The Role of Media in Causing Recessions

Ryan Shuster

University of California Davis

Graduate School of Management

Abstract

In this paper, I examine the effect of newspaper sentiment on causing recessions. To this end, I estimate a probit model using data on newspaper sentiment, consumer confidence, and several known macroeconomic predictors such as the inversion and flatness of the yield curve. Controlling for these known effects, I discover a new result that the ratio of negative to positive sentiment in newspapers significantly increases the likelihood of recession. Thus, the media’s reporting of the economy plays a role in causing recessions.

**1. Introduction**

While our ability to predict recessions has gotten better, our understanding of the cause of recessions remains blurry. Traditional research into leading indicators of recessions have focused on multiple economic factors and the yield curve. The literature shows that the yield curve is a forward-looking indicator that serves as a predictor of recessions. The yield curve has been the best single predictor of recessions in the past (Estrella & Mishkin, 2006)( Zaloom, 2009). In Caitlin Zaloom’s survey of economic literature on the yield curve, she lays out the four theories that could shape the yield curve. One theory is that differing demand for liquidity changes the spread of the yield curve. The second theory states that investor’s changing preferences for maturity length effect the curve’s spread. The “market expectations hypothesis” postulates that the yield curve is shaped by investors predictions of future interest rates. The final theory includes investors’ expectations about the future. If we can directly measure investors’ expectation about the future, we would not need to rely on the yield curve as a predictor.

In their 2013 study, Christiansen, Eriksen, and Moller compared consumer and investor sentiment surveys to traditional recession predictors (2013-14). They found that sentiment out-performed the traditional macroeconomic predictors and the yield curve. This finding reflects the research on the predictive power of the yield curve, but also hints that consumer sentiment, not just investor sentiment, is useful to predict recessions. While the yield curve theoretically includes investor’s expectations it likely leaves out consumer expectations. If we can include consumer expectations in the model we can get closer to understanding the cause of recessions. Ben S. Bernanke, former Chairman of the Board of Governors of the Federal Reserve System, expresses the importance of sentiment in the following way: “As in all past crises, at the root of the problem is a loss of confidence by investors and the public in the strength of key financial institutions and markets.” Hence, this paper attempts to use sentiment analysis of newspaper articles as a proxy for public sentiment traditionally ignored in econometric models.

**2. Data Description**

**2.1 News Paper Sentiment**

The textual data comes from the website newspaperarchive.com. This website has daily newspaper images that have been ran through optical character recognition (OCR). This process results in a database of newspaper articles in a plain text format that can be used for sentiment analysis. I selected the newspaper the “Daily Herald Suburban Chicago” because its issues are from 1960 to present. This paper was also selected because it had fewer missing weeks than most newspapers. I scraped weekly newspaper text from the website from the front page of the “Daily Herald Suburban Chicago” between 01/14/1960 and 01/31/2018. The resulting data had 321 missing weeks out of 2149.

To perform sentiment analysis, I used lists of negative and positive words to “score” texts. These word lists are traditionally tuned to pick up on common positive or negative words. The more positive words in an article the higher the positive sentiment. To understand consumer sentiment in relation to the economy, we need word lists tuned to business activity. In “Textual Analysis in Accounting and Finance: A Survey,” Tim Loughran and Bill McDonald establish positive and negative word lists compiled from financial documents (2016). They correlated stock returns with words in 10k filings and found that these words have significant predictive power. I used these two word lists to create two variables, PosS (Figure 1) and NegS (Figure 2). These variables are the count of the words from the positive and negative word lists. Then these two were combined to make the variable sratio = NegS/PosS (Figure 3). The missing weeks where filled in by linear interpolation. Equation (1) presents the interpolation equation for a missing week j.

(1)

Finally, I averaged the weekly sentiment scores to arrive at the monthly sentiment score.

