Measuring investor sentiment in the stock market

Francisca BEER
California State University of San Bernardino

Mohamed ZOUAOUI♦
University of Franche-Comté

FARGO - Centre de recherche en Finance, ARchitecture et Gouvernance des Organisations

Cahier du FARGO n° 1110901
First Draft - September 28, 2011

Abstract: Recently, the investor sentiment measure has become one of the most widely studied areas in behavioral finance. Various measures have been developed in the literature without being to determine which measure should be used. The purpose of this study is to test their relative performance in predicting stock returns. Using a panel of investor sentiment measures, we develop a new measure of sentiment which combines direct and indirect sentiment measures. Our results show that our composite sentiment index affects the returns of stocks hard to value and difficult to arbitrage consistent with the predictions of noise trader's models. Finally, we find that our composite index has a better predictive ability than the alternative sentiment measures largely used in the literature.

Keywords: sentiment measures, composite index, stock returns.

JEL Classification: G12, G14.

Contact : Mohamed ZOUAOUI, Mohamed.Zouaoui@univ-fcomte.fr, Tel: 0033 381 666 642

♦ We thank Gregory W. Brown, Yaw-Hei Wang, Jeffrey Wurgler as well as Denys Glushkov for generously providing some of the data. All errors are the sole responsibility of the authors.

1. Introduction

Now, the question is no longer, as it was a few decades ago, whether investor sentiment affects stock prices, but rather how to measure investor sentiment and quantify its effects.

Malcolm Baker and Jeffrey Wurgler

In recent years, the usefulness of an investor sentiment measure to predict stock market returns has been the subject of frequent inquiries. A vast number of empirical investigations with different measures of investor sentiment have been conducted. While theoretical models have early incorporated the existence of noise traders into equilibrium asset pricing (Black, 1986; De Long, Shleifer, Summers and Waldmann, 1990), empirical evidence on the correct proxy to quantify sentiment does not provide clear findings.

Several empirical studies have attempted to measure investor sentiment. These studies identified direct and indirect sentiment measures[1]. Direct sentiment measures are derived from surveys directly asking individuals how they feel about current or future economic and stock market conditions while indirect ones represent economic and financial variables susceptible to capture investors' state of mind.

Numerous significant publications focus on the impact of direct sentiment measures on stock returns. Unfortunately, the results appear to be sensitive to the measure used. Solt and Statman [1988] and Clarke and Statman [1998] point out that investor sentiment compiled by Investors Intelligence survey is not useful as a contrary indicator. They report no statistically significant relationship between the sentiment index and subsequent stock returns. Contrary to previous findings, Brown and Cliff [2005] show that excessive optimism leads to periods of market overvaluation and high current sentiment is followed by low cumulative long-run returns. De Bondt [1993] finds that individual investors surveyed by the American Association of Individual

Investors forecast future stock returns. Other studies focusing on indexes of consumer confidence analyzed the impact of sentiment on the stock market. Otoo [1999] reports a strong contemporaneous relationship between the changes of consumer confidence index and the stock returns. However, by examining the causal relation among the variables, she finds that returns Granger-cause consumer confidence at very short horizons but not vice versa. In another study, Lemmon and Portniaguina [2006] find that sentiment index as proxied by consumer confidence is able to forecast the returns of small stocks and those with low institutional ownership. More recently, Schmeling [2009] and Zouaoui, Nouyrigat and Beer [2011] confirm this finding in an international context.

The benefits of using surveys to measure sentiment are significant. Surveys take into account the psychological dimension of individuals (optimism, pessimism and neutrality) in accordance with their socioeconomic characteristics. They use standardized questions making measurement more precise and permit to have large and regular time series. Surveys also provide information about investors' state of mind even without sophisticated financial theory to validate them. Nevertheless, surveys also have boundaries. First, sample size is often limited. Further, the so-called "prestige bias", i.e. the tendency for respondents to answer in a way that make them feel better, often impact the survey results. Sometimes even more pervasive, since most survey opinions are gradually submitted throughout a week or a month period, the results do not correspond to investor sentiment during a given point-in-time, but to a mix of recent and old opinions. Other potential problem is that survey responses are weighted equally regardless of the magnitude of funds managed by respondents. Finally, with the exception of the survey UBS/Gallup, no distinction is made between the different levels of optimism or pessimism. These limitations explain why other papers suggest using economic and market variables as implicit sentiment proxies instead.

Neal and Wheatley [1998] examine the forecast power of three popular measures of investor sentiment: the level of discount on closed-end fund, the ratio of odd-lot sales to purchases and the net mutual fund redemptions. They find that closed-end fund discount and net redemptions predict the size premium better than the odd-lot ratio. Brown and Cliff [2004] scrutinize various direct and indirect sentiment indicators. They report that direct (surveys) and indirect measures of sentiment are correlated. Although indicators of sentiment strongly correlate with contemporaneous market returns, they show that sentiment has little predictive power for near-term future stock returns. Qiu and Welch [2006] find no association between the discount of closed-end fund and the UBS/Gallup sentiment survey. However, they show that the consumer confidence index is highly correlated to UBS/Gallup sentiment survey. Therefore, they recommend the use of the confidence index instead of the closed-end fund discount as a sentiment measure.

Baker and Wurgler [2006] construct a sentiment index as a linear combination of six indirect measures, namely the closed-end fund discount, the logarithm of the NYSE share turnover ratio, the number of IPOs, the average first-day returns on IPOs, the share of equity issues in total equity and debt issues and the dividend premium defined as the log difference of the average market-to-book ratios of dividend payers and non payers. They show that stocks difficult to value and to arbitrage react more strongly to investor sentiment than the other categories of stocks with opposite characteristics.

Indirect measures have a tremendous advantage over direct one, they are relatively easy to construct as they are based on simple market data. Indirect indicators are observed in real time and reflect both the power of market participant and the strength of their bullishness or bearishness. However, using economic and financial variables as investor sentiment measure also raise difficulties. First, some indirect indicators rely on controversial theoretical explanations[2].

Second, the indicators are endogenous to the market and economic activity, so they may not measure exclusively investor sentiment. These indicators are an amalgam of the economic expectations and sentiment and the process of isolating one from the other could prove difficult if not impossible.

Which investor sentiment indicator should be used? To date, the behavioral models are still silent as to what indicator should be used to assess their validity. Indeed, although numerous studies on the issue of investor sentiment have been published, relatively little research has focused on their relative efficacy in predicting future stocks returns. Further, very few studies are identified and those are limited to compare direct with indirect indicators[3] (See Qiu and Welch, 2006; Feldman, 2010).

The purpose of this article is to test which of these measures is the most pertinent predictor of future stock returns. Each indicator presents advantages and limitations. Furthermore, each individual indicator could measure sentiment at a specific point in the market cycle and no during the sample period[4]. These considerations induce us to consider that the best empirical approach is to condense several imperfect indicators into an aggregate index. As a result, we build a new measure of sentiment by combining several well-known direct and indirect sentiment indicators. Specifically, we split the raw sentiment indicator into a rational component related to economic fundamentals and a psychological component linked to investor sentiment. Using a panel of investor sentiment measures, we develop a new composite sentiment indicator and a new composite fundamental indicator. We subsequently incorporate these two indicators in a model studying simultaneously the impact of economic fundamentals and the impact of investor sentiment on the stock returns.

