NBER WORKING PAPER SERIES

INVESTOR SENTIMENT MEASURES

Lily Qiu
Ivo Welch

Working Paper 10794
http://www.nber.org/papers/w10794

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2004

We especially thank Charles Lee, Stephen Ross, Andre Shleifer, and Richard Thaler for making their data generously available to us. Umit Gurun provided newer data, and seminar participants at Mchigan and Rochester had to suffer through a presentation involving some of the tables in this paper. The views expressed herein are those of the author(s) and not necessarily those of the National Bureau of Economic Research.

Investor Sentiment Measures
Lily Qiu and Ivo Welch
NBER Working Paper No. 10794
September 2004, Revised July 2006
JEL No. G12, G14

ABSTRACT

This paper compares investor sentiment measures based on consumer confidence surveys with measures extracted from the closed-end fund discount (CEFD). Our evidence suggests that these two kinds of sentiment measures do not correlate well with one another. For a short 2 −4 year period in which we have direct investor sentiment survey data from UBS/Gallup, only the consumer confidence correlates well with investor sentiment. Further, only the consumer confidence based measure can robustly explain the small-firm return spread and the return spread between stocks held disproportionately by retail investors and those held by institutional investors. Surprisingly, there is even a hint that the consumer confidence measure can explain closed-end fund IPO activity, while the CEFD cannot. In sum, our evidence supports the view that sentiment plays a role in financial markets, but that the CEFD may be the wrong measure of sentiment.

Lily Qiu
Department of Economics
Box B
Brown University
Providence, RI 02912

Ivo Welch
Yale University
School of Management
46 Hillhouse Ave.
New Haven, CT 06520-8200
and NBER
ivo_welch@brown.edu

The behavioral theory of DeLong, Shleifer, Summers, and Waldmann (1990) predicts that noise trader sentiment can persist in financial markets. They argue that *changes* in noise trader sentiment must be difficult to predict to avoid arbitrage. Assets that are disproportionally exposed to noise trader risk are both riskier and have to offer an extra return premium. In sum, the theory predicts that sentiment can influence security pricing under two necessary conditions: [1] the assets are held predominantly by sentiment (noise) traders, and [2] transaction costs are high enough to prevent systematic arbitrage by arbitrageurs.

Lee, Shleifer, and Thaler (1991), henceforth LST, first explore the empirical implications of this theory by assuming that noise traders are individual investors. Because individual investors were already known to disproportionally hold closed-end funds (henceforth, CEF), LST interpret the closed-end fund discount (henceforth, CEFD) as a (negative) sentiment factor. Lee, Shleifer, and Thaler (1991) then wring further implications and empirical support from this insight. Their most important implication is that *decreases* in the CEFD (i.e., more optimism) should be positively correlated with the returns of assets that are disproportionally held by noise sentiment traders. LST identify the smallest decile of firms as such. They find that small firms outperform large firms when the CEFD decreases.

Although Lee, Shleifer, and Thaler (1991) discount other possible factors determining the CEFD, first and foremost agency (transaction) costs, they concede that multiple factors are likely to influence the CEFD. Ross (2005) explores these factors in more detail and argues that transaction costs are more important than LST realized. In any case, the behavioral and the transaction cost views of the closed-end fund discount are *not* mutually exclusive. Moreover, Spiegel (1997) and Berk and Stanton (2004) have recently proposed a rational explanation for some of the time-pattern in the CEFD, although neither focuses on the correlation between the extreme size decile return spread and sentiment.

Interest in the CEFD as a sentiment index has not waned. Indeed, Lee, Shleifer, and Thaler (1991) has been a seminal paper, both in the novelty of its ideas and its subsequent impact: as of March 2004, a quick citation search yields over 100 cites to it. A search in SSRN shows that "investor sentiment" finds 53 matches, compared to 78 for the phrase "APT." If anything, investor sentiment has become a subject of more intense interest. The CEFD remains the only widely used proxy of investor sentiment, and is itself a component in some other measures of investor sentiment, though all of these have remained fairly boutique.