Figure 1

Figure 2

Figure 3

**2.2 Economic Indicators**

In addition to sentiment, macroeconomic indicators have been found to be robust indictors of recessions. Including a robust set of traditional indicators is important because it allows me to examine if sentiment does in fact hold information about recessions. There are many macroeconomic indicators to look at. Consequently, indicators are often condensed into indexes (Stock & Watson, 1989). I will use a few select measures instead, these measures retain most of the predictive power of the indexes (Davig & Hall, 2017). The indicators I selected are given in Table 1

**Table 1**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Data Label** | ISM | M-Sent | UOrders | Emp | Hauth | Inv | S&P |
| **Description** | Institute of supply management index (1-100) | Consumer sentiment from the University of Michigan | $ US unfilled orders | Thousands of Employee, Non-Farm | # of new housing authorization | $ US inventories | $ S&P 500 Market Cap |
| **Paper** | Troy Davig and Aaron Smalter Hall  2017 | Charlotte Christiansen, Jonas Nygaard Eriksen, and Stig V. Møller  2014 | James H. Stock and Mark W. Watson  1989 | Troy Davig and Aaron Smalter Hall  2017 | James H. Stock and Mark W. Watson  1989 | James H. Stock and Mark W. Watson  1989 | Troy Davig and Aaron Smalter Hall  2017 |

The variables UOrders, Hauth, Inv, and S&P have been shown to be positively correlated with economic growth. I therefore expect all of them to be negative indicators of recessions. Employment is a more complicated story. It is highly correlated with inflation. In the periods of 1960 to 2017, the correlation between employment and consumer price index (CPI) was 0.97. Inflation often affects recessions, but both too high or too low inflation can cause a recession (McMahon, 2011). Thus, the Emp variable will increase or decrease the likelihood of recession. The inclusion of these traditional measures will validate or invalidate the hypothesis that consumer sentiment contains additional predictive power over traditional models.

I included two sentiment variables: ISM and M-Sent. These variables control for consumer and professional sentiment in the economy. They are expected to have a negative effect on recessions (Christiansen, Eriksen, & Moller, 2013-14). The better people feel about the economy, the less likely it will go into recession. The inclusion of these variables highlights the effect of the media on recessions because these variables control for the actual sentiment of the period.

The dependent variable is defined by NBER’s recession and expansion dates. NBER business cycle dataset is publicly available and standard in the recession literature (Estrella & Trubin, 2006). This dataset defines the first recession month as the month following a peak and the last month of the recession as the final month of the trough. The data sample runs over the period 1960 – 2017. The dependent variable is binary, where “1” indicates the economy is in recession and a “0” denotes the economy is in expansion. The sample contains 8 recessionary periods.

**2.3 Yield Curve**

The slope of the yield curve has been found to be the best leading indicator, specifically the spread between the 10-year U.S. T-bonds and the yield on the 1-year U.S. T bonds (Estrella & Mishkin, 1996). The level of the yield curve has also been shown to be a significant factor in predicting recessions. To calculate these three variables, I started out with a data set of the yields for the following treasuries: 1 month, 3 month, 1 year, 2 year, 3 year, 5 year, 7 year, 10 year, 20 year, and 30 year. I then used linear regression with the equation 2 to estimate the slope (, curvature (, and level of yield of a zero-duration treasury (.

Equation 2:

From prior research we know that negative slopes should be a leading indicator of recession for up to 6 months (Estrella & Trubin, 2006). The curvature, estimated by , measures the lack of flatness of the yield curve. In other words, as curvature increases, the curve becomes humped in the middle as medium term yields increase. The existing literature suggests that level and curvature of the yield curve are positively correlated with macro-economic factors. Therefore, these variables are expected to have a negative effect on the probability of recession.

**3. Model**

**3.1 Model Selection**

The main interest of this paper is to assess the predictive power of the newspaper sentiment after controlling for the macroeconomic variables. Therefore, I want to build a model to forecast the probability that the binary NBER recession variable = 1 at time t. To this end, I use the Probit model, where the probability of a recession is as follows:

Equation 3:

where denotes the Cumulative Distribution Function of the normal distribution. This transforms the NBER variable to always be in the interval [1,0]. Equation (5) below specifies as a linear combination of various variables in the model. Equation 4 shows the simple version of this. This model suggests that the probability of a recession (Pt) is a function of a macroeconomic, yield curve, and sentiment variable (x). The macroeconomic variables are coincident indicators with a lag (p) of 0. Sentiment and yield curve slope are included with lags of 0 to 6.

