8.495185 | >1 |
| ARMA(4,1) | −8.408595 | 1 |
| ARMA(4,2) | −8.531950 | >1 |
| ARMA(4,3) | −8.435662 | >1 |
| ARMA(4,4) | −8.320768 | >1 |
| ARMA(2,3) | −8.416930 | >1 |
| ARMA(1,3) | −8.545574 | 1 |
| ARMA(1,4) | −8.481509 | 1 |
| ARMA(2.4) | −8.386913 | >1 |
| ARMA(3,4) | −8.420199 | >1 |

This table shows various Schwarz information criteria tests in order to determine the best ARMA structure of the base model with private consumption as dependent variable and news sentiment index, UMICH ICS, personal income, personal savings, consumer price index, unemployment rate, short-term interest rates, S&P 500 stock index and EURUSD exchange rate as independent variables

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

ARMA(1,2) structure, as these two model structures are the only ones to have inverted MA roots that are less than 1. Given that the ARMA(1,2) model has the lower BIC, we take this model as our base model. This is in line with Sommer's [2007] findings, as he applies an ARMA(1,2) structure to modeling private consumption as well. On a similar note, Working [1960] finds that an MA(1) process is necessary because preference choices generate time aggregation. Further, Carroll et al. [2010] find that in a single habit formation or sticky information model, time aggregation generates an MA(2) process in consumption growth. The AR(1) process is important, according to Carroll et al. [2010], because of the stickiness in consumption growth. We have thus accounted for possible measurement errors, as mentioned by Carroll et al. [2010] but also by Sommer [2007], to the best of our knowledge, so that we formulate the model with an ARMA(1,2) structure as

$$\Delta \log c_t = k + \alpha_1 \Delta \log c_{t-1} + \beta \Delta \log Z_t + \gamma S_t + \sum_{i=1}^{2} \theta_i \varepsilon_{t-i} + \varepsilon_t, \tag{7}$$

where $\Delta \log c_t$ refers to differenced logs of private consumption, $k$ is the constant term, $S_t$ refers to levels of news sentiment and/or the ICS according to Sommer [2007], $\Delta \log Z_t$ to differenced logs values of all other independent variables, such as personal income and savings, consumer price index, unemployment, short-term interest rates, the S&P500 stock index as well as the EURUSD exchange rate, and $\varepsilon_t$ represents the error term. 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 and without $S_t$, so that we can specifically identify the value of the ICS and the news sentiment index jointly and individually.

### 2.1. Discussion of Results

We examine private consumption growth with our formulated model and apply the Theil inequality coefficient for comparison, as according to Theil [1958], as well as to the Root Mean Squared Error (RMSE).[11] In all regressions, we derive heteroskedasticity consistent covariance matrices according to White [1980].[12] We further 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 serial correlation in any of the regressions.

As outlined in Eq. (7), our model has an ARMA(1,2) structure. The results of all regressions run are presented in Table III. Regression (1) includes all variables available, as we follow Hendry's [1993] approach. In this regression, almost all independent variables are statistically significant, except unemployment rate and the S&P 500 stock index. All variables show the correct, or expected, coefficient sign except the unemployment rate, which has a positive coefficient sign. We would have expected that the sign is negative, as we assume that with less unemployment there is more private consumption. The news sentiment index, the ICS, personal income, and the S&P 500 have positive coefficient signs, which are as expected. Private consumption increases when news and consumer sentiment increases, as consumers are more willing to consume when they are in a good mood. When the income of a household rises, its consumption grows naturally. When stock markets rise, especially in the U.S., households tend to consume more as they experience rising stock prices with an increase in their wealth, so that they can consume more. When consumer prices fall, one would expect consumers to consume more, justifying the negative coefficient sign for the Consumer Price Index (CPI). With the EURUSD exchange rate, the case is not as clear cut. A rise in the exchange rate pair signals a weakening U.S. dollar. Thus, a positive coefficient sign for the EURUSD exchange rate means that when the domestic currency weakens (i.e., the U.S. dollar), people tend to consume more. The adjusted R-squared of regression (1) is quite high with 0.78. In regressions (2) and (3), we have left out the ICS and the news sentiment index, respectively, to determine the effect of each sentiment variable on the model. It becomes quite apparent that the ICS adds more explanatory power to the model than the news sentiment index. Although not statistically significant, the

[11]See Appendix A.4 for the calculation of the Theil inequality coefficient and the RMSE.
[12]See Appendix A.5 for the exact formulation of the White covariance matrix.

