The Asymmetric Effects of Investor Sentiment

Chandler Lutz*

Copenhagen Business School

cl.eco@cbs.dk

Forthcoming in *Macroeconomic Dynamics*

November 25 2014

Abstract

We use the returns on lottery-like stocks and a dynamic factor model to construct a novel index of investor sentiment. This new measure is highly correlated with other behavioral indicators, but more closely tracks speculative episodes. Our main new finding is that the effects of sentiment are asymmetric: During peak-to-trough periods of investor sentiment (sentiment contractions), high sentiment predicts low future returns for the cross-section of speculative stocks and for the market overall, while the relationship between sentiment and future returns is positive but relatively weak during trough-to-peak episodes (sentiment expansions). Overall, these results match the theories and anecdotal accounts of investor sentiment.

JEL Classification: G11, G12, G17

Keywords: Behavioral Finance, Asymmetric Investor Sentiment

*I'd like to thank two anonymous referees and numerous people and conference participants throughout North America and Europe for their helpful comments.

In this paper, we develop a unique index of investor sentiment (henceforth, SENT) by estimating a latent factor from the returns on lottery-like stocks. This index provides a novel proxy for stock market sentiment while at the same time enabling new insights into the relationship between investor behavior and returns. We find that (1) SENT more closely tracks anecdotal accounts of investor sentiment over our sample period than previously constructed indicators; (2) the index predicts implied volatility, media pessimism, and stock returns; and (3) the predictive effects of sentiment with regard to stock returns are asymmetric in that they are negative, large in magnitude, and highly significant during peak-to-trough episodes of investor sentiment (sentiment contractions), but often positive and small in magnitude during trough-to-peak episodes (sentiment expansions). This third result, the main new finding in our paper, matches the theoretical predictions of investor sentiment outlined in Abreu and Brunnermeier (2003) and holds for a broad cross-section of speculative stocks and for the market overall.

We compile SENT through a dynamic factor model from the returns on lottery-like stocks. Stocks with lottery-like characteristics are used as individual investors are attracted to their speculative features (Kumar (2009)), while the dynamic factor model allows for the construction of a latent common component in levels from a group of return series. Thus, the econometric framework yields a new sentiment proxy (SENT) that is easily comparable to anecdotal accounts of investor behavior over the sample period and other sentiment aggregates.

Our index, which increases with investor optimism and decreases with pessimism, is highly positively correlated with the sentiment measure of Baker and Wurgler (2006) (henceforth, BWsent) and is inversely related to the VIX stock market fear gauge and the media pessimism proxy (henceforth, Pessimism) of Tetlock (2007).[1] The similarities between SENT and these other measures validate our index as a proxy of investor sentiment. Yet over the sample period, our index more accurately tracks speculative episodes than BWsent, the VIX, or Pessimism. Indeed, using the Bry and Boschan (1971) algo-

[1]The overall patterns of SENT and BWsent are substantially similar as the correlation coefficient between the two measures is large in magnitude at 0.588 and significant at the one percent level. The correlation coefficients between SENT and the VIX is -0.235 while the correlation between SENT and Pessimism is -0.292. Both of these coefficients are significant at the one percent level.

rithm, we date the turning points in SENT. We find that these turning points are nearly identical to those documented qualitatively by Baker and Wurgler (2006) from historical records and also lead the turning points found using BWsent, the VIX, and Pessimism. Thus, our index appears to capture investor behavior via different channels than previously developed indicators that allow for a more accurate timing of stock sentiment episodes.

Using our index, we study the effects of stock market sentiment on returns. Like previous empirical studies, we find that high sentiment predicts low future returns over our entire sample period; suggesting that rational arbitrageurs eventually correct sentiment-based mis-pricings for the overall sample.[2] Yet the crux of our investigation is the examination of the predictive effects of investor sentiment during disparate time periods. This analysis is motivated by the asymmetric behavior of sophisticated investors over boom and bust episodes. For example, Brunnermeier and Nagel (2004) find that hedge funds were not a corrective force during the technology bubble in the late 1990s and did not begin to reduce their long positions in tech stocks until September of 1999.[3] Thus, these anecdotal accounts suggest that sophisticated investors do not always act as a corrective influence in the presence of a sentiment based mispricing. To examine these qualitative findings empirically, we study the predictive effects of our index, SENT, during sentiment expansions and contractions. In sentiment expansions (trough-to-peak episodes of investor sentiment), we find that an increase in sentiment predicts higher returns, but that the effect is relatively small in magnitude. During sentiment contractions (peak-to-trough episodes of investor sentiment), however, high sentiment predicts low future returns. These latter predictive effects are large in magnitude, highly significant, and hold for a broad cross-section of stocks as well as the market overall.

Our main new finding implies that sentiment and future returns are positive but

---

[2]Baker and Wurgler (2007) provide an overview of papers that examine sentiment in the stock market. See also Lim and Brooks (2010).

[3]The tech bubble then popped in March of 2000. Similarly, Temin and Voth (2004) and Brunnermeier (2009) contend that sophisticated investors built long positions in highly speculative securities rather than trading against the expanding bubbles associated with South Sea episode in 1720 or the housing boom in the 2000s.

weakly related during sentiment expansions and negative and strongly related during contractions. In other words, the predictive effects of sentiment are asymmetric. These asymmetric predictive effects closely correspond with the theoretical framework of Abreu and Brunnermeier (2003). Abreu and Brunnermeier (2003) build a model where rational arbitrageurs face a synchronization problem that prevents them from immediately attacking a sentiment induced mispricing or bubble. Since the presence of the bubble is never common knowledge, each rational arbitrageur must predict when each other trader will bet against the mispricing. This allows price to differ from value for finite time periods. In equilibrium, Abreu and Brunnermeier find that rational speculators have a profit incentive to ride the bubble before trying to exit just before the crash. Hence, rational speculators build long positions in speculative securities as sentiment expands and then attempt to exit the market before pessimism takes hold; yielding an asymmetric relationship between sentiment and returns across boom and bust episodes that is congruent with our empirical results.

A number of other studies have empirically considered the relationship between investor sentiment and stock returns. Baker and Wurgler (2007) provide an overview of this literature. In general, these papers assume that the predictive effects of investor sentiment are homogeneous over the entire sample and find that high sentiment relates to low future returns. While our work is similar to these previous studies, we extend the literature by considering the effects of a novel sentiment index on returns over sentiment expansions and contractions. Furthermore, other papers, such as Baker and Wurgler (2006) and Stambaugh, Yu, and Yuan (2012), consider the predictive effects of high and low investor sentiment relative to the mean or median. Although these papers are related to our work, they are decidedly different as sentiment can be above average (or above the median) during both a sentiment expansion *and* a sentiment contraction.[4]

Lastly, our econometric methodology builds on a large set of recent papers that employ dynamic factor models within macroeconomics and financial economics. These studies

[4]For example, both our sentiment index and that of Baker and Wurgler (2006, 2007) were well above average just prior to and just after the conclusion of the tech bubble in March of 2000 while sentiment transitioned from a state of expansion to a state of contraction over this period.

use dynamic factor models to extract a relevant common component from set of key variables.[5] Our work is similar to this literature as we extract a latent sentiment proxy from a key set of return series that possess lottery-like characteristics.

The rest of the paper proceeds as follows: We describe the data in section 1; sections 2 and 3 cover the estimation and interpretation of our sentiment measure; section 4 discusses the sentiment contractions and expansions; sections 5 and 6 outline the predictive regressions; 7 provides an interpretation of the results; and 8 concludes.

1 The Data

We consider the returns on lottery-like stocks to measure investor sentiment. These stocks are speculative securities with high betas. In other words, they are high risk, high reward. Kumar (2009) finds that individual investors, the investors most associated with agent sentiment, are attracted to stocks with lottery-like characteristics.

More specifically, our dataset includes the difference in returns between stocks that do not pay dividends and those that do (henceforth, Div), companies with earnings less than or equal to zero and those with positive earnings (henceforth, Earn), and small and large firms (henceforth, Size). A small (large) firm is defined as one whose market cap is in the bottom (upper) 20 percent. Equal weighted returns on low momentum firms (henceforth, Lowmom) are also used to represent companies in distress. We classify firms whose returns are in the bottom ten percent for the previous two to twelve months as having low momentum. Data on these series runs from July 1951 to September 2009.

In general, lottery-like stocks have less information available which allows investors to defend a wide range of valuations. These stocks are also risky to arbitrage due to their high idiosyncratic volatility (Wurgler and Zhuravskaya (2002)). Together, these characteristics make lottery-like stocks speculative in nature.

Table 1 shows the summary statistics for the aforementioned series. For comparison purposes, the summary statistics for the S&P500 are also included. Div, Earn, and

[5]See, for example, Ludvigson and Ng (2009), Chauvet and Potter (2000), Chauvet (1998), Barnett, Chauvet, and Tierney (2008), Chauvet and Piger (2008), Kishor and Neanidis (2012), Eickmeier and Hofmann (2013), and Fuleky and Bonham (2013). See also Balke et al. (2015).

Size all have lower average returns than the S&P500, but higher standard deviations. Lowmom is the most volatile series and produces the largest average monthly returns at 0.790 percent. Moreover, Lowmom yields the largest maximum monthly return at 65.030 percent; this is about five times as large as the maximum monthly return on the S&P500. Similarly, the maximum monthly returns on Div, Earn, and Size are also quite high. The large maximum monthly returns for the these stocks are indicative of their lottery-like nature.

[**Insert Table 1 About Here**]

Table 2 presents the correlation coefficients between Div, Earn, Size, and Lowmom. All four variables are highly correlated. Div and Earn are the most closely related with a correlation coefficient of 0.86, while the relationship between Size and Lowmom is weakest with a correlation of 0.62.

[**Insert Table 2 About Here**]

1.1 Other Investor Sentiment Measures

We consider three common stock sentiment measures used in the literature: Baker and Wurgler's (2006) sentiment index; the VIX index; and media pessimism based on the Wall Street Journal Column "Abreast of the Market."