Equation 4:

Building on this model, I hypothesized that the lags could built on each other. I added 6 lagged slope variables, where i denotes the number of lagged months. Equation 5 shows that X is the sum of multiple lags at time t minus a lag of p months. These are weighted by a weight wi. (to be estimated).

Equation 5:

where . The variables in are linearly combined as shown in Equation 6

Equation 6:

Where n is the number of variables in the model. Note that this model does not have an intercept because the independent variables contain all the information about recession likelihood. In other words, if intercept were included, then when all , the model would predict 50% chance of prediction, which lacks face validity.

The coefficients are then estimated using maximum likelihood estimation given in Equation 7.

Equation 7:

**3.2 Variable Selection**

The model was originally estimated using all the independent variables. I excluded variables that had a p-value > .05. This resulted in the final variables shown in Table 2. To select the best lag for the slope variable, I re-ran the model including 7 slope variables with lags of zero to six months. When including all the lag variables only Slope\_6 had a p-value > .05. Therefore, Slope\_6 was kept in the final model.

All variables affect the probability of recession as expected. All but one macroeconomic indicators had a negative effect on Pt. The only exception was employment, which has a positive effect on recession: the higher the employment the greater the likelihood of recession. This result follows from the fact that employment and inflation are highly correlated, and that inflation is known in the literature to have a positive effect on recessions.

After selecting the macroeconomic variables, I added the two sentiment variables in the model. NegS and PosS had the expected effect (NegS had a positive effect on recessions, PosS had a negative effect on recessions), but only PosS was significant. For the sake of parsiomony, I constructed sentiment ratio, sratio, to captures the *relative* sentiment; it is defined as sratio = NegS/PosS. This transformation increased the R^2 of the model from 54% to 55%. This ratio also had the expected effect on the chance of recession. If a newspaper had more negative words relative to positive words it increases the chance of recession.

To show its power as a leading indicator, I ran the model with lagged 7 sratio variables, with lags varying from zero to six. Sratio with a lag of two was the only significant variable in this model. Adding the two-month lagged sratio increased the R^2 from 55% to 57%.

**4. Empirical Findings**

**4.1 Sentiment Variables**

The results show that the lack of positive sentiment is more predictive of recessions than negative sentiment. This is shown by NegS not being significant while PosS is. More interestingly, the ratio of NegS to PosS is even more predictive than either variable on their own. The ratio captures the true sentiment of the time better than either variable by themselves. This is because while some period may have lots or little of financial news which will change the absolute about of sentiment, the ratio is the true indicator of how the economy is heading.

The finding that the lag of two months is significant supports our hypotheses that sentiment is a cause of recessions, not just a reaction to bad markets. This means that sentiment is a predictive variable. In fact, using sentiment as a coincident indicator is less predictive of recessions than using it as a leading indicator. This shows us that sentiment is one of the main drivers of the economy.

The results from table 2 can be interpreted as follows. We observe that the change in the ratio of negative to positive words greatly affects the probability of recession. If a newspaper moves from 1:1 to a 2:1 negative to positive sratio, the chance of recession is 42.5 times more likely. *To put this in perspective, the S&P would have to fall 373 points to create the same effect.*

**Table 2**

* 1. Results

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Coefficients: |  |  |  |  |  |  |
|  | Estimate | Odds Ratio:  exp(Estimate) | Std. | Error | z-value | Pr(>|z|) |
| sratio\_2 | 3.75E+00 | 4.250782e+01 | 8.35E-01 | 4.492 | 7.05E-06 | \*\*\* |
| S.P | -5.05E-03 | 9.949609e-01 | 1.00E-03 | -5.041 | 4.63E-07 | \*\*\* |
| Emp | 1.23E-04 | 1.000123e+00 | 1.75E-05 | 7.028 | 2.09E-12 | \*\*\* |
| Slope\_6 | -1.38E+01 | 1.055739e-06 | 1.95E+00 | -7.045 | 1.86E-12 | \*\*\* |
| Yield Curve Level | -1.68E-01 | 8.450906e-01 | 8.46E-02 | -1.989 | 0.046715 | \* |
| Curvature | -8.77E+00 | 1.561414e-04 | 2.74E+00 | -3.204 | 0.001353 | \*\* |
| ISM | -9.14E-02 | 9.126164e-01 | 2.61E-02 | -3.507 | 0.000453 | \*\*\* |
| M.Sent | -8.82E-02 | 9.155531e-01 | 1.57E-02 | -5.63 | 1.81E-08 | \*\*\* |