Table III. *Dependent Variable:* Private Consumption (log differenced)

| *Independent Variables* | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| News Sentiment Index (level) | 0.020243* | 0.025059* | | | 0.017516* | 0.022945* | | |
| | (0.010394) | (0.014492) | | | (0.009398) | (0.012918) | | |
| University of Michigan Index of Consumer Sentiment (level) | 0.000275*** | | 0.00029*** | | 0.000259*** | | 0.000260*** | |
| | (0.000029) | | (0.0000363) | | (0.0000259) | | (0.0000323) | |
| Personal Income (log differenced) | 0.333273*** | 0.402467*** | 0.327918*** | 0.432349*** | 0.338315*** | 0.426542*** | 0.332939*** | 0.441900*** |
| | (0.090222) | (0.080346) | (0.091676) | (0.076108) | (0.059511) | (0.075912) | (0.069773) | (0.069031) |
| Personal Savings (log differenced) | −0.005926*** | −0.007340*** | −0.006205*** | −0.008083*** | −0.006103*** | −0.007267*** | −0.006429*** | −0.007812*** |
| | (0.001502) | (0.002181) | (0.001405) | (0.002181) | (0.001307) | (0.002097) | (0.001307) | (0.002094) |
| Consumer Price Index (log differenced) | −0.249226*** | −0.296969*** | −0.207102*** | −0.308459*** | −0.265031*** | −0.263430*** | −0.240178*** | −0.263204*** |
| | (0.066178) | (0.093675) | (0.073135) | (0.088706) | (0.061698) | (0.086789) | (0.071575) | (0.078832) |
| Unemployment Rate (log) | 0.002061 | −0.003496 | 0.002372 | −0.003462 | | | | |
| | (0.002116) | (0.004280) | (0.002231) | (0.006128) | | | | |
| Short-Term Interest Rates (3-month USD LIBOR) (log differenced) | −0.009144*** | −0.010388*** | −0.009500*** | −0.010760*** | −0.009818*** | −0.010896*** | −0.009673*** | −0.011218*** |
| | (0.001699) | (0.002854) | (0.001875) | (0.002937) | (0.001406) | (0.002629) | (0.001662) | (0.002536) |
| S&P 500 Stock Index (log differenced) | 0.000914 | 0.005807 | −0.002164 | 0.007023 | | | | |
| | (0.004273) | (0.004515) | (0.003584) | (0.004527) | | | | |
| EURUSD exchange rate (log differenced) | 0.021243** | 0.010995 | 0.020728** | 0.00753 | 0.020632** | 0.010173 | 0.017135** | 0.007498 |
| | (0.008030) | (0.008449) | (0.008118) | (0.007614) | (0.007927) | (0.007822) | (0.008336) | (0.008322) |
| Constant | −0.035941*** | −0.009074 | −0.025692*** | 0.008807 | −0.029450*** | −0.013754 | −0.019099*** | 0.002233 |
| | (0.009602) | (0.022394) | (0.005680) | (0.011489) | (0.006459) | (0.012316) | (0.002546) | (0.003402) |
| AR(1) | −0.495860 | 0.972162*** | −0.833839*** | 0.906758*** | −0.610716 | 0.965979*** | −0.897728*** | 0.929465*** |
| | (0.647202) | (0.124090) | (0.123994) | (0.173920) | (0.608043) | (0.098362) | (0.103371) | (0.105047) |
| MA(1) | 0.276916 | −1.004680*** | 0.805767*** | −0.991135*** | 0.449735 | −0.859107*** | 0.840121*** | −0.804859*** |
| | (0.657225) | (0.251748) | (0.193306) | (0.255731) | (0.617229) | (0.213260) | (0.181046) | (0.203681) |
| MA(2) | −0.269725 | 0.293325 | −0.164578 | 0.408691** | −0.210402 | 0.223713 | −0.159874 | 0.252208 |
| | (0.222134) | (0.190523) | (0.190398) | (0.158843) | (0.181256) | (0.192118) | (0.181047) | (0.173721) |
| R-squared | 0.825857 | 0.730782 | 0.816862 | 0.720401 | 0.821914 | 0.722671 | 0.815889 | 0.711524 |
| Adjusted R-squared | 0.778363 | 0.664973 | 0.772095 | 0.659619 | 0.783200 | 0.669565 | 0.780634 | 0.663445 |
| N (after adjustments) | 57 | 57 | 57 | 57 | 57 | 57 | 57 | 57 |
| Schwarz criterion | −8.452986 | −8.088272 | −8.473557 | −8.121369 | −8.572458 | −8.200450 | −8.610119 | −8.231975 |
| Root Mean Squared Error (RMSE) | 17.60577 | 22.43139 | 18.16637 | 22.93285 | 17.83218 | 22.73074 | 18.23207 | 23.17145 |
| Theil Inequality Coefficient | 0.001095 | 0.001395 | 0.00113 | 0.001426 | 0.001109 | 0.001413 | 0.001134 | 0.001441 |