Baker and Wurgler (2006) compile their index (henceforth, BWsent) by extracting a common component from traditional sentiment measures.[6] These measures include the closed-end fund discount, NYSE share turnover, the number of IPOs in a given month, the first day return on IPOs for each month, the equity share of new issues, and the dividend premium. The data range from August 1965 to July 2007. BWsent rises during optimistic times and falls as pessimism takes hold. Baker and Wurgler find that high levels of BWsent relate to low future returns for a broad cross-section of speculative stocks.

The VIX index captures the implied volatility from S&P500 options. Practitioners consider the VIX index to be a measure of investor fear (Whaley 2000). We use the old

[6] We obtain BWsent from Jeffrey Wurgler's website.

formula traded under the symbol VXO as it has a longer sample dating back to 1986. The data for the VIX was downloaded from the Chicago Board Options Exchange (CBOE). We average over days to obtain the VIX at the monthly frequency. High values in the VIX index imply high investor fear, and vice versa.

We also construct a measure of media pessimism using the Wall Street Journal column "Abreast of the Market" as in Tetlock (2007). This column is available daily from the Factiva database and ranges from January 1984 through December 2001. We compile media pessimism (henceforth, Pessimism) from the portion of negative words in each article using the Loughran and McDonald (2011) financial dictionary via the General Inquirer software. Pessimism is transformed to the monthly periodicity by averaging over days. Tetlock contends that media pessimism captures sentiment as it is correlated with the VIX index and BWsent.

1.2 Anecdotal Accounts of Investor Sentiment

Baker and Wurgler (2006) qualitatively highlight five major sentiment episodes over our sample period.[7] First, Baker and Wurgler document an electronics boom in 1961 that ended in 1962. Another bubble then developed during the late 1960s as investors preferred young growth stocks; this speculative episode ended in 1969. High sentiment times transpired again in the late 1970s concluding with a "hot-issue" market in 1980. A technology new issue boom then arose before ending in the second half of 1983. Baker and Wurgler also note that sentiment waned in the early to mid 1970s and then again during the mid to late 1980s. Lastly, in the late 1990s, investors flocked to technology stocks creating the greatest speculative mania since the 1920s (Shiller (2006)). This tech bubble burst in March of 2000. Below we compare the turning points of SENT to these episodes; this will help us validate our index as a measure of investor sentiment.

1.3 Stock Market Data

We consider a number of stock portfolios to study the predictive effects of investor sentiment. First, decile portfolios are formed based on the following firm characteristics:

[7] See the references in Baker and Wurgler (2006) for a more detailed analysis of these episodes.

Volatility, age, book-equity over market-equity (BE/ME), dividends, earnings, size, and momentum. Then we create long-short portfolios where we compare the returns across deciles. More specifically, we consider (1) the returns on high volatility stocks minus those on low volatility stocks (henceforth, σ), where a stock has high (low) volatility if its previous 2-12 month standard deviation of returns is in the upper (lower) 30 percent of all stocks; (2) the returns on young stocks minus those on old stocks (henceforth, Age), where a stock is young (old) if its age from the first month that it is listed in the CRSP is in the bottom (top) three deciles; (3) the Fama-French HML factor; (4) the returns on medium value stocks minus those on low value stocks based on book-equity over market-equity, where a stock is medium (low) valued if its BE/ME ratio is in deciles 4, 5, 6, or 7 (1, 2, or 3) of all stocks; (5) the returns on stocks that do not pay dividends less those that do (Div); (6) the returns on stocks with earnings less than zero minus those with positive earnings (Earn); (7) the Fama-French SMB factor; (8) the Fama-French MKT factor (excess market returns); (9) returns on high momentum stocks minus those on medium momentum stocks, where a stock has high (medium) momentum if its returns in the previous 2-12 months are in deciles 8, 9, or 10 (4, 5, 6, or 7) of all stocks; and (10) the returns on medium momentum stocks less those on low momentum stocks.

Portfolios based on size, earnings, dividends, book-equity over market-equity, and momentum are from Kenneth French's website. We compile the portfolios based on age and volatility from the CRSP database using share codes 10 and 11. Portfolios include all stocks listed on the NYSE, NASDAQ, and AMEX.

1.4 Other Data

A number of macroeconomic variables including industrial production, durable and non-durable consumption, and the BAA–AAA corporate bond spread are used as controls. These series are obtained from the FRED economic database of the Federal Reserve Bank of St. Louis. We also control for the monthly volatility of the S&P500. The monthly volatility is calculated by taking the standard deviation of the daily S&P500 returns.

2 Estimation of the Sentiment Index

We construct SENT from the return portfolios based on dividends, earnings, low momentum, and firm size through a dynamic factor model with Bayesian estimation. Dynamic factor analysis is used as it allows us to extract an unobserved common component in levels from our set of return series (log-differenced prices). Thus, by utilizing dynamic factor techniques, we can create a levels sentiment index that can subsequently be used in empirical tests and compared to related aggregates of investor behavior.

Although dynamic factor models are common in empirical macroeconomics, they are used less frequently in the financial literature. Ludvigson and Ng (2009) produce one financial application. They extract a common factor from standard macro variables and use this factor to predict the bond risk premium. Another application to finance is provided by Chauvet and Potter (2000). They use a dynamic factor model to develop coincident and leading indicators for the stock market.[8] In appendix A, we describe the dynamic factor model and its estimation in more detail.

Before we estimate the dynamic factor model, we orthogonalize the return series (Div, Earn, Size, and Lowmom) to macroeconomic indicators and measures of time-varying risk in the stock and bond markets using the following regression:

$$\begin{aligned} r_{it} = \alpha &+ \beta_1 INDPRO_t + \beta_2 IPDCONGD_t + \beta_3 IPNCONGD_t \\ &+ \beta_4 TREAS_t + \beta_5 NBER_t + \beta_6 sp500vol_t + \beta_7 (BAA - AAA) + \varepsilon_{it} \end{aligned} \tag{1}$$

Where r_{it} represents each of the return series, Div, Earn, Size, and Lowmom. We orthogonalize each return series to growth in industrial production ($INDPRO$), consumer durables ($IPDCONGD$), consumer nondurables ($IPNCONGD$), the one month Treasury Rate ($TREAS$), and a dummy variable for NBER recessions ($NBER$). These factors allow us to control for broad macroeconomic effects. We also control for the S&P500 monthly volatility ($sp500vol$) and the BAA–AAA corporate spread ($BAA - AAA$). $sp500vol$ and $BAA - AAA$ allow us to control for time-varying risk in the stock and bond markets.

[8] Other examples in the finance and business cycle literature include Chauvet (1998), Barnett, Chauvet, and Tierney (2008), and Chauvet and Piger (2008).

We estimate the dynamic factor model and derive our sentiment index, the unobserved common factor, using the residuals, the ε_{it}'s, from the regression described in equation 1. We label our index as SENT. SENT is standardized to have zero mean and unit variance. High values in our index correspond to high sentiment times, and vice versa.

3 SENT as a Measure of Investor Sentiment

We interpret SENT as a measure of sentiment as it is constructed from the returns on portfolios favored by individual investors; closely tracks investor behavior over the sample period; is positively correlated with known proxies of investor optimism; and is inversely related to investor fear and media pessimism. First, SENT matches anecdotal accounts of sentiment over the sample period as evidenced by the plot in figure 1: SENT increased during the optimistic times of the late 1960s before crashing as the bear market ensued in 1969; the index then jumped with the hot IPO markets of the late 1970s and the early 1980s; in the second half of the 1980s, SENT retreated as optimism faded; and during the 1990s tech bubble, SENT rose markedly reaching its global maximum in February of 2000. Recall that the tech boom, which concluded in March of 2000, was the largest stock market sentiment episode over our sample period (Shiller (2006)). Finally, our sentiment measure waffled in the mid-2000s before crashing with the onset of the financial crisis in 2007. This last observation is congruent with our expectations as optimism during this recent episode was largely related to the housing boom, while the pessimism associated with the bust permeated through all financial markets.

[**Insert Figure 1 About Here**]

Moreover, Shiller (2006) argues that sentiment jumped in the early 1980s and has remained high ever since. The plot of our sentiment measure in figure 1 matches Shiller's assertion as SENT was markedly higher and above average (above zero) for most of the sample after 1978.

One potential concern in the use of return data for the construction of our index may be that SENT is just reflecting its underlying components and not an unobserved common factor. To alleviate these doubts, we calculate the correlation coefficients between SENT and Div, Earn, Lowmom, and Size. Table 2 displays the results. Clearly, SENT is largely unrelated to its components as the largest correlation coefficient occurs between SENT and Size with a value of just 0.05 that is not statistically significant. Thus, our index appears to capture a common factor that differs substantially from the raw underlying returns but is strongly related to the shared speculative episodes surrounding those returns.

3.1 SENT versus other Investor Sentiment Indicators

To further validate SENT as a measure of investor sentiment, we compare it to other behavioral proxies. We first make the comparison graphically. Figure 2 shows the plot of SENT versus BWsent. The overall pattern between the two series is strikingly similar. Both series rise during sentiment booms and fall as bear markets take hold. Furthermore, both series match the speculative episodes described in Baker and Wurgler (2006) as they were high in the late 1960s, low in the mid 1970s, high in the early 1980s, and spiked in the early 2000s. Yet just by looking at the graph, we can tell that SENT leads BWsent: SENT peaked first following the boom times in the late 1960s, the late 1970s, the early 1980s, and the early 2000s. Thus, our index appears to capture the palpable swings in investor sentiment in a more timely fashion than Baker and Wurgler's index.

[**Insert Figure 2 About Here**]

Next, figure 3 shows SENT versus the VIX index. In figure 3, we multiply the VIX index by (-1) so that VIX*(-1) falls as investors become more pessimistic; this makes it easy to compare the VIX to SENT. Both SENT and VIX*(-1) crashed during bear markets over the sample period, but VIX*(-1) fell further during the 1987 crash and the recent financial crisis. Additionally, SENT spiked during the tech bubble in 2000, but VIX*(-1) did not. This matches our expectations as the VIX index captures fear rather than optimism (Whaley (2000)).