Our own paper revisits the empirical evidence. We focus only on time-series sentiment evidence, and borrow ideas and concepts from LST. But we can improve on the LST metrics. First, our paper has an unfair advantage over LST—hindsight. That is, we can expand our sample from 1985 to 2002 (16 more years), almost doubling the sample, which provides a true out-of-sample test. Second, although sentiment is an abstract concept, we have what we believe to be the best available empirical direct proxy for investor sentiment. It is one

of the components in the *Survey of Investor Sentiment* conducted by *UBS/Gallup*. This question asks respondents with savings and investments of over $10,000 the question:

> *I would like to ask you to think about the factors that could affect the overall investment environment OVER THE NEXT TWELVE MONTHS...how would you rate performance of the stock market OVER THE NEXT TWELVE MONTHS?*

The data also allows us to split respondents into investors with more than $100,000 and less than $100,000 in wealth. We can exploit the UBS/Gallup survey to validate whether or not other proxies correlate well with this measure of investor sentiment. Unfortunately, the UBS/Gallup survey began regularly in 1999 and is only available in the United States, limiting our ability to rely exclusively on it. Third, we follow LST in exploring the role of investor sentiment on financial market pricing, but we also explore whether the pathways through which consumer confidence correlates with financial markets are real (future corporate profits based) or financial (investor preference based).

Our paper offers three primary findings, each standing on their own:

1. We show that the CEFD is not a good proxy for investor sentiment. It has no correlation with the direct *UBS/Gallup* measure (or other proxies of investor sentiment).

2. We show that the alternative candidate, consumer confidence, is a reasonable proxy: consumer confidence changes correlate strongly with changes in the *UBS/Gallup* proxy. That is, the two share a strong common factor, although they sample very different respondents. It is likely that they both pick up on the underlying factor—investor sentiment.

3. We explore the impact of sentiment in financial markets.

   We find that the CEFD has only reliably worked in Januaries in explaining small stock excess returns even in the LST sample. Furthermore, it has not worked after 1985.

   In contrast, consumer confidence predicts small firm returns in non-January months just as well as in January months. When we decompose changes in consumer confidence into a part that does and a part that does not correlate with present and/or future consumption and corporate profits, we find that both help explain the excess return spread between the two extreme firm-size deciles. Consumer confidence also does not diminish in importance if a contemporaneous measure of CEO confidence is included. Thus, the pricing influence of consumer confidence in financial markets is not a phenomenon that arises exclusively through the predictive power of consumer confidence changes on future real economic conditions. There is mild evidence that consumer confidence changes have marginal impact on the returns of retail-held illiquid stocks, holding firm-size constant, but this evidence is fragile.

Our consumer confidence evidence is consistent with DeLong, Shleifer, Summers, and Waldmann (1990). It establishes a necessary, but not a sufficient condition for this behav-

ioral perspective. Yet, absent precise quantitative theories of how sentiment should influence financial markets in an irrational-world perspective vs. a rational-world perspective—so that we can attribute correlation that can be claimed by either perspectives exclusively to one or the other—we cannot determine whether investor sentiment is an entirely behavioral, irrational or a classical, rational phenomenon.

Our paper is similar to Lemmon and Portniaguina (2004), which is both contemporaneous and independent. Our paper differs in that our paper [a] relies on direct validation via the UBS/Gallup proxy; [b] controls for different macroeconomic factors; [c] predicts returns with changes in consumer confidence, not levels in consumer confidence; and [d] does not interpret the retail stock return spread evidence to be robust.

Our next section describes the closed-end fund discount and consumer confidence, and the theories behind their use. Section II describes the principal validation series, the UBS/Gallup investor sentiment poll and the correlation of other proxies therewith. It also addresses issues of bias, noise, investor identification, and the difference between a direct validation test and a theory-based joint hypothesis test. Finally, it discusses other proxies of investor sentiment that we do not explore further. Our paper draws a sharp distinction between this part, and what follows: an exploration of the influence of investor sentiment on financial market pricing in Section III. This section reexamines the LST evidence, decomposes the pathway through which consumer confidence influences returns, and describes why it is intrinsically impossible to separate the influence of investor sentiment into a rational and an irrational component. Section IV shows that exposure to sentiment is not a priced factor in cross-section, and that sentiment itself is influenced by the financial markets. After concluding in Section V, Appendix A offers a FAQ that responds to earlier referee reports.

I The Closed-End Fund Discount and Consumer Confidence

Investor sentiment ranks among the most important theoretical and empirical concepts in behavioral finance. To have meaning, it would seem essential to have an accepted proxy for investor sentiment. (The alternative of attributing everything not explained by simple rational models to sentiment would be a very weak test.) The closed-end fund discount is the closest we have come to such a widely accepted measure. Other proxies, further discussed in Section II.C, have remained "boutique."