Table III. *Continued*

| | (9) | (10) | (11) | (12) | (13) | (14) | (15) | (16) |
|---|---|---|---|---|---|---|---|---|
| News Sentiment Index (level) | | | | | 0.027444** | 0.032305* | | |
| | | | | | (0.011974) | (0.018723) | | |
| University of Michigan Index of Consumer Sentiment (level) | | | | | 0.000313*** | | 0.000311*** | |
| | | | | | (0.0000245) | | (0.0000457) | |
| Personal Income (log differenced) | 0.449762*** | 0.296702*** | 0.247588** | 0.133287* | | | | |
| | (0.072144) | (0.102941) | (0.102301) | (0.076513) | | | | |
| Personal Savings (log differenced) | −0.008203*** | −0.006392*** | −0.005809** | | | | | |
| | (0.001952) | (0.002043) | (0.002203) | | | | | |
| Consumer Price Index (log differenced) | −0.250422*** | −0.113268 | | | | | | |
| | (0.078985) | (0.157751) | | | | | | |
| Unemployment Rate (log) | | | | | | | | |
| Short-Term Interest Rates (3-month USD LIBOR) (log differenced) | −0.011716*** | | | | | | | |
| | (0.002526) | | | | | | | |
| S&P 500 Stock Index (log differenced) | | | | | | | | |
| EURUSD exchange rate (log differenced) | | | | | | | | |
| Constant | 0.002095 | 0.004511* | 0.00428* | 0.005694** | −0.037941*** | −0.012673 | −0.020812*** | 0.007236*** |
| | (0.003342) | (0.002352) | (0.002351) | (0.002172) | (0.008377) | (0.010952) | (0.004221) | (0.002076) |
| AR(1) | 0.924658*** | 0.866438*** | 0.879640*** | 0.797431*** | 0.693802*** | 0.837459*** | 0.097748 | 0.792840*** |
| | (0.109641) | (0.170278) | (0.138179) | (0.149098) | (0.177278) | (0.110104) | (2.070965) | (0.135577) |
| MA(1) | −0.788219*** | −0.727921*** | −0.691983*** | −0.587716*** | −0.937085*** | −0.59034*** | −0.218277 | −0.510896*** |
| | (0.198986) | (0.174748) | (0.163513) | (0.164330) | (0.218829) | (0.141943) | (2.022967) | (0.151669) |
| MA(2) | 0.237741 | 0.127817 | 0.090309 | 0.245821 | −0.034714 | 0.206638 | 0.078435 | 0.277411** |
| | (0.177152) | (0.170614) | (0.183632) | (0.152013) | (0.213227) | (0.142503) | (0.280370) | (0.133844) |
| R-squared | 0.707264 | 0.554463 | 0.541382 | 0.45684 | 0.648000 | 0.456235 | 0.586521 | 0.428464 |
| Adjusted R-squared | 0.665445 | 0.502047 | 0.497284 | 0.415847 | 0.614153 | 0.415197 | 0.555315 | 0.396712 |
| N (after adjustments) | 57 | 58 | 58 | 58 | 58 | 58 | 58 | 58 |
| Schwarz criterion | −8.288247 | −7.955795 | −7.996864 | −7.897687 | −8.261451 | −7.896574 | −8.170484 | −7.916771 |
| Root Mean Squared Error (RMSE) | 23.39116 | 30.00512 | 30.35707 | 32.44914 | 25.45901 | 32.36906 | 27.80289 | 33.10148 |
| Theil Inequality Coefficient | 0.001454 | 0.00186 | 0.001882 | 0.002011 | 0.001578 | 0.002007 | 0.001724 | 0.002052 |