[**Insert Figure 3 About Here**]

In figure 4, we graph SENT versus Pessimism. Pessimism is multiplied by (-1) to ensure that the series falls as news stories become more negative. Even though

Pessimism*(-1) is volatile over the sample period, it did peak around the start of the bear markets in 1987, 1990, and 2000. Furthermore, Pessimism*(-1) shares a similar pattern with SENT. Most noticeably, the two series spiked around the tech bubble in 2000. Yet during this episode SENT peaked first while Pessimism*(-1) hit its high point after the onset of the bear market. This suggests that SENT leads Pessimism.

[Insert Figure 4 About Here]

Table 3 shows the correlations between SENT and the other behavioral indicators. All of the coefficients have the expected sign. SENT and BWsent are closely related with a correlation coefficient of 0.588 that is statistically significant at the 1 percent level. Our sentiment proxy is also negatively related to the VIX index. Lastly, SENT and Pessimism are negatively correlated and thus suggesting that media pessimism is low when sentiment is high.

[Insert Table 3 About Here]

In table 4, we present regression results where we use our index to predict the other behavioral indicators. Overall, SENT predicts the other proxies in the expected direction: High levels of SENT lead to high levels in BWsent, lower levels in the VIX index (investor fear), and low media pessimism. All of the regression coefficients on SENT are significant at the one percent level. Thus, SENT acts as a leading indicator. Moreover, SENT has more predictive power when BWsent is the dependent variable as the R^2 is largest for this regression. This latter result is not surprising as SENT and BWsent are similar in shape.

[Insert Table 4 About Here]

In sum, the figures and regression results in this section suggest that SENT is related to BWsent, the VIX, and media pessimism, but acts as a leading indicator and more closely tracks speculative episodes of investor behavior.

4 Sentiment Contractions and Expansions

Above we compared SENT to the sentiment indicators graphically and through predictive regressions. In this section, the Bry and Boschan (1971) algorithm is used to date the quantitative turning points that represent sentiment contractions and expansions.[9] We undertake this analysis so that we can quantitatively compare the peaks and troughs in our index and related indicators to the highs and lows associated with notable sentiment episodes over the sample period and later study the predictive effects of our index during disparate time periods. Overall, the results from this section show that the cycles in SENT almost exactly match anecdotal accounts of speculative episodes and lead the cycles in the other behavioral indicators.

The Bry and Boschan algorithm is a set of conditional rules used to determine the cycles and turning points in a time series. Pagan and Sossounov (2003) employ the Bry and Boschan algorithm to date bull and bear markets. We outline the exact rules for the algorithm in appendix B. Future research may find our analysis in this section helpful as we provide specific dates for sentiment contractions and expansions over the sample period.

We apply the Bry and Boschan algorithm to SENT, BWsent, VIX*(-1), and Pessimism*(-1). As previously noted, the VIX and Pessimism are multiplied by (-1) so they fall as fear and negativity grow. The results from the analysis are presented in table 5. The left column within each panel lists local peaks; the right column lists local troughs. A peak represents the most optimistic point in the cycle; a trough represents the most pessimistic point. As noted above, we follow the business cycle literature and define a sentiment contraction as a peak-to-trough episode of investor sentiment. Thus, the date in the left panel represents the beginning of each sentiment contraction, while the date in the right panel marks the end of each sentiment contraction. The cycle for the sentiment indicator that occurred first during a given time period is listed in boldface font.

[9] For an overview of business-cycle dating techniques see Chauvet and Potter (2003) and Chauvet and Hamilton (2006). We used the updated version of the Bry and Boschan algorithm as in Pagan and Sossounov (2003).

[Insert Table 5 About Here]

Overall, the dates of the sentiment contractions listed in table 5 for SENT are nearly identical to those described anecdotally by Baker and Wurgler (2006): Our index peaked, high sentiment times ended, and sentiment contractions began in 196901 as a speculative market fizzled in 1969; in 198306 at the conclusion of a technology new issue boom in the second half 1983; and in February of 2000 as the 1990s technology bubble popped. The algorithm recorded other contractions that began in 1981, in the mid 1990s around the 1997 Asian Financial Crisis, and in 2004 prior to the recent financial crisis.

Furthermore, SENT, BWsent, VIX*(−1), and Pessimism*(−1) record cycles at similar times, but SENT appears to act as a leading indicator. For example, SENT closely tracked speculative episodes during the first half of the sample and peaked in 1969, 1981, and 1983, while BWsent hit its high points after optimism began to subside during all of these episodes.[10] After reaching local maxima in the mid 1980s, both SENT and BWsent fell until the late 1980s and early 1990s as optimism faded. In comparison, VIX*(−1) and Pessimism*(−1) hit local highs just prior to the October 1987 crash in 198707 and 198701, respectively. These fear and negativity measures then went through numerous cycles from the late 1980s to the mid 1990s. Cycles occur more frequently in VIX*(−1) and Pessimism*(−1) as these series are much more volatile.

The 1990s technology bubble was the largest sentiment boom and bust over our sample period and is only comparable to the stock market mania in the 1920s (Shiller (2006)). The bubble burst in March of 2000 after the tech driven NASDAQ Composite Index reached its all time high. Even though all of the sentiment measures peaked around the conclusion of the tech bubble, only our index timed the transition correctly: SENT reached its max in February of 2000 while BWsent, VIX*(−1), and Pessimism*(−1) hit their high points in 200102, 200008, and 200004, respectively. Thus, the turning point in SENT appears to correctly time the end of the 1990s speculative bubble and lead that of the other sentiment indicators.

Overall, the analysis from this section indicates that SENT closely tracks anecdotal

[10] BWsent was the only sentiment indicator available prior to 1984.

accounts of investor behavior and provides accurate dates for sentiment turning points over the sample period. Below, we use the dates for contractions and expansions in SENT to examine the asymmetric predictive effects of investor sentiment.

5 Predictive Return Regressions

In this section, we run regressions to determine the predictive power of sentiment on future returns in a model-based framework over our entire sample period. We regress the 10 aforementioned long-short portfolios on SENT_{t-1} using data from September 1951 to August 2009. This will give us the average predictive relationship between SENT and returns for the whole sample. In line with previous work, we include the three Fama-French factors and a momentum factor as controls. The model becomes

$$z_t = \alpha + \beta_1 \text{SENT}_{t-1} + \beta_2 MKT_t + \beta_3 SMB_t + \beta_4 HML_t + \beta_5 UMD_t + \varepsilon_t \qquad (2)$$

where z_t represents any one of the ten long-short portfolios. As usual, MKT, SMB, HML, and UMD are the market, small minus big, high value minus low value, and the momentum factors, respectively. If one of the controls is the dependent variable, we exclude it from the set of regressors. Table 6 shows the results. Bootstrapped p-values are listed in parentheses.

[**Insert Table 6 About Here**]

The coefficient on SENT_{t-1} is significant when the long-short portfolios based on age, BE/ME, dividends, earnings, or momentum are the dependent variables. For these regressions, the coefficient on SENT_{t-1} has the expected sign. Thus, high sentiment predicts low returns for (1) young stocks; (2) low value stocks based on BE/ME (e.g. growth stocks); (3) stocks without earnings; (4) stocks without dividends; and (5) stocks in distress. SENT also predicts returns in the expected direction when σ and MKT are the dependent variables, but the regression coefficients are not significant at the 15 percent level.

Overall, the predictive results in table 6 are similar to those found in earlier studies when the Fama-French and momentum factors are included as additional controls.[11]

6 Asymmetric Predictive Regressions

Above we studied the average predictive effects of investor sentiment over our entire sample period. In this section, we examine the predictive power of SENT during sentiment expansions and contractions. This will allow us to test for asymmetric effects. As noted above, we define a sentiment contraction as a time of diminishing investor sentiment or a peak-to-trough episode in SENT. The dates for the sentiment contractions are listed in the left most panel of table 5.

To study the relationship between our index and returns during sentiment contractions and expansions we use the following regression model:

$$z_t = \alpha + \beta_1 \mathrm{SENT}_{t-1} + \beta_2 \mathrm{SENT}_{t-1}^{Contr} + \varepsilon_t \quad (3)$$

where z_t is any one of the long-short portfolios described above and $\mathrm{SENT}_{t-1}^{Contr}$ equals SENT_{t-1} during a sentiment contraction and 0 otherwise. Hence, the total effect of sentiment on returns during a contraction will be the sum of the coefficients on SENT_{t-1} and $\mathrm{SENT}_{t-1}^{Contr}$ ($\beta_1 + \beta_2$), while the predictive effect of sentiment on returns during an expansion will be captured just by the coefficient on SENT_{t-1} (β_1). Below we augment the model and control for the Fama-French factors and momentum.

Table 7 shows the regression results based on the model in equation 3. In table 7, we report the coefficients on SENT_{t-1} (β_1), $SENT_{t-1}^{Contr}$ (β_2), and the sum of the coefficients on SENT_{t-1} and $\mathrm{SENT}_{t-1}^{Contr}$ ($\beta_1 + \beta_2$). In the far right column, the p-value from the F-statistic that tests the null that hypothesis that $\beta_1 + \beta_2 = 0$ is listed in parentheses.