The analysis yields three important results. First, we find that our composite sentiment index produces a faithful reproduction of the bubbles and crashes during our study period, i.e. July

1981 to December 2010. For instance, our composite index records a significant decline when the speculative bubble of October 1987 burst. Between 1998 and 2003, it peaks in March 2000 at the beginning of the internet bubble. During the period studied, our indicator is at its lowest level during the recent sub-prime crisis.

Second, we show that the rational approach can be advantageously completed by the behavioral approach. In agreement with the classic theory, we report a significantly relationship between the investors' rational expectations and the stock returns. Specifically, we find that small stocks, value stocks, young stocks, unprofitable stocks and intangible stocks are less (more) vulnerable to the business cycle after periods characterized by good (bad) economic expectations. In agreement with the behavioral approach, we find that investor sentiment impact stocks hard to value and difficult to arbitrage even after controlling for economic fundamentals and structural risks factors. Specifically, these stocks earn low (high) returns following periods when investor sentiment is high (low).

Third, we find that our measure outperforms the other popular measures routinely used in the literature (direct, indirect and other composite measures) in predicting returns of long-short portfolios based on different characteristics designed to reflect the stocks hard to value and difficult to arbitrage. This finding suggests that the strength of our composite measure comes from the simultaneous effect of the combination of both direct and indirect indicators.

The remainder of the paper proceeds as follows. The second section presents the data used to study the impact of investor sentiment on the stock returns. This section also develops our composite indicators and builds the long-short portfolios. The third section presents the methodology and analyses the empirical results obtained. The fourth section compares our composite index to other sentiment indicators usually used in the literature. The fifth section concludes the study.

2. Data

The data used in this study were collected from several sources. Our study includes financial and macro-economic variables and survey results. Table 1 presents the list of variables and sources of data used.

a. Fundamental and sentiment indicator

Investor sentiment can be defined as a belief about future cash flows and investment risks that is not warranted by fundamentals. Each individual sentiment proxy is likely to include a sentiment component as well as idiosyncratic component that are unrelated to sentiment. To circumvent this problem, we use principal components analysis to remove these idiosyncratic components by combining them together into an aggregate index. Therefore, given that there is no uncontroversial and universally accepted sentiment measure, we construct a synthetic indicator which combines direct and indirect sentiment measures. The strength of the composite indicators is that they take into account multiple sources of information. Hence, a composite indicator should better reflect the changes of investors' sentiment than any measure used individually[5].

Our study focuses on two direct sentiment indicators and four indirect sentiment indicators. This choice is supported by the following three facts. First, our selection is the result of the established relationship between each selected sentiment index and the equity market. Prior studies show that each sentiment index seizes some of the stock market aspects not already contained in traditional macro-economic indicators. Second, although both direct and indirect measures of investor sentiment have been employed in previous studies, it remains unclear which measures are actually the most appropriate, and to which extent they represent the same informational content. We find it reasonable to infer that combining several imperfect measures

using both direct and indirect will lead to a better one. Third, data availability narrows the options; some indicators are only available over short time periods.

An aggregate index is constructed from a principal component analysis (first component) of six sentiment measures identified in previous studies: University of Michigan consumer confidence index (UMI), the Investors Intelligence spread Bull-Bear (II), the number of IPOs in a given month (NIPO), the average monthly first-day returns on IPOs (RIPO), the net new cash flows of US equity mutual funds (FLOW) and finally the closed-end fund discount (CEFD). All proxies are measured monthly over the period from June 1981 to December 2010.

Panel A of Table 2 reports the correlations between the raw sentiment indicators. We note that direct sentiment measures are highly correlated with indirect measures. This result sheds some light on the fact that certain indirect measures have acquired a solid reputation as sentiment measures. The II index is positively correlated with FLOW, RIPO and NIPO and negatively correlated with CEFD. When II index increases, CEFD drops indicating that investors are more optimistic. We also find a significant negative correlation between CEFD and FLOW; large discounts reduce the magnitude of FLOW. Consistent with the hypothesis which states that IPOs occur during periods of high investor optimism, we note a significant correlation between UMI and the variables RIPO and NIPO. Finally, we observe non-significant correlations between UMI and the variables CEFD and FLOW. Thus, the correlation between the indicators is not perfect; this underlines the benefit of extracting the common component of the indicators.

The raw sentiment indicators encompass a psychological component related to sentiment and a rational component related to economic fundamentals. As noted by several studies[6], investors' sentiment varies in part for entirely rational reasons related to the macroeconomic conditions. When an investor is bullish or bearish, this could be a rational reflection of future period's expectation or irrational enthusiasm or a combination of both. To isolate sentiment consists

precisely of identifying investors' optimism (pessimism) although there is not a good (bad) economic reason for being so. To separate these two aspects, it is necessary to split the raw sentiment indicators in two components: a rational one, reflecting the economic fundamentals and a psychological component, reflecting investor sentiment. To this end, we regress the raw sentiment indicators on contemporaneous variables that capture the component related to the business cycle, i.e. macroeconomic "fundamentals" and we treat the residuals from the regression as our purer measure of sentiment unwarranted by fundamentals:

$$SENT_{i,t} = \alpha_i + \beta_i \sum_{i=1}^{I} FUND_{i,t} + \varepsilon_{i,t} \qquad (1)$$

Where $SENT_i$ is one of the raw sentiment variables, α_i is the constant and β_i are the parameters to be estimated. $FUND_i$ is the set of fundamental variables representing rational expectations. These variables include growth of industrial production (IP), inflation (INF), term spread (TS), default spread (DS), and growth in durable (DC), nondurable (NDC) and services consumption (SC). Similar to the literature, we assume that the information conveyed by these variables is quite complete and sufficient to capture the economic rational expectations. The fitted values of Equation 1 capture the rational component and the residual capture the psychological component.

Panel B of Table 2 reports the correlations between the raw sentiment indicators and the macroeconomic variables. The largest coefficient of correlation is between University of Michigan consumer confidence index and the default spread (-58.3%). The smallest coefficient is between the Investors Intelligence index and the growth of non-durable goods (-0.2%). The range of these coefficients shows that the macroeconomic variables reflecting rational anticipations can explain a trivial part or a large part of the raw sentiment indicators. The adjusted R^2 range from 7.4 to 39.1%.

In constructing our synthetic sentiment index, we follow closely the methodology of Baker and Wurgler [2006]. We start by estimating the first principal components of six current (*t*) investor sentiment proxies and their lags (*t-1*)[7]. This gives a first-stage index with six loadings for the current proxies and six loadings for the lagged proxies. Finally, our sentiment index represents the first principal component of six proxies (each proxy or its lag) most highly correlated with the first stage-index. This procedure leads to the following composite sentiment index:

$$\begin{aligned} CSI_t &= 0.208\,UMI_t^{\perp} + 0.191\,II_{t-1}^{\perp} + 0.204\,NIPO_t^{\perp} \\ &\quad + 0.183\,RIPO_{t-1}^{\perp} + 0.234\,FLOW_{t-1}^{\perp} - 0.211\,CEFD_{t-1}^{\perp} \end{aligned} \qquad (2)$$

The first principal component explains about 52% of the sample variance in the macro-adjusted sentiment proxies[8]. We can see that all individual sentiment measures obtain a similar weight (around 0.2) within the overall equation for the CSI. We find that the coefficients of the sentiment indicators have all the expected signs. They are positive for the survey data, the variables related to IPOs and mutual fund flows. The negative sign on the closed-end fund discount is consistent with the interpretation of investor sentiment, the greater the discount, the more investors are bearish.