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

coefficient sign of unemployment in regression (2) is conceptionally correct, whereas in regression (3) it is not. For the S&P 500 stock index coefficient sign, the same applies. All other coefficient signs of the explanatory variables are correct and identical in the two models. This could mean that news sentiment adds more “quality” to the model than the ICS, even though this assumption is neither manifested in the statistical significance nor in the explanatory power. It is nevertheless an interesting aspect that should be kept in mind. For comparison reasons, we have run regression (4) without the news and consumer sentiment variables. Notably, all coefficient signs are correct, as we would have anticipated. However, the unemployment rate, stock index, and foreign exchange rate coefficients are not statistically significant. The adjusted R-squared value of regression (4) is less than in regressions (1) to (3), indicating that news and consumer sentiment add value to the consumption model. It also becomes apparent that the ICS adds much more explanatory power than news sentiment, although the “quality” of the results is better for the models that include the news sentiment index because the coefficient signs are correct for all variables when news sentiment is included. In regressions (5) to (8), we have dropped the insignificant variables unemployment rate and S&P 500 stock index from our base regression (1), while alternating between the news sentiment index and the ICS to highlight the differences between these sentiment variables again. In all four regressions, the explanatory power increases slightly after dropping the unemployment rate and stock market variables. We obtain the same pattern as in regressions (1) to (4) with regression (5) having the highest adjusted R-squared, as it incorporates both sentiment variables. Regression (7), which includes the ICS only, shows a much higher adjusted R-squared (0.78) than regression (6) with the news sentiment index (0.66). Regression (8) without both sentiment variables has the lowest explanatory power.

So far, we can summarize three findings: first, both the consumer and news sentiment variables, jointly and individually, add explanatory power to consumption behavior models. Second, consumer sentiment (i.e., the ICS) adds more explanatory power to consumption models than news sentiment. And third, although consumer sentiment adds more explanatory power to the models, news sentiment appears to add more “quality” to general consumer behavior models, manifested in correct coefficient signs of all independent variables of regression (1) as opposed to regression (3).

In regressions (9) to (12), we reduce regression (8) further to be able to compare those models with sentiment variables to more “traditional” consumption behavior models, as, for example, in Hayashi [1982] or in Campbell and Mankiw [1989], which do not include consumer or news sentiment. Regression (9) incorporates personal income and savings, inflation, and short-term interest rates, regression (10) consists only of personal income and savings as well as inflation, regression (11) of personal income and savings, and (12) of personal income. Almost all coefficients (except consumer prices in (10)) are statistically significant in the regressions run, speaking for strong models with an adjusted R-squared ranging from 0.66 to 0.41. The coefficient signs are all correctly specified in these regressions. In regressions (13) to (15), we consider solely the sentiment variables, jointly and individually, for explaining private consumption growth. Again, we see that the ICS has the higher explanatory power than news sentiment individually in regressions (14) and (15), respectively. When combined, both sentiment variables achieve much higher explanatory power as well as higher statistical significance of the news sentiment coefficient.Comparing regressions (9) to (12) with regressions (13) to (15), we can claim that both consumer and news sentiment achieve similar explanatory power and statistical significance in consumer behavior models than traditional macroeconomic variables. For example, news sentiment is as good a proxy for consumption growth as personal income. Regression (16) is the most basic form of the ARMA(1,2) model without any independent variable. The adjusted

R-squared is the lowest among all regressions run, indicating that all independent variables specified in the model are of some value.

Overall, we confirm our previous three findings after having considered the results of all regressions run. First, explanatory power and statistical significance are added when including both consumer and news sentiment in the model. Second, it appears that consumer sentiment, measured by the ICS, performs markedly better than the news sentiment index in consumption behavior models. Last, news sentiment adds more "quality" to our general consumer behavior model, as only when news sentiment is included in the model are all of the regression coefficient signs correct. In general, we can conclude that news sentiment is a useful variable to add in consumer behavior models, especially when coupled with consumer sentiment and other macroeconomic variables. News sentiment also performs as well as personal income in explaining private consumption growth.