A The CEFD Measure and Theory-Based Validation

The Lee, Shleifer, and Thaler (1991) proxy for investor sentiment strongly relies on one empirical observation and one theoretical hypothesis. The empirical observation is that closed-end funds are predominantly held by retail investors. This is widely accepted. The

theoretical conjecture is that the closed-end fund discount is driven by retail investor sentiment. Under LST's theoretical conjecture, the CEFD becomes a proxy for investor sentiment. This may not be the case, and for two reasons. First, there can be other factors that can influence the CEFD and changes therein, for example transaction costs as noted by Ross (2005) or time-varying liquidity premia and agency costs. Second, closed-end funds could be disproportionally held by unusual retail investors (specifically, trust accounts), which may not represent ordinary retail investors.

LST primarily base their validation on the correlation of the CEFD with the small-firm decile excess rate of return.[1] However, this can be criticized because it validates one financial measure (the CEFD) with another (decile stock returns). This makes it more likely that both are driven by a financial markets phenomena that are not yet fully understood. One can further debate whether other regularities should be controlled away, or whether they are integral. It could be that any other factor "steals" significance from the CEFD precisely because it, too, reflects investor sentiment. This was partly at issue in a well-known (and amusing) exchange in the Journal of Finance in 1993. Chen, Kan, and Miller (1993) point out that the correlation between the CEFD and the size spread declined in the latter half of LST's sample. They also find that small firms with low institutional ownership have similar coefficients as those with high institutional ownership. In their rejoinder, Chopra, Lee, Shleifer, and Thaler (1993) respond that all small firms are generally noise-trader sensitive with low institutional ownership, and that expecting to find an effect after splitting subsamples again is asking for too much, thus imposing an incorrect null hypothesis.

In sum, the CEFD sentiment measure is a proxy that is based on the investor sentiment theory itself. The tests then measure consistency of one implication of the theory (the proxy extraction from the CEFD) with other implications (e.g., the rate of return on assets disproportionally held by noise traders). To the extent that other costs can matter to the CEFD (e.g., agency or fund fee costs), the CEFD proxy identification with sentiment can be weak. With both the test and the proxy based on financial data—the CEFD is essentially a book-to-market ratio—it is also relatively more likely that another theory (e.g., agency- or liquidity-based) could eventually offer an explanation for both findings, which could be different from investor sentiment.

[1] Lee, Shleifer, and Thaler (1991) also document that new CEF's tend to appear in LST's CEF data base when investor sentiment *levels* are very positive, and that the CEFD on different funds should be positively correlated. When using the SDC data base, we could not confirm more initiations of CEFs when investor sentiment is high. Lee, Shleifer, and Thaler (1991) can also hint that the theory is not inconsistent with a negative CEFD upon fund inception, followed by a sharp drop from a premium into a discount, and then a positive drift in CEFD reduction [to account for the need to offer a positive expected rate of return]. However, the dynamics are weak: the process by which the negative discount becomes positive is not clear, and we are not aware of evidence that the CEFD systematically narrows over time.

B CEFD Variable Construction

Both Lee, Shleifer, and Thaler (1991) and Ross (2005) generously shared their CEFD data series with us; descriptions of their constructions can be found in their respective papers. Our intent is to work with one long series, rather than with two separate series. Ex-ante, both sources provide equally valid measures of the CEFD. Our only concern is that the splicing (by averaging) the two series into one series could introduce a sharp discrepancy that is calculation based, especially around the breakpoints.

For the 72 months in which we have both the Ross (2005) closed-end fund data and the Lee, Shleifer, and Thaler (1991) data, the correlation between their value-weighted measures is 95%, The value-weighted regression coefficients are

$$\text{Ross CEFD Measure}_t \;=\; -0.21 \;+\; 0.973 \cdot \text{LST CEFD Measure}_t \;+\; \text{Noise}_t \qquad . \qquad (1)$$

The correlation between the two sources' equal-weighted CEFD measures is "only" 83%, however. Nevertheless, this is comfortingly close to a one-to-one relation with a high R^2. The correlations actually seem remarkably high, given that there is no standard as to which CEFs are included. Moreover, around the breakpoints where the two series overlap, we see

Month	Ross	LST	Predicted	Average
1980-01		16.96%		
1980-02	15.78%	18.85%	18.13%	17.31%
⋮	⋮	⋮	⋮	⋮
1985-12	2.10%	4.17%	3.85%	3.13%
1986-01	7.16%			