The same procedure is applied to fundamental proxies in order to extract a composite fundamental index. The first principal component explains about 42% of the sample variance. We also find that all individual fundamental measures have similar weight and display the expected sign. The composite fundamental index (*CFI)* is as follows:

$$\begin{aligned} CFI_t &= 0.134 UMI_t^{*} + 0.151 II_{t-1}^{*} + 0.121 NIPO_t^{*} \\ &\quad + 0.101 RIPO_{t-1}^{*} + 0.198 FLOW_{t-1}^{*} - 0.118 CEFD_{t-1}^{*} \end{aligned} \qquad (3)$$

Does our composite index capture the fluctuations in sentiment? Unfortunately, there is not yet a method to evaluate the correctness of a specific proxy for sentiment. Baker and Wrugler

[2007] recommend an eyeball test as the best method to validate that the indicator adequately reflects investor sentiment. They state that if a proxy succeeds in capturing sentiment, it lines up with recognized bubbles and crashes during the period analysed.

Figure 1 illustrates the evolution of the sentiment composite indicator during the period from July 1981 to December 2010. The sentiment indicator drops sharply in the year 1987, it reaches its lowest level in November 1987. This situation coincides with the market crash of October 19, 1987. Significant decreases are also seen during the collapse of the bonds market in 1994 and during the collapse of LTCM in 1998. Moreover, we note an increase of the composite index at the peak of the market in 2000. Over the period 1998-2003, the composite index reaches its highest level in March 2000[9]. This date coincides with the peak of the Dot.com. The composite sentiment index starts to decrease in April 2000; this decline accelerates after the attacks of September 11, 2001. As anticipated, the index also shows a large decrease in 2008 during the sub-prime crisis. Overall, the composite sentiment index produces a faithful reproduction of the bubbles and crashes during our study period.

b. Firm characteristics

Our sample includes all common stocks (share codes 10 and 11) currently or formerly listed on the NYSE, AMEX and NASDAQ between July 1981 and December 2010, thus avoiding survivorship bias. Stock returns and firm characteristics are from the merged CRSP-Compustat database. We form long-short portfolios based on characteristics designed to reflect the stocks that are hard to value and difficult to arbitrage. Following Baker and Wurgler [2006, 2007], the firm characteristics considered are: size, firm's growth potential and distress, age, profitability, dividend policy, tangibility and arbitrage costs.

Size is the market capitalization measured as price time shares outstanding from CRSP. Firm's growth potential and distress characteristic include market-to-book computed as the market value at the end of the calendar year divided by the book value reported anytime during the fiscal year *t*. Age is the number of months since the firm's first appearance on the CRSP tapes. Profitability is captured by the return on assets defined as the earnings divided by total assets. Asset tangibility is captured by property, plant and equipment over total assets. Dividend policy is dividends per share at the ex date multiplied by compustat shares outstanding divided by book equity. Arbitrage costs are measured by idiosyncratic volatility measured by the standard deviation of residuals (over 60 months preceding month *t*) in the regression of individual stock returns on Fama and French (1993) risk factors[10].

We construct long-short portfolios based on each of firm characteristics, using different *k* time horizons (1, 6, 9 and 12 months). In particular, we use the calendar approach of Jegadeesh and Titman [2001] to minimize the bias of autocorrelation arising from the construction of event-time portfolios over multi-period horizons. Each month, we form three equally weighed portfolios by sorting the firms on each corresponding characteristic[11]. The bottom (top) one-third ranked stocks of each characteristic are designated as a Low (High) portfolio. We then calculate in every month of horizon *k* the difference between the average returns of the constructed extreme portfolios *(Low-high*). Finally, the holding period return for long-short portfolios is the average monthly return over the previous *k* months of the *k* differential portfolios estimated for that month.

3. Methodology and empirical results

We use the model of Lemmon and Portniguina [2006] to test the impact of investor sentiment on the stock returns. This model offers a considerable benefit; it allows studying simultaneously

the impact of economic fundamentals (estimated using the raw sentiment index) and the impact of investor sentiment on stock returns. For this reason, it is really useful to examine the contribution of behavioral finance theory to the classical finance theory. The empirical model used is of the following form:

$$R^{j}_{(Low),t} - R^{j}_{(High),t} = (\alpha_1 + \alpha_2 CSI_{t-1}) + (\beta_1 + \beta_2 CFI_{t-1})RMRF_t + \varepsilon^{j}_{t} \qquad (4)$$

Where $R^{j}_{(Low),t}$-$R^{j}_{(High),t}$ is one-month holding period return[12] for long-short portfolios based on characteristics *j*. Low is defined as a firm in the bottom three CRSP deciles and High is defined as a firm in the top three CRSP deciles. CSI_{t-1} and CFI_{t-1} are the lagged values of the composite sentiment index and the fundamental composite index, respectively. $RMRF_{,t}$ is the excess return on the CRSP value-weighted index for month *t*.

Lemmon and Portniguina's model extends the CAPM by integrating both rational component reflecting the economic fundamentals and psychological component reflecting investor sentiment. This model allows the conditional market beta to be a function of economic fundamentals and allow the pricing error to depend on investor sentiment.

Under the rational hypothesis, the time variation in the expected returns of the long-short portfolios is related to the economic fundamentals, as investors rationally forecast future macroeconomic conditions. There is now substantial evidence that firms on the fringe, and firms with a higher likelihood of financial distress, are more sensitive to changes in the business cycle (Chan and Chen, 1991; Jagannathan and Wang, 1996). The conditional market betas of these firms should increase (decrease) following periods of low (high) expectations about economic fundamentals. Therefore, we expect that the coefficient β_2 will be significantly less than zero.

Under the behavioral hypothesis, the time variation in the expected returns of the long-short portfolios is affected by investor sentiment if the coefficient α_2 is significantly less than zero. The

proposed explanation is that investors tend to overvalue stocks that are difficult to arbitrage and hard to value when they are optimistic, and to undervalue these stocks when they are pessimistic. However, as every mispricing is eventually corrected, excessive optimism is inevitably followed by a decrease in stock prices when markets return to fundamentals. Therefore, these difficult to arbitrage and hard to value stocks earn low (high) returns following periods when investor sentiment is high (low). The opposite outcome should be observed during period of low investor sentiment.

The regression results from the model (4) are presented in Table 3. For panel A, the empirical results support both hypotheses. The coefficients estimates α_2 and β_2 are negative and statistically significant at conventional levels. A negative significant β_2 coefficient confirms that small stocks are more (less) vulnerable to change in the business cycles after periods of pessimistic (optimistic) economic prospects. This established finding supports the classical finance theory that small stocks fluctuate more with business cycle and have higher risk exposure to the changing risk premium. Further, consistent with the behavioral hypothesis, the significant negative α_2 coefficient shows a relationship between the lagged sentiment variable and the size premium. This result provides supports to that of Lemmon and Portniguina [2006] and reinforces the behavioral finance theory that establishes the overvaluation (undervaluation) of small cap relative to large cap when investors are bullish (bearish).

Similar results are summarized in panel B. The coefficients estimates α_2 and β_2 validate both the rational and behavioral hypotheses. The negative sign of the coefficient β_2 corroborates the main argument of the classical finance theory that the market-to-book ratio shows how vulnerable value stocks are to financial distress. Furthermore, the value premium is negatively correlated with the lagged psychological factor. Value stocks show higher (lower) expected returns than growth stocks after periods characterized by excessive investor pessimism (optimism). This result

reflects the undervaluation (over-valuation) of value stocks relative to growth stocks when investors are bearish (bullish).