[Insert Table 7 About Here]

In general, the coefficients on SENT_{t-1}, $\mathrm{SENT}_{t-1}^{Contr}$, and the sum of the coefficients on SENT_{t-1} and $\mathrm{SENT}_{t-1}^{Contr}$, $\beta_1 + \beta_2$, are highly significant. For the cross-section of

[11]Previous studies also use the model similar to that in equation 2 to test the predictive effects of investor sentiment. See Baker and Wurgler (2007) for an overview.

speculative stocks and for the market overall, the predictive effects of sentiment are positive but relatively small in magnitude during sentiment expansions (as evinced by the coefficient on SENT_{t-1}), but negative and large in magnitude during sentiment contractions ($\beta_1 + \beta_2$ in the far right column). Thus, the predictive effects of investor sentiment are asymmetric. For a specific example, consider the long-short portfolio based on Age. During sentiment expansions, a one standard deviation increase in SENT leads to an increase in returns of 0.025 percent for the next month that is statistically significant at the 1 percent level. In marked contrast, a one standard deviation increase in SENT during contractions leads to a decrease in returns of $0.025 - 0.083 = -0.059$ percent for the next month that is significant at the 1 percent level. Thus, high sentiment leads to low returns only during sentiment contractions, while an increase in SENT predicts higher future returns during sentiment expansions. Moreover, the total predictive effects of SENT are over twice as large during a contraction. As a second example, let excess market returns be the dependent variable. In this case, the predictive effects of SENT on returns during a sentiment expansion are positive and significant at the one percent level as a one standard deviation increase in SENT leads to an increase in excess market returns of 0.018 percent. Yet during contractions, a one standard deviation increase in sentiment predicts a decrease in returns of $0.018 - 0.084 = -0.067$ percent per month that is significant at the 1 percent level. Hence, high sentiment relates to low future excess market returns only during contractions.

The reversal of the predictive effects of investor sentiment with regard to expansions and contractions persists for all of the portfolios listed in table 7. Thus, high sentiment predicts elevated returns during sentiment expansions, but low future returns during sentiment contractions for volatile stocks, young stocks, low value or growth stocks, stocks without earnings or dividends, small stocks, the market overall, high momentum stocks, and low momentum stocks.[12]

Next, we analyze the effects of SENT on future returns during sentiment contractions

[12]All of the coefficients in table 7 are significant at the 1 percent level except for the total effect during a sentiment contraction when the Medium − Low portfolio based on momentum is the dependent variable.

and expansions while controlling for the Fama-French factors and momentum in the following regression framework:

$$z_t = \alpha + \beta_1 \text{SENT}_{t-1} + \beta_2 \text{SENT}_{t-1}^{Contr} + \beta_3 MKT_t + \beta_4 SMB_t + \beta_5 HML_t + \beta_6 UMD_t + \varepsilon_t \quad (4)$$

As above, z_t is any of the long-short portfolios based on the various firm or stock characteristics. If one of the Fama-French factors or the momentum factor is the dependent variable, then it is not included in the set of regressors. Table 8 shows the results. In general, the findings are substantially similar to those described above: High values of SENT predict high returns during sentiment expansions, but low returns in sentiment contractions.[13] Moreover, in accordance with our previous findings, the predictive magnitudes are much larger during sentiment contractions. Overall, the results based on the model outlined in equation 4 suggest that the asymmetric predictive effects of SENT persist even after incorporating the usual controls.

[Insert Table 8 About Here]

In sum, the analyses from this section imply that the effects of sentiment on returns are asymmetric and match the theoretical predictions of Abreu and Brunnermeier (2003).[14] Abreu and Brunnermeier contend that sophisticated investors advantageously hold long positions in speculative securities when optimism permeates through markets. Then when sentiment fizzles, these same sophisticated investors reduce their holdings and attempt to exit the market. In accordance with this theory, we find that high values of SENT predict high returns during sentiment expansions, but low future returns during sentiment contractions. The results also suggest that the predictive effects are much larger during times of falling sentiment. These findings are also consistent with the anecdotal accounts documented in Temin and Voth (2004), Brunnermeier and Nagel (2004), and Brunnermeier (2009).

[13]The results in 8 do diverge in a couple of cases. When excess market returns or the returns on the High−Medium portfolio based on the momentum represent the dependent variable, the coefficient on SENT_{t-1} becomes insignificant. All other coefficients are significant at the 15 percent level.

[14]See also Patterson and Douglas (2010).

6.1 Asymmetric Predictive effects using Consumer Confidence and Baker and Wurgler's Sentiment Index

For comparison purposes, we conduct the above analysis using the University of Michigan Consumer Confidence Surveys and Baker and Wurgler's index, BWsent, as a substitute for SENT. First, we start with the consumer confidence index published by the University of Michigan. We follow Lemmon and Portniaguina (2006) and orthogonalize the index to various macroeconomic time series.[15] The regression model now takes the following form:

$$\begin{aligned} z_t = \alpha &+ \beta_1 \mathrm{MICH}_{t-1} + \beta_2 \mathrm{MICH}_{t-1}^{Contr} + \beta_3 MKT_t + \beta_4 SMB_t \\ &+ \beta_5 HML_t + \beta_6 UMD_t + \varepsilon_t \end{aligned} \tag{5}$$

where z_t is any one of the long-short portfolios described above, MICH is the orthogonalized Consumer Confidence Index, and $\mathrm{MICH}_{t-1}^{Contr}$ equals MICH_{t-1} during sentiment contractions and 0 otherwise. We use the sentiment contraction dates based on SENT as those dates closely correspond with anecdotal accounts of investor behavior. This approach also helps avoid data mining.

We report the results in the left panel of table 9. In general, the sum of the coefficients on MICH_{t-1} and $\mathrm{MICH}_{t-1}^{Contr}$ $(\beta_1 + \beta_2)$ has the expected sign and is significant at the 15 percent level. Thus, during sentiment contractions, high levels of the Michigan Consumer Confidence Index predict low future returns for volatile stocks, low value stocks, stocks without earnings or dividends, small stocks, and low momentum stocks. These results are substantially similar to those found above using SENT. Moreover, the coefficients on MICH_{t-1} and $\mathrm{MICH}_{t-1}^{Contr}$ often have opposite signs, suggesting a reversal in the predictive effects of Michigan Consumer Confidence between sentiment expansions and contractions. Again, these results match our above findings, the anecdotal accounts documented in Temin and Voth (2004) and Brunnermeier and Nagel (2004), and the

[15] We retain the residuals from a regression of the Michigan Consumer Confidence Index on durable and non-durable consumption, industrial production, the risk free interest rate, and dummy variable for NBER recessions. Our approach differs slightly from that in Lemmon and Portniaguina as we use the monthly Michigan Consumer Confidence surveys and monthly macroeconomic time series; they use quarterly data.

theoretical predictions of Abreu and Brunnermeier (2003).

[Insert Table 9 About Here]

Next, we examine the predictive effects of BWsent after accounting for sentiment expansions and contractions. The regression model becomes

$$
\begin{aligned}
z_t = \alpha + \beta_1 \text{BWsent}_{t-1} + \beta_2 \text{BWsent}_{t-1}^{Contr} + \beta_3 MKT_t + \beta_4 SMB_t \\
+ \beta_5 HML_t + \beta_6 UMD_t + \varepsilon_t \qquad (6)
\end{aligned}
$$

where $\text{BWsent}_{t-1}^{Contr}$ equals BWsent_{t-1} during a sentiment contraction and 0 otherwise. As before, we use the sentiment contraction dates based on SENT. This approach may adversely affect our results as the turning points in SENT and BWsent differ in some cases. Yet we continue to use contractions based on SENT for the model in equation 6 as these dates more closely correspond to anecdotal accounts of investor sentiment.

The results are in the right panel of table 9. In accordance with our previous findings, the sum of the coefficients on BWsent_{t-1} and $\text{BWsent}_{t-1}^{Contr}$ is large in magnitude and statistically significant. Thus, during sentiment contractions, high BWsent predicts low future returns. During sentiment expansions, however, the coefficient on BWsent_{t-1} is often insignificant; suggesting that the predictive effects of BWsent only persist during sentiment contractions.[16] Overall, these results are similar to those documented above which imply that the predictive effects of investor sentiment are large during sentiment contractions, but weak and relatively small in magnitude during sentiment expansions.

7 Interpretation of the Results

Thus far, the results suggest the interpretation of our index as a measure of sentiment and that the predictive relationship between sentiment and returns is asymmetric. We contend that SENT is a measure of investor sentiment as (1) it is constructed from the returns on lottery-like stocks that attract individual investors; (2) SENT closely tracks anecdotal episodes of investor behavior over the sample period; (3) the quantitative turning points in SENT found using the Bry and Boschan algorithm are nearly identical to the sentiment

[16]Only when SMB is the dependent variable is coefficient on BWsent_{t-1} significant at the 10 percent level.

episodes outlined qualitatively in Baker and Wurgler (2006); (4) our index is markedly higher in the second half of the sample which matches the assertions by Shiller (2006) who argues that a number of social and cultural factors have led to increased sentiment since 1980; and (5) SENT is highly correlated with the sentiment proxy developed by Baker and Wurgler (2006) and is inversely related to the VIX stock market fear gauge and the media pessimism variable of Tetlock (2007).

Our index not only captures the salient swings in investor behavior, but also more accurately times sentiment booms and busts than previously developed proxies. Furthermore, we show that SENT acts as a leading indicator over the sample period. For example, around the conclusion of the tech bubble SENT peaked one month prior to the NASDAQ crash in February of 2000 while BWsent, VIX*(−1), and Pessimism*(−1) hit their high points in February of 2001, August of 2000, and April of 2000, respectively.[17] Thus, our index timed the conclusion of the bubble precisely while the other proxies lagged behind. We also show that SENT predicts these other behavioral proxies in a regression framework. Together, these analyses suggest that our index captures investor sentiment via different channels than previously developed measures that allow for a more accurate timing of speculative episodes.