Can we average the two series? The difference between the Ross and LST series has a median of −0.30%, a mean of −0.44%, a standard deviation of this difference of +2.3%, and an interquartile range of −1.95% to +1.34%. Further, the typical month-to-month standard deviation for LST is 2.3%; for Ross 2.7%. Therefore, around the 1980 break, the average of 17.31% differs from the regression prediction of 18.13% by about 0.8%, which is a discrepancy of about 0.36 extra standard deviations in the time series between using an average and the prediction. The 1985 break point shows even less discrepancy. We also confirmed that none of our results is sensitive to omission of the observations adjacent to the breakpoints.

We negate the CEFD so that according to the LST theory, it becomes a positive sentiment measure—we can then talk unambiguously about improvements (increases) and deterioration (decreases). This measure is called bullish.cefd. We use the prefix "d." to denote the first difference of monthly values, and the suffix "vw." ("ew.") to denote value-weighting (equal-weighting). Therefore, our naming convention demands that we call the first differences in the negative of the CEFD d.bullish.cefd.vw and d.bullish.cefd.ew, for the value-weighted and equal-weighted discounts, respectively. In subsequent tests, we rely on

equal-weighted CEFD changes, because they tend to work better than value-weighted CEFD changes.[2]

[Table 1]
[Figure 1]

Figure 1 and Table 1 give a feel for this data. The figure plots the time-series of CEFD and its annual changes. The table shows that the typical d.bullish.cefd, both equal-weighted and value-weighted, in our sample was around 0%. Drops in the CEFD in excess of −8% occurred in March 1968, November 1976, and August 1998. Increases in excess of 7% occurred in November 1967, January 1974 and September 1998.

C The Consumer Confidence

Our proposed alternative proxy for investor sentiment, the consumer confidence, relies less on the sentiment theory and financial measures itself, but seeks to identify the proxy from direct survey questions.

Intrinsically, consumer confidence seems to be a concept similar to investor sentiment. Many investors are likely to be bullish about the economy when they are bullish about the stock market and vice-versa. To qualify as a proxy for investor sentiment, consumer confidence and investor sentiment must be positively correlated. Thus, a maintained—and later tested—hypothesis is that general exuberance would translate into both consumption and investment optimism. In the context of DeLong, Shleifer, Summers, and Waldmann (1990), the consumer confidence measure further requires an identification of consumers as the individual retail investors that are the relevant noise traders—or, more accurately, that the important marginal sentiment noise investors feel and act like the sampled consumers. We can not fully identify the abstract theoretical concepts, but we can examine the correlation among randomly chosen investors sampled in the UBS/Gallup survey (to see if there is a common factor across individuals), we can look only at the subset of relatively wealthy investors among the UBS/Gallup investors, and we can correlate changes in the UBS/Gallup measure with changes in consumer confidence. In sum, although the survey-based consumer confidence index is not without drawbacks as an investor sentiment measure, it does rely on two auxiliary hypotheses. Fortunately, we have the ability to directly correlate consumer confidence (and other hypothesized investor sentiment measures) with the UBS/Gallup proxy for investor sentiment to test these hypotheses.

[2] We also tried whitening the series. The following model fit best:

$$\begin{array}{lcccccc} \text{d.bullish.cefd.ew}_t & = & -0.256 & + & -0.262 \cdot \text{d.bullish.cefd.ew}_{t-1} & + & -0.0961 \cdot \text{d.bullish.cefd.ew}_{t-2} \\ \overline{R}^2 = 7.9\% & & (\text{T}=-1.67) & & (\text{T}=-5.48) & & (\text{T}=-2.03) \\ & & & + & -0.031 \cdot \text{bullish.cefd.ew}_{t-1} & + & \text{noise}_t \\ & & & & (\text{T}=-2.22) & & \end{array}$$

The R-square is moderate, but significant. All results in the paper are robust to using the whitened or the plain difference series. For comparability with LST, our paper relies on the latter.

D The Michigan Consumer Confidence Index

Our paper relies primarily on the *Michigan Consumer Confidence Index*, provided by the *Michigan Consumer Research Center*.[3] Post-1978, it is released monthly, enjoys great prominence, and is available from the vendor. Quarterly data from 1952 to 1977 is available as series UMCSENT1 from FRED (Federal Reserve Economic Data). For these months, we interpolated linearly. Interpolating reduces the power of the consumer confidence (CC) measure, but the results remain stable if we ignore the quarterly data (as we do in our post-1985 subsamples). The following descriptions of the Michigan Consumer Confidence Index borrows heavily from the websites of the provider and the *Market Harmonics* website. (A deeper description in the context of forecasting consumer spending and some other variables appears in Ludvigson (2004).)