Baker and Wurgler [2006] find a U-shaped relationship between investor sentiment and the returns on portfolios sorted by book-to-market. Specifically, they show that both extreme value and growth stocks are more sensitive to investor sentiment than more established firms. These stocks are performing better (worse) than those with a medium book-to-market ratio when sentiment is low (high). To compare our results to those reported by Baker and Wrugler [2006], we re-estimate model (4) by sequentially replacing the dependent variable with the excess returns of value stocks and with the excess returns of growth stocks. Our results[13] only partially mirror those of Baker and Wurgler [2006]. We find evidence that only value stocks significantly react to investor sentiment. Specifically, we find that the excess returns on value stocks are higher after periods of low sentiment and lower following periods of high sentiment. Although our results only partially mirror those of Baker and Wurgler [2006], they are similar to those of Kumar and Lee [2006]. These authors find that individual investors tend to overweight their portfolios with value stocks. They also show that while there was a strong correlation between investor sentiment and the returns of value stocks, the relationship between investor sentiment and growth stocks is weak.

In panels C, D and E, the returns of the long-short portfolios are significantly and negatively correlated with the lagged economic fundamentals. The coefficients β_2 are also negative although not significant in panels F and G. In all the panels, evidence indicates a significant negative relationship between the sentiment factor and the future portfolio returns. This last result supports the hypothesis that investor sentiment is an important factor in the return generating process of common stocks. As shown in table 3, the incremental adjusted R^2 due to the addition of the variable sentiment range from 3.1 to 5.9%.

To assess the robustness of our results, we run regressions in which we allow the market betas to vary directly with other fundamental variables usually used in the literature to predict stock returns. In addition to the composite fundamental index, we include the dividend yield on the market (DIV), the default spread (DS) and the three-month treasury rate (INT). We also include the premium on a portfolio of small stocks relative to large stocks (SMB), the premium on a portfolio of high book-to-market stocks relative to low book-to-market stocks (HML) and the premium on a portfolio of stocks with high past returns relative to stocks with low past returns (UMD)[14], as control variables. To do so, we run the following model[15]:

$$R^{j}_{(Low),t} - R^{j}_{(High),t} = (\alpha_1 + \alpha_2 CSI_{t-1}) + (\beta_1 + \beta_2 CFI_{t-1} + \beta_3 DIV_{t-1} + \beta_4 DS_{t-1} + \beta_5 INT_{t-1}) RMRF_t + \delta_1 SMB_t + \delta_2 HML_t + \delta_3 UMD_t + \varepsilon^{j}_{t} \quad (5)$$

The results in table 4 are qualitatively identical to those reported previously, although the statistical significance of the variable CSI is slightly weaker. Except for the panel F, the results show that the variable CSI remains significant with the expected negative sign even after controlling for traditional variables and structural risk factor. This results suggest that the predictive power of the variable CSI is unrelated to time-varying expected returns or economic cycles. Due to the addition of the variable sentiment, the incremental adjusted R^2 increases from 2.9 to 5.1%[16]. Additionally, the large majority of the coefficients estimates for the indicator CFI are negative and significant at the conventional level. This finding is thought-provoking and suggests that our aggregate fundamental index is a better candidate to forecast stock returns than other fundamental variables usually used in the literature.

Overall, the rational and behavioral hypotheses are both validated by our results[17]. We find that small stocks, value stocks, young stocks, unprofitable stocks and intangible stocks are less (more) vulnerable to the business cycle after periods characterized by good (bad) economic

expectations. Our results also show that periods characterized by excessive investors' optimism (pessimism) are followed by low (high) returns for stocks difficult-to-value and hard-to-arbitrage.

4. Comparing the relevance of sentiment indicators

To measure investor sentiment we developed a composite index that combines direct and indirect sentiment indicators. This index was preferred to using exclusively a direct or indirect sentiment measure. Is our composite indicator is more relevance than the alternative measures used in previous studies? This section scrutinizes the relevance of our synthetic sentiment index.

One way to operate is to compare the results obtained in the previous section with those obtained on the basis of an alternative sentiment measure. To do so, we re-estimate the model (4) using some popular sentiment measures usually used in the literature and compare findings with those reported above. Specifically, we contrast our indicators with a set of direct, indirect and composite indicators. For the sake of brevity, we present only the coefficient estimates (β_2) and (α_2) respectively assessing the rational and the behavioral hypotheses.

- **The Composite Sentiment Index vs. direct measures**

First, we compare our composite sentiment index to each of the direct measures used in its construction, i.e.: the University of Michigan consumer confidence index (UMI) and the Investors Intelligence index (II). Second, we compare our indicator with the Conference Board survey of consumer confidence (CBIND) used by Otoo [1999] and Lemmon and Portniguina [2006]. Finally, we compare our composite indicator with that of the American Association of Individual Investors (AAII) used by Brown [1999] and Fisher and Statman [2003][18].

The results depicted in Table 5 show that almost all the coefficients estimates α_2 are not significant. Among the 28 coefficients estimates α_2, there is only three significant coefficients. Results indicate that the II index forecasts the size premium and the value premium and that AAII

indicator predicts only the size premium. We also observe that the two consumer sentiment index (UMI and CBIND) do not significantly impact the stock returns. These results diverge from those reported by Lemmon and Portniguina [2006] who report a significant negative relationship between the consumer sentiment index and the size premium in the following 3-month. Moreover, we note the low correlation between the rational expectations from each sentiment indicator and the future stock returns. Indeed, only the rational component from both consumer sentiment indexes predicts the value premium.

- **The Composite Sentiment Index vs. indirect measures**

We first compare the predictive power of our composite index with each indirect measure used in its construction, i.e.: the number of IPOs (NIPO), the average monthly first-day returns on IPOs (RIPO), the net new cash flows of US equity mutual funds (FLOW) and the closed-end funds discount (CEFD). We also compare our composite index with two indirect indicators commonly used in the literature: the ratio of odd-lot purchases to sales (ODDLOT)[19] and the NYSE ARMS index (ARMS).

The sentiment indicator ODDLOT is employed by Neal and Wheatley [1998]. A ratio that is greater (less) than one indicates bullish (bearish) sentiment. The ARMS index is used by Simon and Wiggins III [2001] and Wang, Keswani and Taylor [2006]. This ratio represents the number of advancing issues scaled by the advancing volume, divided by the number of declining issues scaled by the declining volume. Richard Arms, who first introduced this ratio, argues that a ratio that is lower (higher) than one indicates that the market is overbought (oversold) and that this should be treated as a bullish (bearish) sign.

Table 5 shows that with the exception of the ODDLOT ratio, the predictive power of indirect indicators is quite disappointing. The ODDLOT ratio has significant predictive abilities for returns on portfolios sorted by size, profitability and tangibility. When CEFD is used as a

sentiment indicator, only the rational expectations component predicts the portfolio returns. This result is consistent with Swaminathan's [1996] argument to link the information in discounts to expectations of future earnings growth and expectations of future inflation.

- **The Composite Sentiment Index vs. other composite index**

We now evaluate two other popular sentiment composite indicators routinely used in the literature, i.e. the Brown and Cliff composite index (BC) and the Baker and Wurgler composite index (BW)[20]. The results in Table 5 show that the BC index forecasts portfolio returns better than the BW index. The BC index forecasts better the returns of portfolios ranked by age, tangibility and dividend policy. The results also show that the two composite indicators outperform the direct and the indirect individual measures in predicting future portfolio returns. This result confirms that the integration of several proxies is a fruitful approach for measuring investor sentiment.