Using our index, we study the effects of investor sentiment on stock returns. Like previous studies, we find that high sentiment relates to low future returns for young stocks, growth stocks, stock without earnings or dividends, and stocks in distress for our sample overall. Yet our main new insights pertain to the relationship between sentiment and returns over various subperiods. More specifically, we compare sentiment to returns during sentiment contractions and expansions. Our results indicate that the predictive effects of sentiment on returns during expansions are positive but relatively small in magnitude, while high SENT predicts low future returns during sentiment contractions. This latter predictive relationship is highly significant, large in magnitude, and often represents a reversal compared to the predictive effects under sentiment expansions. Moreover, our main finding is robust when we include the Fama-French and momentum

[17]Recall that we multiply the VIX and Pessimism by -1 so that they rise with investor optimism and fall with pessimism.

factors as additional controls and use Baker and Wurgler's sentiment index or the University of Michigan Consumer Confidence Index as a substitute for SENT. Together, our findings match the analyses of Brunnermeier and Nagel (2004), Temin and Voth (2004), and Brunnermeier (2009) that contend that sophisticated investors do not always act as corrective force in the presence of a sentiment based mispricing. As such, the results in this paper support the hypothesis that investor sentiment may affect stock returns through the synchronization risk of Abreu and Brunnermeier (2003).

8 Conclusion

In this paper, we apply the returns on lottery-like stocks to develop and test a new index of investor sentiment for the stock market. Lottery-like stocks are used as individual investors are attracted to their speculative features. We find that our index accurately times speculative episodes and predicts other measures of investor sentiment. Using our index, we study the predictive effects of sentiment on stock returns. Like previous empirical studies, we find that high sentiment relates to low future returns over our entire sample period. Yet we also consider the relationship between sentiment and returns over various subperiods using regression analysis. The research results indicate that the predictive effects of SENT are weak but positive during trough-to-peak episodes of investor sentiment (sentiment expansions), but negative, large in magnitude, and highly significant in peak-to-trough periods (sentiment contractions). This suggests that the relationship between sentiment and returns is asymmetric. Overall, our findings closely correspond to the theories of investor sentiment involving synchronization risk where rational arbitrageurs build long positions as sentiment expands and then attempt to reduce their holdings of speculative securities before the crash.

References

Abreu, D. and M. K. Brunnermeier (2003). Bubbles and Crashes. *Econometrica 71*(1), 173–204.

Baker, M. and J. Wurgler (2006). Investor Sentiment and the Cross-Section of Stock Returns. *Journal of Finance 61*(4), 1645–1680.

Baker, M. and J. Wurgler (2007). Investor Sentiment in the Stock Market. *Journal of Economic Perspectives 21*(2), 129–151.

Balke, N. S., J. Ma, and M. E. Wohar (2015). A Bayesian Analysis of Weak Identification in Stock Price Decomposiitions. *Macroeconomic Dynamics 19*(4).

Barnett, W. A., M. Chauvet, and H. L. Tierney (2008). Measurement Error in Monetary Aggregates: A Markov Switching Factor Approach. *Macroeconomic Dynamics*.

Brunnermeier, M. K. (2009). Deciphering the Liquidity and Credit Crunch 2007-2008. *Journal of Economic Perspectives 23*(1), 77–100.

Brunnermeier, M. K. and S. Nagel (2004). Hedge Funds and the Technology Bubble. *The Journal of Finance 59*(5), 2013–2040.

Bry, G. and C. Boschan (1971). Cyclical Analysis of Time Series: Selected Procedures and Computer Programs. *NBER Books*.

Carter, C. and R. Kohn (1994). On Gibbs Sampling for State Space Models. *Biometrika 81*(3), 541–553.

Chauvet, M. (1998). An Econometric Characterization of Business Cycle Dynamics with Factor Structure and Regime Switching. *International Economic Review*, 969–996.

Chauvet, M. and J. D. Hamilton (2006). Dating Business Cycle Turning Points. *Nonlinear Time Series Analysis of Business Cycles*.

Chauvet, M. and J. Piger (2003). Identifying Business Cycle Turning Points in Real Time. *Working Paper*.

Chauvet, M. and J. Piger (2008). A Comparison of the Real-Time Performance of Business Cycle Dating Methods. *Journal of Business & Economic Statistics 26*(1), 42–49.

Chauvet, M. and S. Potter (2000). Coincident and Leading Indicators of the Stock Market. *Journal of Empirical Finance 7*(1), 87–111.

Eickmeier, S. and B. Hofmann (2013). Monetary Policy, Housing Booms, and Financial (Im)balances. *Macroeconomic Dynamics 17*(04), 830–860.

Fuleky, P. and C. S. Bonham (2013). Forecasting with Mixed-Frequency Factor Models in the Presence of Common Trends. *Macroeconomic Dynamics*, 1–23.

Harvey, A. C., S. J. Koopman, and N. Shephard (2004). *State Space and Unobserved Component Models: Theory and Applications*. Cambridge University Press.

Kim, C.-J. and C. R. Nelson (1998). Business Cycle Turning Points, a New Coincident Index, and Tests of Duration Dependence Based on a Dynamic Factor Model with Regime Switching. *Review of Economics and Statistics 80*(2), 188–201.

Kishor, N. K. and K. C. Neanidis (2012). What is Driving Financial Dollarization in Transition Economies? A Dynamic Factor Analysis. *Macroeconomic Dynamics*, 1–20.

Kumar, A. (2009). Who gambles in the stock market? *The Journal of Finance 64*(4), 1889–1933.

Lemmon, M. and E. Portniaguina (2006). Consumer Confidence and Asset Prices: Some Empirical Evidence. *Review of Financial Studies 19*(4), 1499.

Lim, K.-P. and R. D. Brooks (2010). Why Do Emerging Stock Markets Experience More Persistent Price Deviations from a Random Walk Over Time? A Country-Level Analysis. *Macroeconomic Dynamics 14*(S1), 3–41.

Loughran, T. and B. McDonald (2011). When is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance 66*(1), 35–65.

Ludvigson, S. C. and S. Ng (2009). Macro Factors in Bond Risk Premia. *Review of Financial Studies 22*(12), 5027–5067.

Pagan, A. R. and K. A. Sossounov (2003). A Simple Framework for Analysing Bull and Bear Markets. *Journal of Applied Econometrics 18*(1), 23–46.

Patterson, D. M. and V. Sharma (2010). The Incidence of Informational Cascades and the Behavior of Trade Interarrival Times during the Stock Market Bubble. *Macroeconomic Dynamics 14*(S1), 111–136.

Shiller, R. J. (2006). *Irrational Exuberance.* Crown Business.

Stambaugh, R. F., J. Yu, and Y. Yuan (2011). The Short of It: Investor Sentiment and Anomalies. *Journal of Financial Economics*.

Stock, J. H. and M. W. Watson (1991). A Probability Model of the Coincident Economic Indicators. *Leading Economic Indicators: New Approaches and Forecasting Records 66.*

Temin, P. and H.-J. Voth (2004). Riding the South Sea Bubble. *The American Economic Review 94*(5), 1654–1668.

Tetlock, P. C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance 62*(3), 1139–1168.

Whaley, R. E. (2000). The Investor Fear Gauge. *The Journal of Portfolio Management 26*(3), 12–17.

Wurgler, J. and E. Zhuravskaya (2002). Does Arbitrage Flatten Demand Curves for Stocks? *The Journal of Business 75*(4), 583–608.

9 Tables

Table 1: Summary Statistics of the Sentiment Components and the S&P500

	Average	Std	Min	Max
Div	0.004	3.973	-14.730	26.213
Earn	0.108	4.634	-17.160	29.683
Lowmom	0.790	8.510	-28.460	65.030
Size	0.365	4.298	-13.530	23.350
S&P500	0.615	3.520	-20.391	12.021

Notes: This table presents the summary statistics of the sentiment components and the S&P500. Div represents the difference in returns between stocks that do not pay dividends and those that do; Earn represents the difference in returns between stocks without earnings and those with positive earnings; Size represents the difference in returns between small and large firms; and Lowmom represents the returns on low momentum firms. Std is the standard deviation of returns.

Table 2: Correlations of SENT and its Components

	SENT	Div	Earn	Lowmom	Size
SENT	1.000				
Div	0.027	1.000			
Earn	-0.026	0.862	1.000		
Lowmom	0.050	0.797	0.695	1.000	
Size	0.056	0.798	0.733	0.621	1.000

Notes: This table presents the correlations between SENT and its components.

Table 3: Correlations of Sentiment Indicators

	SENT	BWsent	VIX	Pessimism
SENT	1.000			
BWsent	0.588***	1.000		
VIX	-0.235***	0.317***	1.000	
Pessimism	-0.292***	-0.133***	0.191***	1.000

Notes: Correlations between SENT and the other sentiment indicators. BWsent is Baker and Wurgler's (2006, 2007) sentiment index, VIX is the VIX investor fear gauge, and Pessimism is media pessimism based on Tetlock (2007). One, two, or three asterisks represents significance at the 15, 10, or 5 percent levels respectively.

Table 4: Predictive Regressions of Sentiment Indicators on SENT

	BWsent	VIX	Pessimism
(Intercept)	-0.899***	28.99***	3.029***
	(0.000)	(0.000)	(0.000)
SENT_{t-1}	0.074***	-0.547***	-0.015***
	(0.000)	(0.000)	(0.003)
N	510	284	216
$RMSE$	0.796	8.592	0.369
R^2	0.367	0.069	0.04
AIC	1219.227	2031.626	186.714

Notes: Predictive regressions of the sentiment indicators on SENT. BWsent is Baker and Wurgler's sentiment index; VIX is the investor fear gauge described in Whaley (2000); and Pessimism is the media pessimism variable of Tetlock (2007). N is the number of observations in each regression and RMSE is the root mean-squared error. Bootstrapped p-values are listed in parentheses. One, two, and three asterisks represents significance at the 10, 5, and 1 percent levels respectively.

Table 5: Turning Points in the Sentiment Indicators

SENT		BWsent		VIX*(-1)		Pessimism*(-1)	
Peak	Trough	Peak	Trough	Peak	Trough	Peak	Trough
196901	**197312**	196909	197112				
		197303	197709				
198105	**198209**	198203	198305				
198306	**198912**	198406	199108				
				198707	198710	198701	198803
				198906	199010	198904	199011
						199312	199412
						199507	199607
199605	199808	199704	199908	**199401**	**199711**	199705	199711
200002	**200012**	200102	200308	200008	200109	200004	200011
200112	**200209**			200203	200209		
200401	**200812**			200611	200810		

Notes: This table shows the turning points in SENT, Baker and Wurgler's (2006, 2007) sentiment index (BWsent), negative one times the VIX index (VIX*(-1)), and negative one times media pessimism (Pessimism*(-1)). Dates are in the form YYYYMM. In each time period, the series that peaked first is listed in bold face font. Appendix B describes the Bry and Boschan (1971) algorithm.