The Consumer Confidence focuses on five questions:

1. *"We are interested in how people are getting along financially these days. Would you say that you (and your family living there) are better off or worse off financially than you were a year ago?"*
2. *"Now looking ahead—do you think that a year from now you (and your family living there) will be better off financially, or worse off, or just about the same as now?"*
3. *"Now turning to business conditions in the country as a whole—do you think that during the next twelve months we'll have good times financially, or bad times, or what?"*
4. *"Looking ahead, which would you say is more likely—that in the country as a whole we'll have continuous good times during the next five years or so, or that we will have periods of widespread unemployment or depression, or what?"*
5. *"About the big things people buy for their homes—such as furniture, a refrigerator, stove, television, and things like that. Generally speaking, do you think now is a good or bad time for people to buy major household items?"*

Answers are coded on a scale from 1 (good) to 5 (bad), and averaged (equal-weighted). The reported Michigan Consumer Confidence Index is a linear transform thereof. The survey methods themselves are described in detail by Curtin (2000). An excerpt:

> *The monthly survey of consumers is an ongoing nationally representative survey based on approximately 500 telephone interviews with adult men and women living in households in the coterminous United States (48 States plus the District of Columbia). The sample is designed to maximize the study of change by incorporating a rotating panel sample design in an ongoing monthly survey program. For each monthly sample, an*

[3] Our earlier draft also explored other sentiment and confidence indices in more detail. The results using the Consumer Conference board consumer confidence measure were similar, even though the Pearson correlation between Michigan *changes* and CB *changes* was "only" +52% (T-stat of 10.95). The Michigan index focuses more on financial conditions (and especially the individual's own condition), while the Conference Board index focuses more on macroeconomic conditions. Thus, for our purposes, the Michigan index is more suitable. Moreover, because the surveyed individuals do not overlap, the measures do seem to pick up some economy-wide sentiment factors, rather than merely the answers of just their own respondents.

independent cross-section sample of households is drawn. The respondents chosen in this drawing are then reinterviewed six months later. A rotating panel design results, and the total sample for any one survey is normally made up of 60% new respondents, and 40% being interviewed for the second time. The rotating panel design of the Surveys of Consumers has several distinct advantages over a simple random sample. This design provides for the regular assessment of change in attitudes and behavior both at the aggregate and at the individual level. The ability to gauge individual change expands the study of aggregate change by permitting a better assessment of the underlying causes of that change. The rotating panel design also permits a wide range of research strategies made possible by repeated measurements. In addition, the sample design supports the pooling of up to six of the independent monthly samples to achieve larger samples, or to screen for rare populations or events.

The telephone sample is obtained by list-assisted random sampling. A great deal of consideration has been expended on appropriate sampling. Further, the survey documentation gives great emphasis to demographic sampling, sampling error, sample coverage and non-response errors, sample weighting questionnaire design, telephone interviewing (and interviewer training), coding methods, and institutional independence. New survey respondents are rotated in only slowly. Interviews are spread rather evenly across the entire month, and the survey is never revised.

Figure 1 and Table 1 also describe the consumer confidence series. The figure plots the time-series of *annual* observations for our sentiment measures. There is no correlation between the Michigan Consumer Confidence Index and the closed-end fund discount, even in levels. Lemmon and Portniaguina (2004) work with levels in consumer confidence,[4] which are non-stationary. (The auto-coefficient in our data is 0.94, the auto-correlation is 96%.) This requires considerably more delicate statistics. Therefore, our paper works with sentiment changes, the first difference in the Michigan Consumer Confidence Index, called d.bullish.mich.cc. These changes are stationary.[5] [Table 1]

Table 1 shows that the average Michigan Consumer Confidence Index is around 90. The most bullish Michigan indices in the sample occurred in Feb 1998, and from Jan through May of 2000. The most bearish time occurred in April and May of 1980. d.bullish.mich has a month-to-month standard deviation of around 3. The Michigan Consumer Confidence Index had the most pronounced drops in Dec 1980 and Aug 1990, and the most pronounced improvements in Jan 2004, Mar 1991, and Nov 1992.