Finally, we perform an additional robustness test to evaluate the pertinence of the proxies selected to construct our composite sentiment index. We build a "new CSI"[21] that incorporates two additional proxies for sentiment, i.e. AAII and ARMS and re-estimate model (4). Finding[22] shows that the "new CSI" has lower predictive ability than the initial CSI. This result suggests that the greater predictive power of the CSI is probably due to its ability to capture more and better the variable sentiment[23].

Overall, our composite indicator exhibits superior predictive power than those all other sentiment indicators habitually used in the literature. Indeed, none of direct, indirect and alternative composite sentiment indicators records predictive power as significant as those of our composite index. This result led us to believe that the superiority of our measure comes from the simultaneous effect of the combination of both direct and indirect indicators.

6. Conclusion

Measuring investor sentiment and quantifying its effects on the stock market is at the center of academic research in behavioral finance. This study enriches the literature on the relationship between investor sentiment and stock returns. In particular, we split the raw sentiment indicator into a rational component related to economic fundamentals and a psychological component linked to investor sentiment. Using a panel of direct and indirect sentiment measures, we develop a new composite sentiment indicator and a new composite fundamental indicator. As a result, we investigate simultaneously the impact of economic fundamentals and the impact of investor sentiment on the stock returns. Our results show that the rational approach can be usefully completed by the behavioral approach.

Consistent with the traditional approach, we report a significantly relationship between the investors' rational expectations and the stock returns. Certain characteristics make companies more vulnerable to changes in the business cycle. In particular, we found that small stocks, value stocks, young stocks, unprofitable stocks and intangible stocks are less (more) vulnerable to the business cycle after periods characterized by good (bad) economic expectations.

Consistent with the behavioral approach, we find that investor sentiment impact stocks hard to value and difficult to arbitrage. After controlling for economic fundamentals and structural risks factors, we find that pricing error covariate significantly with the sentiment component. In particular, the stocks hard to value and difficult to arbitrage earn low (high) returns following periods when investor sentiment is high (low).

We also compare our composite sentiment index to each measure used in its construction and to several other popular measures routinely used in the literature. We report a dominance of our composite measure in forecasting stock returns. Our composite index provides a better measure

of sentiment by condensing the state of mind of a very large sample of investors (consumer confidence index, investors intelligence index, closed-end fund discount, mutual funds flows, the average monthly first-day returns on IPOs and the number of IPOs). This result led us to believe that the superiority of our composite measure comes from the wealth of information exploited in its construction using both direct and indirect measures.

NOTES

[1] Other researchers use exogenous non-economic factors that affect investor mood such as sunshine [Hirschleifer and Shumway, 2003], lunar cycles [Yuan, Zheng and Zhu, 2005], sport [Edmans, Garcia and Norli, 2007] or aviation disasters [Kaplanski and Levy, 2010].

[2] For example, several controversial theories have been suggested to explain the closed-end fund discount, e.g. agency costs, fund liquidity, fund diversification, market segmentation and investor sentiment.

[3] The results of these studies take into account the Baker and Wrugler composite indicator that include only indirect indicators.

[4] It's difficult to use a unique indicator to measure investor sentiment. The closed-end fund discount, for instance, will not be a worthwhile proxy if a large number of investors have come to prefer open-end funds. During some months, the number of IPOs is equal to zero although the market is not necessarily at the lower level during the period studied.

[5] This hypothesis will be revisited in the fourth section.

[6] See for example Brown and Cliff [2005], Lemmon and Portniaguina [2006] and Baker and Wurgler [2006, 2007].

[7] Baker and Wrugler [2006] argue that the timing of the variable must be taken into account because some variables may reflect a change in sentiment earlier than other.

[8] The sentiment proxies are standardized to have a mean of zero and a standard deviation of one.

[9] The fluctuations of the composite index during the speculative bubble of 2000 (Internet bubble) are much less significant than the fluctuations during the crash of October 1987. One possible explanation is that the fall in prices in 1987 has been more drastic (about 23% in one day) than during the Internet bubble (the decrease took place over several months).

[10] High idiosyncratic risk makes arbitrage risky (see Wurgler and Zhuravskaya, 2002).

[11] In order to make results easier to interpret, we have rescaled the proxy for the variable "Arbitrage costs" such that lower value corresponds to the stocks the most difficult to arbitrage. The rescaling is accomplished by taking the reciprocal of the proxy used to capture arbitrage costs, i.e. (1/Arbitrage costs).

[12] For the sake of brevity, we present only the results for a one-month holding period return.

[13] Results are available upon request.

[14] The monthly time series of theses factors are obtained from Ken French's data library.

[15] The variable SMB is not included as a control variable in panel A because it is highly correlated with the dependent variable by construction. Similarly, HML is not included as a control variable in panel B because of its high correlation with the dependent variable.

[16] As robustness check we also run the regression in Equation (2) using the first principal component from the six raw investor sentiment indicators used in this study. The basic results, not reported, are unchanged.

[17] Overall, results are significant for the 6-month and the 9-month periods and not significant for the 12-month period. By and large, the statistical significance of the coefficients estimate decreases as the holding period increases. Results are available upon request

[18] AAII data are available for the period July 1987 to December 2010.

[19] ODDLOT data are available for the period July 1981 to December 1998.

[20] Brown and Cliff's data are available for the period July 1981 to December 1998. As far as Baker and Wurgler's composite sentiment index is concerned, data are available for the period July 1981 to December 2007.

[21] The "new CSI" is estimated for the period July 1987 to December 2010.

[22] The results are not reported due to space limitation.

[23] The new first principal component captures a lower proportion of the sample variance (about 36 %). Likewise, all of the individual sentiment indicators of the "new CSI" have smaller factor loading than those of initial CSI. The exclusion of some variables does not also improve the results.

References

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

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

Black, F. "Noise." *Journal of Finance*, 41, (1986), pp. 529-543.

Brown, G.W. "Volatility, Sentiment, and Noise Traders." *Financial Analysts Journal*, 55, (1999), pp. 82-90.

Brown, G.W. and M.T. Cliff. "Investor sentiment and the near-term stock market." *Journal of Empirical Finance*, 11, (2004), pp. 1-27.

Brown, G.W. and M. T. Cliff. "Investor Sentiment and Asset Valuation." *Journal of Business*, 78, (2005), pp. 405-440.

Chan, K.C. and N-F. Chen. "Structural and Return Characteristics of Small and Large Firms." *Journal of Finance*, 46, (1991), pp. 1467-1484.

Clarke, R.G. and M. Statman. "Bullish or Bearish?" *Financial Analysts Journal*, 54, (1998), pp. 63-52.

De Bondt, W. F., M. "Betting on trends: Intuitive forecasts of financial risk and return." *International Journal of Forecasting*, 9, (1993), pp. 355-371.

De Long, J. B., A. Shleifer, L. H. Summers and R. J. Waldmann. "Noise trader risk in financial markets". *Journal of Political Economy*, 98, (1990), pp. 703-738.

Edmans, A., D. Garcia, and Ø. Norli. "Sports Sentiment and Stock Returns." *Journal of Finance, 62*, (2007). pp.1967-1998.

Feldman, T. "A More Predictive Index of Market Sentiment." *Journal of Behavioral Finance*, 11, (2010), pp. 211-223.

Fisher, K.L., and M. Statman. "Consumer Confidence and Stock Returns." *Journal of Portfolio Management*, 30, (2003), pp.115-127.