## 3. CONCLUSION

Although media coverage has become extremely important in the past decades, the effects on private consumers of sentiment published in newspaper articles have been barely explored in the literature. We introduce a novel dataset and procedure by creating a news sentiment index that accounts for positive and negative sentiment from over 300,000 newspaper articles of the economics section of some of the most widely-read newspapers in the U.S. from 1995 to 2009. This novel dataset was constructed by utilizing a new text mining approach.

We examine empirically the connection and impact of news sentiment on consumer behavior. In accordance with previous findings of Campbell and Mankiw [1989], Carroll et al. [1994, 2010] and Sommer [2007] among others, we formulate consumption behavior models that incorporate consumer and news sentiment among other macroeconomic variables, such as personal income and savings, inflation, unemployment, interest rates, stock index, and the exchange rate. Following Doms and Morin [2004], we extend their information flow chart (refer to Figure 1) with news sentiment. We draw up two possible channels how private consumption can be influenced and driven. In the first channel, we show that the consumer can be influenced in her consumption behavior by expectations and sentiment, and, in the second channel by news sentiment. Additionally, consumption behavior can be proxied by other macroeconomic variables, such as a household's income and savings, among others. In line with the findings of Sommer [2007] and Carroll et al. [2010], 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, in personal income, the EURUSD exchange rate, and a decrease in personal savings, inflation, and interest rates lead to an increase in private consumption. The unemployment rate and the S&P 500 stock index do not have a statistically significant impact on consumer behavior in our model. With regard to the sentiment variables, we find that consumer sentiment (i.e., the University of Michigan Index of Consumer Sentiment) has more explanatory power than news sentiment in our model. However, news sentiment is a valid variable to add in consumption behavior models, as it has explanatory power and it adds more "quality" to these models, which is manifested in the fact that the coefficient signs of all variables in the model are specified correctly, whereas this is not the fact with consumer sentiment. Tested individually, news sentiment is as good a proxy as personal income in private consumption growth models.

This first long-term analysis of news sentiment 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 its consumption behavior.

## APPENDIXES

### A.1

The sentiment algorithm is based on one of the most popular classifiers used in machine learning science: the naive Bayes classifier.[13] The sentiment algorithm Java program was obtained from a free Web-based sentiment algorithm provider and is tested for accuracy in this study. 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 predefined database of positive and negative words and phrases. The classification is then done by applying the 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 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 $Pr(a|b, c) = Pr(a|c)$ for all possible values of $a$, $b$, and $c$, whenever $Pr(c) > 0$. 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 the 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 a 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. 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.

### A.2

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.

### A.3

In this article, the augmented Dickey-Fuller unit root test is utilized because the model follows a higher-order AR process. Take the following equation

$$\Delta y_t = \alpha_0 + \gamma y_{t-1} + \sum_{i=2}^{p} \beta_i \Delta y_{t-i+1} + \varepsilon_t, \tag{8}$$

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

where $\gamma = -(1 - \sum_{i-1}^{p} a_i)$ and $\beta_i = -\sum_{j-1}^{p} a_j$. The null hypothesis tests whether $\gamma = 0$, and if so, the equation is entirely in first differences and so has a unit root. If $\gamma \neq 0$, then the equation does not have a unit root.

### A.4

The Theil inequality coefficient is calculated as in Theil [1958]. We have

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

where the forecast sample is $j = T+1, T+2, \ldots, T+h$, and the actual and forecast value in period $t$ is $y_t$ and $\hat{y}_t$, respectively.

The Root Mean Squared Error (RMSE) is calculated as follows.

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

### A.5

The White covariance matrix as in White [1980] is given by

$$\hat{\sum}_W = \frac{T}{T-k}(X'X)^{-1}\left(\sum_{t=1}^{T} u_t^2 x_t x_t'\right)(X'X)^{-1},$$

where $T$ is the number of observations, $k$ is the number of regressors, $X$ is the variable matrix, and $u_t$ is the least squares residual.

## REFERENCES

ACEMOGLU, D. AND SCOTT, A. 1994. Consumer confidence and rational expectations: Are agents' beliefs consistent with the theory? *Econ. J. 104*, 1–19.