Table 6: Predictive Regressions of Returns on SENT controlling for the Fama-French Factors and Momentum

Dep Var		SENT_{t-1}
σ	High − Low	-0.004
		(0.591)
Age	Young − Old	-0.009**
		(0.041)
BE/ME	HML	0.011
		(0.193)
BE/ME	Medium − Low	0.008**
		(0.050)
Div	$= 0 - > 0$	-0.018**
		(0.030)
EARN	$\leq 0 - > 0$	-0.044***
		(0.001)
ME	SMB	0.007
		(0.440)
MKT		-0.014
		(0.313)
Mom	High − Medium	0.000
		(0.980)
Mom	Medium − Low	0.020***
		(0.001)

Notes: Predictive regressions of the returns on long-short portfolios on SENT using the regression equation $z_t = \alpha + \beta_1 \text{SENT}_{t-1} + \beta_2 MKT_t + \beta_3 SMB_t + \beta_4 HML_t + \beta_5 UMD_t + \varepsilon_t$ where SENT is the sentiment index and z_t, the dependent variable, is one of the following long-short portfolios: (1) high volatility stocks minus low volatility stocks (σ); (2) young stocks minus old stocks (Age); (3) high value stocks minus low value stocks and medium value stocks minus low value stocks based on book-equity over market-equity (BE/ME); (4) stocks that do not pay dividends less those that do (Div); (5) stocks with earnings less than or equal to zero minus those with positive earnings (EARN); (6) the Fama-French small minus big factor (SMB); (7) excess market returns (MKT); (8) high momentum stocks minus medium momentum stocks and medium momentum stocks less low momentum stocks (MOM). If one of the Fama-French factors or the momentum factor is the dependent variable, we do not include it in the set of regressors. Bootstrapped p-values are listed in parentheses. One, two, and three asterisks represents significance at the 10, 5, and 1 percent levels respectively.

Table 7: Predictive Regressions of Returns on SENT and SENT^{Contr}

Dep Var		SENT$_{t-1}$	SENT$_{t-1}^{Contr}$	$\beta_1 + \beta_2$
σ	High − Low	0.08***	-0.207***	-0.127***
		(0.000)	(0.000)	(0.000)
Age	Young − Old	0.025***	-0.083***	-0.059***
		(0.000)	(0.000)	(0.000)
BE/ME	HML	-0.024***	0.087***	0.063***
		(0.000)	(0.000)	(0.000)
BE/ME	Medium − Low	-0.02***	0.085***	0.065***
		(0.000)	(0.000)	(0.000)
DIV	$=0- >0$	0.064***	-0.2***	-0.135***
		(0.000)	(0.000)	(0.000)
EARN	$\leq 0- >0$	0.046***	-0.208***	-0.162***
		(0.000)	(0.000)	(0.000)
ME	SMB	0.04***	-0.087***	-0.047***
		(0.000)	(0.000)	(0.001)
MKT		0.018***	-0.084***	-0.067***
		(0.000)	(0.000)	(0.001)
Mom	High − Medium	0.027***	-0.071***	-0.044***
		(0.000)	(0.000)	(0.000)
Mom	Medium − Low	0.014***	-0.023***	-0.009
		(0.000)	(0.000)	(0.681)

Notes: Predictive regressions of the returns on long-short portfolios on SENT using the regression equation $z_t = \alpha + \beta_1 \text{SENT}_{t-1} + \beta_2 \text{SENT}_{t-1}^{Contr} + \varepsilon_t$ where SENT is the sentiment index, SENT^{Contr} equals SENT during sentiment contractions and zero otherwise, and z_t, the dependent variable, is one of the following long-short portfolios: (1) high volatility stocks minus low volatility stocks (σ); (2) young stocks minus old stocks (Age); (3) high value stocks minus low value stocks and medium value stocks minus low value stocks based on book-equity over market-equity (BE/ME); (4) stocks that do not pay dividends less those that do (Div); (5) stocks with earnings less than or equal to zero minus those with positive earnings (EARN); (6) the Fama-French small minus big factor (SMB); (7) excess market returns (MKT); (8) high momentum stocks minus medium momentum stocks and medium momentum stocks less low momentum stocks (MOM). If one of the Fama-French factors or the momentum factor is the dependent variable, we do not include it in the set of regressors. Bootstrapped p-values are listed in parentheses. One, two, and three asterisks represents significance at the 10, 5, and 1 percent levels, respectively.

Table 8: Predictive Regressions of Returns on SENT and SENT^{Contr} controlling for the Fama-French Factors and Momentum

Dep Var		SENT_{t-1}	$\text{SENT}_{t-1}^{Contr}$	$\beta_1 + \beta_2$
σ	High − Low	0.038***	-0.098***	-0.06***
		(0.000)	(0.000)	(0.000)
Age	Young − Old	0.012***	-0.055***	-0.043***
		(0.001)	(0.000)	(0.000)
BE/ME	HML	-0.015***	0.060***	0.045***
		(0.000)	(0.000)	(0.000)
BE/ME	Medium − Low	-0.008**	0.052***	0.044***
		(0.016)	(0.000)	(0.000)
DIV	$= 0 - > 0$	0.025***	-0.1***	-0.075***
		(0.000)	(0.000)	(0.000)
EARN	$\leq 0 - > 0$	0.013***	-0.132***	-0.119***
		(0.005)	(0.000)	(0.000)
ME	SMB	0.034***	-0.062***	-0.028**
		(0.000)	(0.000)	(0.038)
MKT		-0.002	-0.026***	-0.029
		(0.440)	(0.000)	(0.140)
Mom	High − Medium	0.012***	-0.025***	-0.013
		(0.004)	(0.000)	(0.146)
Mom	Medium − Low	0.003	0.037***	0.04***
		(0.193)	(0.000)	(0.000)

Notes: Predictive regressions of the returns on long-short portfolios on SENT using the regression equation $z_t = \alpha + \beta_1 \text{SENT}_{t-1} + \beta_2 \text{SENT}_{t-1}^{Contr} + \beta_3 MKT_t + \beta_4 SMB_t + \beta_5 HML_t + \beta_6 UMD_t + \varepsilon_t$ where SENT is the sentiment index, SENT^{Contr} equals SENT during sentiment contractions and zero otherwise, and z_t, the dependent variable, is one of the following long-short portfolios: (1) high volatility stocks minus low volatility stocks (σ); (2) young stocks minus old stocks (Age); (3) high value stocks minus low value stocks and medium value stocks minus low value stocks based on book-equity over market-equity (BE/ME); (4) stocks that do not pay dividends less those that do (Div); (5) stocks with earnings less than or equal to zero minus those with positive earnings (EARN); (6) the Fama-French small minus big factor (SMB); (7) excess market returns (MKT); (8) high momentum stocks minus medium momentum stocks and medium momentum stocks less low momentum stocks (MOM). If one of the Fama-French factors or the momentum factor is the dependent variable, we do not include it in the set of regressors. Bootstrapped p-values are listed in parentheses. One, two, and three asterisks represents significance at the 10, 5, and 1 percent levels respectively.

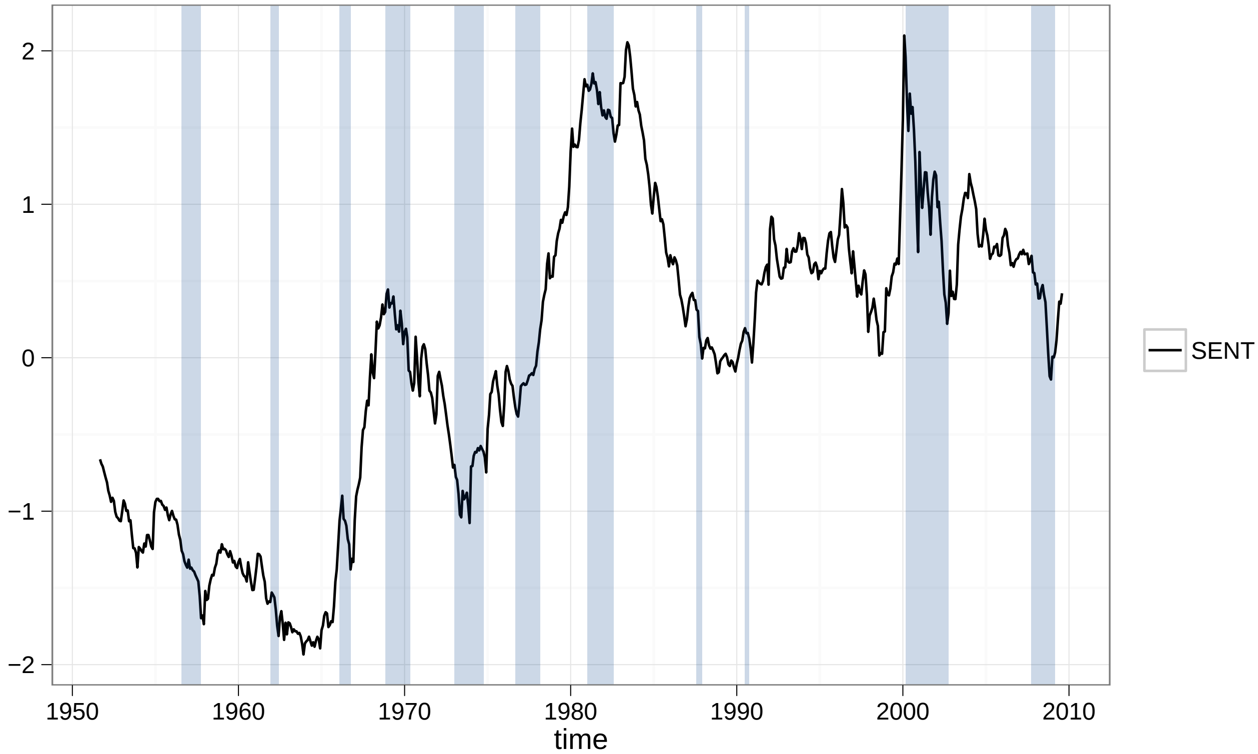
Table 9: Predictive Regressions of Returns on MICH and MICH^{Contr} or BWsent BWsent^{Contr} controlling for the Fama-French Factors and Momentum