Hirshleifer, D. and T. Shumway. "Good Day Sunshine: Stock Returns and the Weather." *Journal of Finance*, 58, (2003), pp. 1009-1032.

Jagannathan, R. and W. Zhenyu. "The Conditional CAPM and the Cross-Section of Expected Returns." *Journal of Finance*, 51, (1996), pp. 3-53.

Jegadeesh, N. and S. Titman. "Profitability of Momentum Strategies: An Evaluation of Alternative Explanations." *Journal of Finance*, 56, (2001), pp. 699-720.

Kaplanski, G. and H. Levy. "Sentiment and stock prices: The case of aviation disasters." *Journal of Financial Economics*, 95, (2010), pp.174-201.

Kumar, A., and C.M.C. Lee. "Retail Investor Sentiment and Return Comovements." *Journal of Finance*, 61, (2006), pp. 2451-2486.

Lemmon, M. and E. Portniaguina. "Consumer Confidence and Asset Prices: Some Empirical Evidence." *Review of Financial Studies*, 19, (2006), pp.1499-1529.

Neal, R. and S.M. Wheatley. "Do measures of investor sentiment predict returns?" *Journal of Financial & Quantitative Analysis*, 33, (1998), pp. 523-547.

Otoo, M. "Consumer sentiment and the stock market." *Federal of Reserve System*, 1999.

Qiu, L. and Ivo W. "Investor Sentiment Measures." *Brown University and NBER*, 2006.

Schmeling, M. "Investor sentiment and stock returns: Some international evidence." *Journal of Empirical Finance* 16, (2009), pp. 394-408.

Simon, D.P. and R.A. Wiggins III "S&P futures returns and contrary sentiment indicators." *Journal of Futures Markets*, 21, (2001), pp. 447-462.

Soft, M.E., and M. Statman. "How Useful is the Sentiment Index?" *Financial Analysts Journal* 44, (1988), pp. 45-55.

Swaminathan, B. "Time-varying expected small firm returns and closed-end fund discounts." *Review of Financial Studies*, 9, (1996), pp. 845-887.

Wang, Y-H., A. Keswani and S.J. Taylor. "The relationships between sentiment, returns and volatility." *International Journal of Forecasting*, 22, (2006), pp. 109-123.

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

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

Zouaoui, M., G. Nouyrigat and F. Beer. "How does investor sentiment affect stock market crises: Evidence from panel data." *The Financial Review*, (2011), forthcoming.

Table 1: Description of the variables used for the construction of the sentiment index

Code	Variables	Measures	Sources
Investor sentiment indicators			
UMI	University of Michigan consumer sentiment index	Five questions making up the consumer sentiment index	University of Michigan Survey Research Center
II	Investors Intelligence index	Bull minus Bear spread	Investors Intelligence
NIPO	Number of IPOs	Number of IPOs in a given month	http://bear.cba.ufl.edu/ritter
RIPO	First-day returns on IPOs	Average monthly first-day returns on IPOs	http://bear.cba.ufl.edu/ritter
FLOW	Net new cash flows of US equity mutual funds	(Inflows-outflows)/Total asset	Investment Company Institute http://www.ici.org/index.html
CEFD	Closed-end fund discount	Equal-weighted average difference between the NAV of closed-end fund and the stock price of fund	Wall Street Journal
CBIND	Conference Board survey of consumer confidence	Five questions making up the consumer sentiment index	Conference Board
AAII	American Association of Individual Investors index	Ratio of bullish to bearish responses	American Association of Individual Investors
ODDLOT	ODDLOT ratio	Ratio of odd-lot purchases to sales	New York Stock Exchange
ARMS	ARMS index NYSE	$\frac{\text{Advancing Issues} / \text{Declining Issues}}{\text{Advancing Volume} / \text{Declining Volume}}$	Datastream
CSI	Composite sentiment index	First component from the principal component analysis of six sentiment measures	
BC	Brown and Cliff composite index	First component from the principal component analysis of numerous sentiment measures	Brown and Cliff
BW	Baker and Wurgler composite index	First component from the principal component analysis of six sentiment measures	Baker and Wurgler http://people.stern.nyu.edu/jwurgler

(Continued)

Table 1 (continued): Description of the variables used for the construction of the sentiment index

Macroeconomics variables			
IP	Growth of industrial production	Change in the natural logarithm of industrial production index	Federal reserve system
INF	Inflation	Change in the natural logarithm of the Consumer Price Index	Federal reserve system
TS	Term spread	Difference between the yields on 10-year U.S. government bonds and 3-month Treasury bills	Federal reserve system
DS	Default spread	Moody's Baa-rated corporate bond yield less the Aaa-rated corporate bond yield	Datastream
DC, NDC and SC	Growth of durable goods, non-durable goods and services consumption expenditures	Change in the natural logarithm of durable goods, non-durable and services consumption expenditures	Federal reserve system

Tableau 2: Monthly correlations between sentiment indicators and macroeconomic variables

This table reports the Pearson correlation coefficients between sentiment indicators and macroeconomic variables. UMI is the University of Michigan consumer confidence index. II is the Investors Intelligence spread Bull-Bear. NIPO represents the number of IPOs in a given month and RIPO denotes the average monthly first-day returns on IPOs. FLOW is the net new cash flows of US equity mutual funds and CEFD is the closed-end funds discount. IP is the growth in industrial production. INF is the inflation rate. ST is the term spread. DS is the default spread. CD, CND and CS are the growth in durable, nondurable, and services consumption, respectively. The sample period includes monthly data from June 1981 to December 2010. ***, **, * denote statistical significance at 1%, 5%, and 10%, respectively.

Panel A	**II**	**UMI**	**CEFD**	**NIPO**	**RIPO**	**FLOW**	**INF**	**IP**	**TS**	**DS**	**CD**	**CND**	**CS**
Mean	12.144	87.977	7.287	31.946	17.017	0.698	0.252	0.168	1.964	1.094	0.565	0.406	0.534
Std	15.276	12.236	6.234	24.700	17.579	0.928	0.273	0.681	1.112	0.489	0.521	0.627	0.363
II	1												
UMI	0.312***	1											
CEFD	-0.287***	0.082	1										
NIPO	0.165**	0.341***	-0.062	1									
RIPO	0.288***	0.371***	0.082	0.079	1								
FLOW	0.392***	0.026	-0.347***	0.412***	0.110	1							
Panel B													
INF	-0.309***	-0.192***	0.223***	-0.028	-0.051	-0.156***	1						
IP	0.024	0.326***	-0.021	0.351***	0.091	0.120*	0.042	1					
TS	0.036	-0.332***	-0.436***	0.024	-0.321***	0.111	-0.135**	0.066	1				
DS	-0.028	-0.583***	0.078	-0.116	-0.178***	-0.161***	0.134**	-0.258***	0.281***	1			
CD	-0.012	0.022	-0.010	0.121**	0.104	-0.053	0.051	0.228***	0.032	0.106	1		
CND	-0.002	0.042	0.088	0.048	0.141**	-0.021	0.271***	0.232***	-0.041	-0.041	0.549***	1	
CS	0.065	-0.067	0.057	0.110	0.046	0.027	0.017	0.032	0.092*	0.281***	0.401***	-0.008	1

Table 3: Time series regressions of portfolio returns on the composite sentiment index and the fundamental composite index

This table presents the results of the estimation of Lemmon and Portniguina's model. The estimation is based on the following model:

$$R^{j}_{(Low),t} - R^{j}_{(High),t} = (\alpha_1 + \alpha_2 CSI_{t-1}) + (\beta_1 + \beta_2 CFI_{t-1}) RMRF_t + \varepsilon^{j}_{t}$$

$R_{(Low),t}$-$R_{(High),t}$ is one-month holding period return for long-short portfolios based on characteristics j. Low is defined as a firm in the bottom three CRSP deciles and High is defined as a firm in the top three CRSP deciles. CSI_{t-1} and CFI_{t-1} are the lagged values of the composite sentiment index and the fundamental composite index respectively. $RMRF_{,t}$ is the excess return on the CRSP value-weighted index for month t. The long-short portfolios are formed based on firm characteristics, i.e.: Size, Market-to book ratio, Age, Profitability, Tangibility, Dividend Policy and 1/Arbitrage costs. The sample period includes monthly data from July 1981 to December 2010. Δ adj.R^2 shows the improvement of the adjusted R^2 after the addition of the sentiment indicator. The Newey-West adjusted *t*-statistics of the coefficient estimates are reported in the parentheses. ***, **, * denote statistical significance at 1%, 5%, and 10%, respectively.