ANG, A., BEKAERT, G., AND WEI, M. 2007. Do macro variables, asset markets, or surveys forecast inflation better? *J. Monet. Econ. 54*, 1163–1212.

BARON, D. 2006. Persistent media bias. *J. Public Econ. 90*, 1–36.

BOX, G. AND JENKINS, G. 1970. *Time Series Analysis Forecasting and Control*. Holden-Day, San Francisco.

BREEDEN, D. T. 1986. Consumption, production, inflation and interest rates. *J. Finan. Econ. 16*, 3–39.

BREUSCH, T. S. AND PAGAN, A. R. 1979. A simple test for heteroskedasticity and random coefficient variation. *Econometrica 48*, 1287–1294.

CAMPBELL, J. AND MANKIW, G. 1989. Consumption, income, and interest rates: Reinterpreting the time series evidence. In *NBER Macroeconomics Annual.* O. Blanchard and S. Fischer (Eds.). MIT Press, Cambridge, MA.

CARROLL, C. 2003. Macroeconomic expectations of households and professional forecasters. *Quart. J. Econ. 118*, 269–298.

CARROLL, C. 2009. Precautionary saving and the marginal propensity to consume out of permanent income. *J. Monet. Econ. 56*, 780–790.

CARROLL, C., FUHRER, J., AND WILCOX, D. 1994. Does consumer sentiment forecast household spending? If so, why? *Amer. Econ. Rev. 84*, 1397–1408.

CARROLL, C., SLACALEK, J., AND SOMMER, M. 2010. International evidence on sticky consumption growth. *Rev. Econ. Statist*.

DICKEY, D. A. AND FULLER, W. A. 1979. Distribution of the estimators for autoregressive time series with a unit root. *J. Amer. Statist. Assoc. 74*, 427–431.

DOMS, M. AND MORIN, N. 2004. Consumer sentiment, the economy, and the news media. FRBSF working paper.

FRIEDMAN, M. 1957. *A Theory of the Consumption Function*. Princeton University Press, Princeton, NJ.

FRIEDMAN, N., GEIGER, D., AND GOLDSZMIDT, M. 1997. Bayesian network classifiers. *Mach. Learn. 29*, 131–163.

GODFREY, L. G. 1978. Testing for multiplicative heteroskedasticity. *J. Econometrics 8*, 227–236.

GRANGER, C. W. J. AND NEWBOLD, P. 1974. Spurious regressions in econometrics. *J. Econometrics 2*, 111–120.

HALL, R. E. 1978. Stochastic implications of the life-cycle/permanent income hypothesis: Theory and evidence. *J. Polit. Econ. 96*, 971–987.

HAND, D. J. AND YU, K. 2001. Idiot's Bayes—Not so stupid after all? *Int. Statist. Rev. 69*, 385–399.

HAYASHI, F. 1982. The permanent income hypothesis: Estimation and testing by instrumental variables. *J. Polit. Econ. 90*, 895–916.

HENDRY, D. F. 1993. *Econometrics: Alchemy or Science? Essays in Econometric Methodology*. Blackwell Publishers, Oxford, UK.

KOTSIANTIS, S. B. AND PINTELAS, P. E. 2004. Increasing the classification accuracy of simple Bayesian classifier. In Lecture Notes in Artificial Intelligence, vol. 3192. Springer, 198–207.

LEWIS, D. 1998. Naive (Bayes) at forty: The independence assumption in information retrieval. In *Proceedings of the 10th European Conference On Machine Learning (ECML'98)*.

LLOYD, B. T. JR. 1999. Survey measures of expected U.S. inflation. *J. Econ. Perspect. 13*, 125–144.

RISH, I. 2001. An empirical study of the naive Bayes classifier. In *Proceedings of the Workshop on Empirical Methods in Artificial Intelligence (IJCAI)*.

SOMMER, M. 2007. Habit formation and aggregate consumption dynamics. *B.E. J. Macroecon. Advances. 7*, Article 21.

THEIL, H. 1958. *Economic Forecasts and Policy*. North Holland, Amsterdam.

WHITE, H. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica 48*, 817–838.

WORKING, H. 1960. Note on the correlation of first differences on averages in a random chain. *Econometrica 28*, 916–918.

Received October 2010; revised January 2011; accepted April 2011