		Michigan Consumer Sentiment			Baker and Wurgler's Index		
Dep Var		MICH$_{t-1}$	MICH$^{Contr}_{t-1}$	$\beta_1+\beta_2$	BWsent$_{t-1}$	BWsent$^{Contr}_{t-1}$	$\beta_1+\beta_2$
σ	High – Low	0.259	-0.654*	-0.395	-0.077	-0.491*	-0.569***
		(0.223)	(0.061)	(0.123)	(0.649)	(0.067)	(0.002)
Age	Young – Old	0.222*	-0.155	0.067	-0.029	-0.203	-0.232**
		(0.058)	(0.414)	(0.651)	(0.760)	(0.175)	(0.021)
BE/ME	HML	-0.583***	1.364***	0.781***	-0.226	0.536*	0.31*
		(0.004)	(0.000)	(0.001)	(0.197)	(0.052)	(0.096)
BE/ME	Medium – Low	-0.269	0.902***	0.633***	-0.085	0.421*	0.336**
		(0.110)	(0.001)	(0.002)	(0.548)	(0.058)	(0.024)
DIV	$=0->0$	0.169	-0.77**	-0.6**	0.127	-0.606**	-0.479**
		(0.440)	(0.033)	(0.024)	(0.506)	(0.031)	(0.018)
EARN	$\leq 0->0$	0.406	-1.168**	-0.762**	-0.197	-0.764**	-0.961***
		(0.150)	(0.012)	(0.025)	(0.405)	(0.041)	(0.000)
ME	SMB	-0.278	-0.175	-0.453*	-0.453**	0.393	-0.06
		(0.221)	(0.639)	(0.097)	(0.027)	(0.227)	(0.782)
MKT		-0.471	0.712	0.242	0.237	-0.667	-0.43
		(0.140)	(0.174)	(0.530)	(0.369)	(0.108)	(0.124)
Mom	High – Medium	0.257	-0.252	0.006	0.122	-0.337	-0.215
		(0.101)	(0.328)	(0.976)	(0.364)	(0.111)	(0.130)
Mom	Medium – Low	-0.298*	0.595**	0.297	0.047	0.218	0.265*
		(0.063)	(0.024)	(0.123)	(0.719)	(0.289)	(0.056)

Notes: Predictive regressions of the returns on long-short portfolios on MICH and MICH^{Contr} or BWsent and BWsent^{Contr} controlling for the Fama-French Factors and momentum. MICH^{Contr} equals MICH during sentiment contractions and 0 otherwise, and BWsent^{Contr} equals BWsent during sentiment contractions and 0 otherwise. In the left panel, using Michigan Consumer Sentiment, the regression model is $z_t = \alpha + \beta_1 \text{MICH}_{t-1} + \beta_2 \text{MICH}^{Contr}_{t-1} + \beta_3 MKT_t + \beta_4 SMB_t + \beta_5 HML_t + \beta_6 UMD_t + \varepsilon_t$ where MICH is the Michigan Consumer Sentiment index orthogonalized to various macroeconomic indicators. In the right panel, using Baker and Wurgler's sentiment index, the regression model is $z_t = \alpha + \beta_1 \text{BWsent}_{t-1} + \beta_2 \text{BWsent}^{Contr}_{t-1} + \beta_4 MKT_t + \beta_5 SMB_t + \beta_6 HML_t + \beta_7 UMD_t + \varepsilon_t$. In both panels, the dependent variable, z_t, is one of the following long-short portfolios: (1) high volatility stocks minus low volatility stocks (σ); (2) young stocks minus old stocks (Age); (3) high value stocks minus low value stocks and medium value stocks minus low value stocks based on book-equity over market-equity (BE/ME); (4) stocks that do not pay dividends less those that do (Div); (5) stocks with earnings less than or equal to zero minus those with positive earnings (EARN); (6) the Fama-French small minus big factor (SMB); (7) excess market returns (MKT); (8) high momentum stocks minus medium momentum stocks and medium momentum stocks less low momentum stocks (MOM). If one of the Fama-French factors or the momentum factor is the dependent variable, we do not include it in the set of regressors. Bootstrapped p-values are listed in parentheses. One, two, and three asterisks represents significance at the 10, 5, and 1 percent levels respectively.

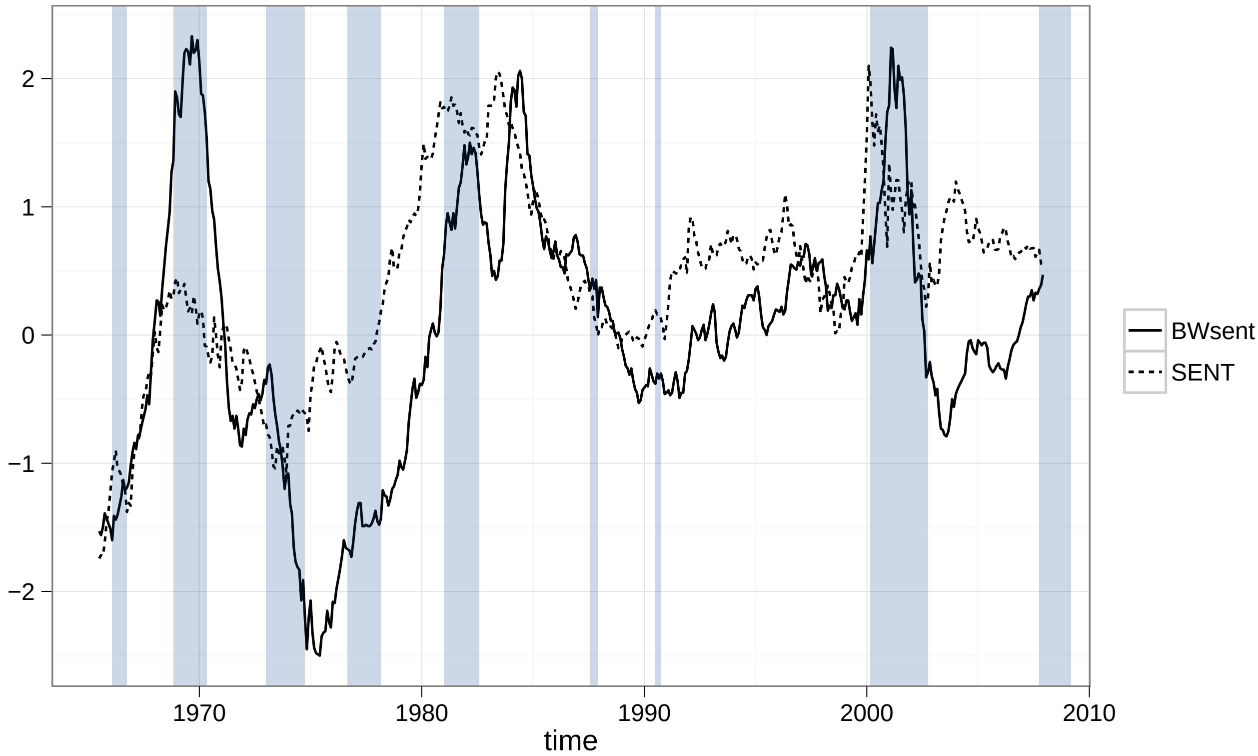
10 Figures

Figure 1: The Sentiment Index (SENT)



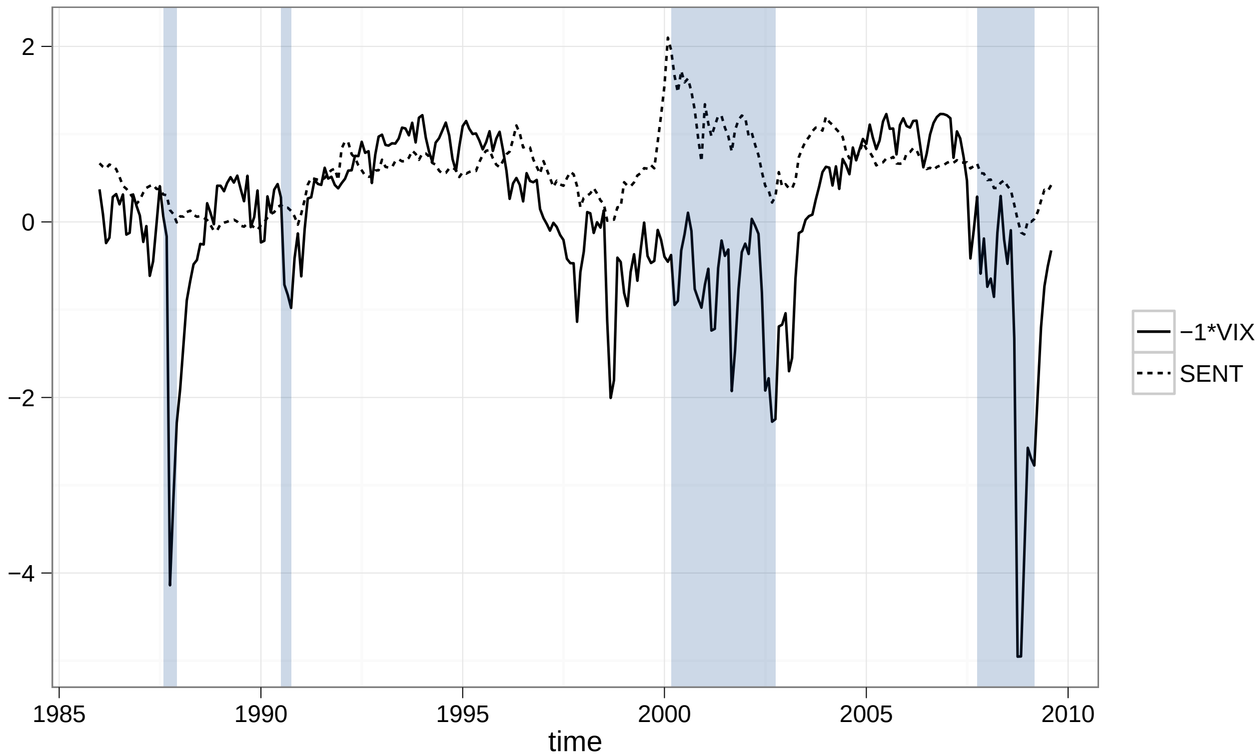
Notes: The sentiment index, SENT. High values of SENT correspond to high levels of agent sentiment. The shaded vertical bars represent bear markets defined as 20 percent or more drop in the S&P 500 over a period of two or more months.