	Panel A: *size*	***Panel B:*** *Market to-book ratio*	***Panel C:*** *Age*	***Panel D:*** *Profitability*	***Panel E:*** *Tangibility*	***Panel F:*** *Dividend Policy*	***Panel G:*** *1/Arbitrage costs*
Intercept	0.002*	-0.005 *	0.003	0.000	0.004**	0.000	0.002
	(1.74)	(-1.68)	(0.11)	(0.97)	(2.15)	(0.230)	(0.655)
CSI_{t-1}	-0.032**	-0.021**	-0.024**	-0.020 **	-0.031***	-0.011*	-0.020**
	(-2.06)	(-1.99)	(-2.01)	(-2.04)	(-2.81)	(-1.87)	(-2.12)
$RMRF_t$	0.812***	0.711***	0.961***	0.816 ***	0.959***	0.870***	1.120***
	(3.45)	(3.13)	(2.86)	(3.49)	(2.66)	(3.305)	(3.976)
$CFI_{t-1} \times RMRF_t$	-0.058***	-0.042**	-0.039*	-0.032 *	-0.023*	-0.003	-0.002
	(-3.18)	(-1.98)	(-1.87)	(-1.85)	(-1.83)	(-0.907)	(-1.357)
Adj.R^2	0.239	0.198	0.184	0.201	0.185	0.156	0.168
Δ adj.R^2	0.059	0.045	0.052	0.039	0.041	0.031	0.053

Table 4: Time series regressions of portfolio returns, controlling for the other fundamental risk factors

This table presents the results of the estimation of Lemmon and Portniguina's model augmented by the other fundamental risk factors. The estimation is based on the following model:

$$R^{j}_{(Low),t} - R^{j}_{(High),t} = (\alpha_1 + \alpha_2 CSI_{t-1}) + (\beta_1 + \beta_2 CFI_{t-1} + \beta_3 DIV_{t-1} + \beta_4 DS_{t-1} + \beta_5 INT_{t-1}) RMRF_t + \delta_1 SMB_t + \delta_2 HML_t + \delta_3 UMD_t + \varepsilon^{j}_{t}$$

$R_{(Low),t}$-$R_{(High),t}$ is one-month holding period return for long-short portfolios based on characteristics j. Low is defined as a firm in the bottom three CRSP deciles and High is defined as a firm in the top three CRSP deciles. CSI_{t-1} and CFI_{t-1} are the lagged of composite sentiment index and the fundamental composite index respectively. *DIV* is the dividend yield on the market; *DS* is the default spread and *INT* is the three-month treasury rate. $RMRF_{,t}$ is the excess return on the CRSP value-weighted index for month t. *SMB* is the premium on a portfolio of small stocks relative to large stocks; *HML* is the premium on a portfolio of high book to market stocks relative to low book to market stocks and *UMD* is the premium on a portfolio of stocks with high past returns relative to stocks with low past returns. The long-short portfolios are formed based on firm characteristics, i.e.: Size, Market-to book ratio, Age, Profitability, Tangibility, Dividend Policy and 1/Arbitrage costs. The sample period includes monthly data from July 1981 to December 2010. Δ adj.R^2 shows the improvement of the adjusted R^2 after the addition of the sentiment indicator. The Newey-West adjusted *t*-statistics of the coefficient estimates are reported in the parentheses. ***, **, * denote statistical significance at 1%, 5%, and 10%, respectively.

	Panel A: *size*	***Panel B:*** *Market-to book ratio*	***Panel C:*** *Age*	***Panel D:*** *Profitability*	***Panel E:*** *Tangibility*	***Panel F:*** *Dividend Policy*	***Panel G:*** *1/Arbitrage*
Intercept	0.009* (1.73)	-0.003 (-1.21)	0.004 (1.12)	-0.003 (-0.87)	0.006*(1.89)	0.002(0.34)	0.007(0.87)
CSI_{t-1}	-0.029** (-2.02)	-0.020** (-1.97)	-0.019**(-1.98)	-0.018** (-1.96)	-0.027***(-2.51)	-0.009 (-1.55)	-0.017**(-1.99)
$RMRF_t$	0.806*** (3.65)	0.706** (3.09)	0.897***(3.66)	0.855***(3.38)	0.897***(3.76)	0.870***(4.30)	1.110***(3.26)
$CFI_{t-1} \times RMRF_t$	-0.044** (-3.07)	-0.038** (-2.02)	-0.035*(-1.92)	-0.033**(-1.99)	-0.019*(-1.71)	-0.009(-1.10)	-0.002(-1.355)
$DIV_{t-1} \times RMRF_t$	0.098 (1.432)	0.101 (1.52)	0.143** (1.97)	0.105 (1.57)	0.132* (1.66)	0.099 (1.22)	0.128** (2.11)
$DS_{t-1} \times RMRF_t$	-0.023** (-2.24)	-0.019** (-1.98)	-0.016 (-0.87)	-0.014 (-0.98)	-0.017 (-1.54)	-0.021** (-2.09)	-0.022*** (-2.67)
$INT_{t-1} \times RMRF_t$	0.012 (1.45)	0.014 (0.71)	0.027* (1.76)	0.024** (1.98)	0.013 (1.61)	0.019** (1.97)	0.011 (1.23)
SMB		0.132 (1.16)	0.234* (1.74)	0.099 (1.12)	0.245** (1.98)	0.097 (0.91)	0.145* (1.66)
HML	0.142 (0.92)		0.07 (0.56)	-0.082 (-0.98)	0.133 (0.65)	-0.198** (-2.06)	-0.087 (-0.99)
UMD	0.050 (1.12)	-0.081 (-0.89)	0.064 (1.47)	0.072** (1.98)	0.066* (1.76)	0.098* (1.92)	-0.049 (-1.27)
Adj.R^2	0.269	0.203	0.206	0.212	0.196	0.176	0.258
Δ adj.R^2	0.051	0.041	0.046	0.029	0.034	----	0.030

Table 5: The relevance of alternative sentiment indicators

This table reports the coefficients estimates (β_2) and (α_2) of Lemmon and Portniguina's model using an alternative sentiment indicator. UMI is the University of Michigan consumer confidence index. II is the Investors Intelligence spread Bull-Bear. NIPO represents the number of IPOs in a given month and RIPO denotes the average monthly first-day returns on IPOs. FLOW is the net new cash flows of US equity mutual funds. CEFD is the closed-end fund discount. AAII is the American Association of Individual Investors sentiment index. ODDLOT is the ratio of odd-lot purchases to sales. ARMS is the NYSE ARMS index. BC is the Brown and Cliff composite sentiment index and BW is the Baker and Wurgler composite sentiment index. The sample period includes monthly data from July 1981 to December 2010, except for AAII, ODDLOT and the two alternative composite indexes. AAII data are available for the period July 1987 to December 2010. ODDLOT and Brown and Cliff's data are available for the period July 1981 to December 1998. Baker and Wurgler's data are available for the period July 1981 to December 2007. The Newey-West adjusted t-statistics of the coefficient estimates are reported in the parentheses. ***, **, * denote statistical significance at 1%, 5%, and 10%, respectively.