Newspaper sentiment is as descriptive as some other traditional indicators. In table 3, we can see that sratio\_2 is tied to be the 5th most descriptive variable. It is tied with the flatness of the yield curve, measured by decreasing curvature. It is also more important than ISM as well as the level of the yield curve. Hence, newspaper sentiment possesses predictive ability and offers a new variable to predict recessions.

**Table 3**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Variables Removed | None | sratio\_2 | S&P | Emp | Slope\_6 | Curvature  (low values mean yield curve becomes flat) | Yield Curve Level | ISM | M-Sent |
| Months Lag | 0 | 2 | 0 | 0 | 6 | 0 | 0 | 0 | 0 |
| R2 | 57 | 53 | 47 | 38 | 36 | 53 | 56 | 54 | 49 |
| Change |  | 4 | 10 | 19 | 21 | 4 | 1 | 3 | 8 |
| Importance |  | 5th | 3rd | 2nd | 1st | 5th | 7th | 6th | 4th |

**4. Conclusion**

This paper sheds light on how sentiments in newspapers predict recessions. It suggests that the news media has a role in causing recessions. The ratio of negative to positive sentiment is especially important, revealing that negative stories in media spread fear and initiates a recession in subsequent months. In today’s charged media landscape where negative stories draw more attention than positive stories, the media should not only be cognizant of this finding, but also It maintain positive stories in the headlines to keep the sentiment ratio good and bolster the confidence of the market.

References

Christiansen, C., Eriksen, J. & Moller, S. (2013-14). Forecasting US Recessions: The Role of Sentiments. CREATES. Retrieved from   
<ftp://ftp.econ.au.dk/creates/rp/13/rp13_14.pdf>

Blanchard, J. & Fischer S. (1989). New Indexes of Coincident and Leading Economic Indicators. National Bureau of Economic Research, Volume 4. Retrieved from   
<http://www.nber.org/chapters/c10968.pdf>

Davig, T. & Hall, A. (2017). Recession Forecasting Using Bayesian Classification. The Federal Reserve Bank of Kansas City Research Working Papers, Issn 1936-5330. Retrieved from   
<https://www.kansascityfed.org/~/media/files/publicat/reswkpap/pdf/rwp16-06.pdf>

Estrella, A., & Trubin, M. (2006). The Yield Curve as a leading Indicator: Some Practical Issues. Current Issues in Economics and Finance, Volume 12 Number 5. Retrieved from   
https://www.newyorkfed.org/research/current\_issues/ci12-5.html

Estrella, A., & Mishkin, F. (1996). The Yield Cure as a Predictor of U.S. Recessions. Current Issues in Economics and Finance, Volume 2 Number 7. Retrieved from   
<https://www.newyorkfed.org/research/current_issues/ci2-7.html>

Loughran, T. & McDonald, B. (2016). Textual Analysis in Accounting and Finance: A Survey. Journal of Accounting Research, Volume 54 Issue 4. Retrieved from   
<https://onlinelibrary.wiley.com/doi/abs/10.1111/1475-679X.12123>

Mcmahon, T. (2011). Recessions and Inflation since 1914. InflationData. Retrieved from   
<https://inflationdata.com/Inflation/Inflation_Articles/Inflation_and_Recession.asp>

Stock, J. & Watson, M. (1993). A Procedure for Predicting Recessions with Leading Indicators: Economic Issues and Recent Experience. National Bureau of Economic Research, Business Cycles, Indicators and Forecasting. Retrieved from   
<http://www.nber.org/chapters/c7190.pdf>

Zaloom, C. (2009). How to Read the Future: The Yield Curve, affect, and Financial Prediction. Public Culture, Volume 21 Issue 2. Retrieved from   
https://read.dukeupress.edu/public-culture/article-pdf/21/2/245/488573/PC21.2\_05\_Zaloom\_sh.pdf