Figure 2: BWsent vs. SENT



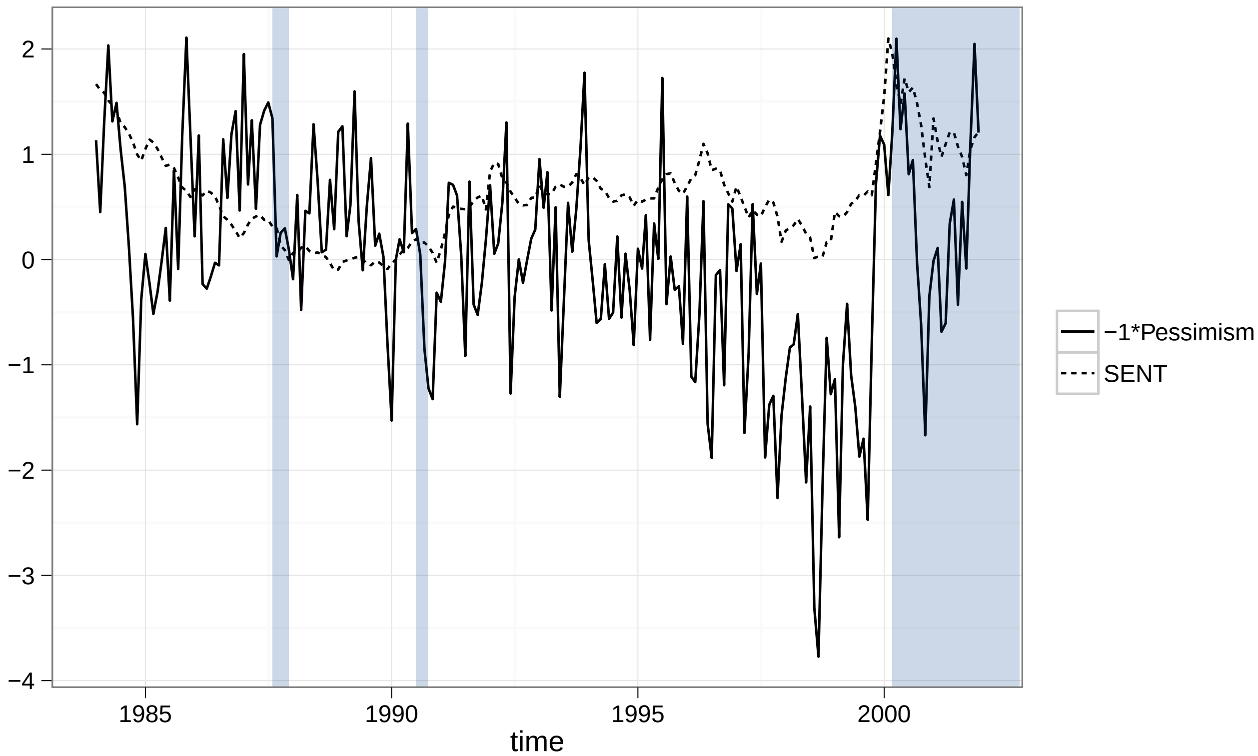
Notes: Shaded areas are bear markets defined by a 20 percent or more drop in the S&P 500 over a period of two or more months. The solid line is BWsent; SENT is the dotted line. We standardize both series to have zero mean and unit variance.

Figure 3: SENT and VIX*(-1)



Notes: SENT and VIX*(−1). We standardize both series to have zero mean and unit variance. The solid line is VIX*(−1); the dashed line is SENT. Shaded areas are bear markets.

Figure 4: SENT and Pessimism



Notes: SENT and Pessimism*(−1). We standardize both series to have zero mean and unit variance. The solid line is Pessimism*(−1); the dashed line is SENT. Shaded areas are bear markets.

A Appendix: The Dynamic Factor Model

In this section we describe the dynamic factor model. Like principal component analysis (PCA), dynamic factor models extract a component from a set of time series. Unlike PCA, dynamic factor models allow us to compile a factor in levels from differenced series. Let r_{it} stand for each the orthogonalized return series, Div, Earn, Size, and Lowmom. We specify the model as follows:

$$r_{it} = \gamma_i \Delta c_t + \varepsilon_{it}, i = 1, \ldots, 5 \tag{7}$$

$$\phi(L)\Delta c_t = \omega_t, \omega_t \sim N(0, 1) \tag{8}$$

$$\psi_i(L)\varepsilon_{it} = v_{it}, v_{it} \sim N(0, \sigma_i^2) \tag{9}$$

$\Delta c_t = \Delta C_t - \delta$ is the component common to all series, ε_{it} is an idiosyncratic component, γ_i is the factor loading, and L is the lag operator. To derive the unobserved common component representing the levels index, we need to identify δ. Our approach follows Stock and Watson (1991) who noticed that ΔC_t is a function of past lags of ΔY_t. Stock and Watson then derive an estimate of δ by taking the expected value of ΔC_t. Once we have an estimate for δ, we can compile C_t. C_t is the common component representing the levels sentiment index (SENT).

We place the model into state-space form and then estimate the model using the Bayesian multimove Gibbs-sampling approach based on Carter and Kohn (1994) and Kim and Nelson (1998). The Bayesian method takes into account parameter uncertainty by jointly estimating the state vector and the model parameters.

To implement the estimation algorithm, we use the MCMC Gibbs-sampling method. We run the algorithm 10,000 times and drop the first 2000 iterations. Using the Bayesian Information Criterion (BIC), we choose two lags for $\phi(L)$ and two lags for each $\psi_i(L)$. For a further explanation of these techniques, see Kim and Nelson (1998) or Harvey, Koopman and Shephard (2004).

Table 10 shows parameter estimates. First, the coefficients on the lags of Δc_t are positive. This suggests that sentiment is positively autocorrelated with its own past values.

Second, the factor loadings, $\gamma_1, \ldots, \gamma_4$, all have the expected positive sign. Hence, Δc_t is positively related to all of its components. Third, in the idiosyncratic equations, the coefficients on ψ_{i1} and ψ_{i2} are all small in magnitude. This implies that common component captures most of the dynamics in the index components, Div, Earn, Lowmom, and Size. Lastly, the estimated variance for Lowmom, σ_3^2, is much larger in magnitude than the variance estimated for the other parameters. This corresponds with our expectations; as evinced in table 1, the standard deviation of Lowmom is approximately twice as large as that estimated for the other variables.

Table 10: Parameter Estimates from the Dynamic Factor for SENT

	Mean	Std	Median
	Panel 1: ϕ		
ϕ_1	0.210	0.014	0.210
ϕ_2	0.009	0.015	0.009
	Panel 2: Div		
γ_1	3.648	0.096	3.646
ψ_{11}	0.007	0.055	0.007
ψ_{12}	0.001	0.058	0.000
σ_1^2	0.544	0.125	0.535
	Panel 3: Earn		
γ_2	3.757	0.120	3.760
ψ_{21}	-0.158	0.042	-0.158
ψ_{22}	0.035	0.043	0.033
σ_2^2	4.707	0.292	4.695
	Panel 4: Lowmom		
γ_3	6.167	0.231	6.159
ψ_{31}	0.158	0.040	0.157
ψ_{32}	-0.092	0.041	-0.091
σ_3^2	23.789	1.434	23.772
	Panel 5: Size		
γ_4	3.202	0.122	3.201
ψ_{41}	0.028	0.040	0.028
ψ_{42}	0.061	0.039	0.061
σ_4^2	6.289	0.367	6.283
	Panel 6: δ		
δ	0.015	0.003	0.015

Notes: Parameter estimates from the dynamic factor model outlined in appendix A. Panel 1 shows the coefficient estimates on the lags of Δc_t in equation 8. Panels 2, 3, 4, and 5 show the factor loadings, the γ_i's, and the coefficient estimates from the idiosyncratic equation outlined in equation 9 for Div, Earn, Lowmom, or Size, respectively. Panel 6 shows the estimate of δ. For each parameter, we show the mean, standard deviation, and median of the estimates.

B Appendix: Procedure For Determination of Turning Points Using the Bry and Boschan Algorithm

1. Filter the series using the HP Filter with the smoothing parameter set to 150.

2. Remove outliers from the data and replace them using the Spencer curve.

3. Determination of initial turning points in raw data:

   (a) Determination of initial turning points in raw data by choosing local peaks (troughs) as occurring when they are the highest (lowest) values in a window 12 months on either side of the date.

   (b) Enforcement of alternation of turns by selecting highest of multiple peaks (or lowest of multiple troughs).

4. Censoring operations (ensure alternation after each):

   (a) Eliminate turns within 6 months of the beginning and end of the series.

   (b) Eliminate peaks (or troughs) at both ends of the series that are lower (or higher) than values closer to the end.

   (c) Eliminate cycles whose duration are less than 24 months.

   (d) Eliminate phases whose duration are less than 4 months or whose magnitude are smaller than 1.5 standard deviations of the sentiment index.

5. State final turning points.