Sentiment measures	**Panel A** *size*		**Panel B** *Market-to book*		**Panel C** *Age*		**Panel D** *Profitability*		**Panel E** *Tangibility*		**Panel F** *Dividend Policy*		**Panel G** *1/Arbitrage costs*	
Direct measures	**α_2**	**β_2**	**α_2**	**β_2**	**α_2**	**β_2**	**α_2**	**β_2**	**α_2**	**β_2**	**α_2**	**β_2**	**α_2**	**β_2**
II	**-0.013****	-0.008	**-0.041*****	-0.009	-0.001	0.005	-0.000	0.005	-0.001	-0.001	-0.001	0.005	0.001	-0.001
	(-1.981)	(-1.246)	**(-2.404)**	(-1.157)	(-0.728)	(1.114)	(-0.352)	(1.099)	(-0.547)	(-0.269)	(-1.098)	(0.170)	(0.309)	(-0.313)
UMI	-0.000	0.006	-0.000	**-0.021***	-0.001	0.011	-0.003	0.002	-0.002	-0.009	-0.000	0.005	-0.005	-0.002
	(-0.762)	(0.307)	(-0.727)	**(-1.667)**	(-0.639)	(1.106)	(-1.035)	(0.989)	(-1.029)	(-1.244)	(-0.136)	(0.663)	(-1.265)	(-0.278)
AAII	**-0.009***	0.007	-0.001	-0.005	-0.009	0.009	-0.008	-0.009	-0.008	-0.008	0.009	0.004	-0.003	-0.009
	(-1.781)	(0.010)	(-0.127)	(-1.267)	(-1.239)	(0.106)	(-1.009)	(-0.987)	(-0.919)	(-1.244)	(1.009)	(0.928)	(-0.876)	(-1.098)
CBIND	-0.001	0.004	-0.002	**-0.018****	0.007	-0.009	-0.003	-0.004	0.009	0.008	-0.001	0.005	-0.005	-0.009
	(-0.106)	(0.207)	(-0.654)	**(-1.989)**	(-0.435)	(0.106)	(-1.085)	(-0.989)	(-1.009)	(1.244)	(-0.105)	(0.639)	(-1.098)	(-1.159)
Indirect measures	**α_2**	**β_2**	**α_2**	**β_2**	**α_2**	**β_2**	**α_2**	**β_2**	**α_2**	**β_2**	**α_2**	**β_2**	**α_2**	**β_2**
NIPO	-0.008	**-0.040*****	0.000	**-0.013***	0.000	-0.009	-0.000	-0.003	-0.002	-0.005	-0.000	**-0.021*****	-0.006	-0.019
	(-1.012)	**(-2.739)**	(0.261)	**(-1.802)**	(-1.475)	(-1.629)	(-0.799)	(-0.632)	(-1.337)	(-1.201)	(-0.307)	**(-2.721)**	(-1.334)	(-1.618)
RIPO	0.000	0.006	-0.000	0.002	0.000	0.025	0.000	0.009	0.000	-0.004	0.000	0.006	0.000	-0.004
	(1.261)	(0.315)	(-1.607)	(0.544)	(1.450)	(1.225)	(1.246)	(1.446)	(1.256)	(-0.823)	(1.198)	(1.604)	(1.234)	(-1.591)
FLOW	-0.005	-0.004	0.000	-0.002	-0.000	0.096	-0.002	0.002	0.000	-0.007	-0.001	-0.000	-0.001	-0.002
	(-0.627)	(-0.694)	(0.435)	(-0.968)	(-0.062)	(0.473)	(-0.869)	(1.033)	(0.178)	(-1.056)	(-0.731)	(-0.198)	(-0.368)	(-0.596)
CEFD	-0.001	**0.082****	0.000	**0.036****	0.000	**0.028*****	0.000	0.009	0.000	0.008	0.000	**-0.027****	0.000	**0.018***
	(-0.859)	**(1.985)**	(0.229)	**(2.258)**	(0.838)	**(2.553)**	(0.710)	(1.613)	(0.710)	(1.131)	(0.119)	**(-2.246)**	(0.288)	**(1.922)**
ODDLOT	**-0.028***	0.002	-0.001	-0.002	0.001	**-0.022***	**-0.021***	0 .002	**-0.024***	**-0.019***	0.004	0.008	0.007	0.001
	(1.869)	(0.096)	(-0.987)	(-0.627)	(0.861)	**(1.752)**	**(1.771)**	(0.770)	**(1.895)**	**(-1.798)**	(0.098)	(1.361)	(1.474)	(0.243)
ARMS	-0.002	0.008	-0.005	0.002	-0.015	-0.005	-0.007	0.005	0.006	**0.006***	0.007	**0.005***	-0.002	0.003
	(-0.543)	(1.086)	(-0.377)	(0.240)	(-1.093)	(-0.461)	(-0.573)	(1.161)	(0.04)	**(2.104)**	(0.588)	**(2.203)**	(-0.844)	(1.080)
Composite measures	**α_2**	**β_2**	**α_2**	**β_2**	**α_2**	**β_2**	**α_2**	**β_2**	**α_2**	**β_2**	**α_2**	**β_2**	**α_2**	**β_2**
BW	**-0.012***	**-0.022***	**-0.031****	**-0.018***	-0.008	0.038	-0.002	0.007	-0.008	**-0.049****	-0.001	-0.001	**-0.019****	-0.001
	(-1.687)	**(-1.765)**	**(-2.686)**	**(-1.894)**	(-0.985)	(0.730)	(-0.907)	(1.352)	(-1.369)	**(-1.991)**	(-0.995)	(-0.449)	**(-2.214)**	(-0.198)
BC	**-0.019****	-0.004	**-0.042*****	**-0.027***	**-0.032****	0.035	0.006	0.005	**-0.042****	-0.055	**-0.017***	-0.002	**-0.011***	-0.002
	(-1.981)	(-0.442)	**(-2.881)**	**(-1.786)**	**(-2.072)**	(0.680)	(1.354)	(1.205)	**(-2.202)**	(-1.299)	**(1.945)**	(-0.623)	**(-1.891)**	(-0.442)

Figure 1: Composite sentiment index and composite fundamental index, July 1981 to December 2010

This figure shows the development of the composite sentiment index (CSI) and the composite fundamental index (CFI) over time. We regress each raw sentiment indicator on the growth in industrial production, the inflation, the term spread, the default spread and the growth in durable, nondurable, and services consumption. The fitted values of the regression capture the fundamental component and the residual capture the sentiment component. The composite (fundamental) sentiment index is constructed from a principal component analysis (first component) of six sentiment (fundamental) indicators identified in previous studies: the consumer confidence index, the investors intelligence index, the number of IPOs, the average monthly first-day returns on IPOs, the mutual funds flows and the closed-end fund discount. The sentiment proxies are standardized to have a mean of zero and a standard deviation of one.