[4] The theory suggests that levels of sentiment correlate with levels of pricing, and therefore that changes in sentiment correlate with returns. It does not easily suggest that levels of sentiment correlate with returns.

[5] They are stationary, but they are not i.i.d. Lagged changes have no further influence, but lagged levels do. A time-series model suggests

$$\begin{array}{lccccc} \text{d.bullish.mich.cc}_t & = & -0.3719 & + & (-0.0444) \cdot \text{bullish.mich.cc}_{t-1} & + & \text{noise}_t \\ \overline{R}^2 = 7.2\% & & (\text{t}=-2.38) & & (\text{t}=-3.18) & & \end{array}$$

The adjusted R-square is modest, and more importantly, our results that follow are similar if we use the unexpected series instead of simple changes in the two sentiment indices. For simplicity, and analogous treatment relative to the CEFD, we use the differenced but unwhitened series.

II Direct Validation

Our paper's first goal is the validation of proxies for investor sentiment. Although sentiment in DeLong, Shleifer, Summers, and Waldmann (1990) is an abstract concept, an optimal empirical validation measure would be a time-series of complete surveys of all investors in the financial markets, perhaps value-weighted by portfolio size. This data is not available. However, the UBS/Gallup survey provides a reasonably plausible sampled substitute.

A The UBS/Gallup Validation Measure

The *UBS/GALLUP Index of Investor Optimism* offers a representative and carefully sampled survey of investors. This measure is constructed from surveys of (randomly chosen) investors with more than $10,000 in wealth. During the first two weeks of every month, UBS/Gallup conducts 1,000 interviews of investors. It reports the results on the last Monday of the month. (http://www.ropercenter.uconn.edu/ubs.html provides more detail.) Our own study does not rely on their raw index, but on one question from this index:

> 6c. Now, I would like to ask you to think about the factors that could affect the overall investment environment OVER THE NEXT TWELVE MONTHS. On the same five-point scale, as far as the general condition of the economy is concerned, how would you rate (read and rotate A-D) Performance of the stock market, OVER THE NEXT TWELVE MONTHS? (NOTE TO INTERVIEWER: Do NOT repeat the scale unless it is necessary to remind the respondent. If respondent says "optimistic" or "pessimistic", be sure to clarify if that is "very" or "somewhat").

Value	Label
1	very pessimistic
2	somewhat pessimistic
3	neither
4	somewhat optimistic
5	very optimistic
6	don't know
7	refused
8	no answer

Our paper codes a "very optimistic" as $+2$, a "somewhat optimistic" as $+1$, a somewhat pessimistic as -1, and a "very pessimistic" as a -2, and henceforth work with the average. Table 1 shows that the median month-to-month change in the UBS/Gallup poll was around 0.16. The most optimistic months were Dec 1999 to Feb 2000, the most pessimistic months were Jul and Oct 2002.

The biggest drawback is that although UBS/Gallup is careful to retain the same investor profile every month, it does not retain the same investors every month. This introduces noise into their measure. (In contrast, the consumer confidence measure tends to follow the same consumers.) Nevertheless, we believe that the UBS/Gallup survey is a good validator for other investor sentiment proxies today. (See Page 14 for more alternatives.)

Unfortunately, the UBS/Gallup survey has only been conducted beginning in October 1996, and with regularity only beginning in 1999. We have data until December 2002. The data series is too short for serious use in financial research.[6]

B Validation Evidence

[Table 2]

The top part of Table 2 presents the first important piece of evidence that our paper offers. It shows the correlation among changes in hypothesized proxies for sentiment. Even though respondents, questions, and resampling procedures differ, there is a factor that is common to the Michigan Consumer Confidence Index and to the UBS/Gallup investor sentiment surveys. The correlation between changes in consumer confidence and changes in the UBS/Gallup investor sentiment series is a positive +55.2%, which is highly statistically significant. The table also reports the UBS/Gallup correlation with two other confidence measures, the Conference Board measure of consumer confidence, and the Michigan survey of consumers about their confidence in future business conditions. Both measures seem to be valid, but poorer proxies for investor sentiment, which is why we do not entertain them beyond this table.[7]

In contrast, there is no correlation between changes in the UBS/Gallup survey sentiment and changes in the closed-end fund discount. The two do not seem to have a strong common component. The disconnect between CEFD based and survey based sentiment measure is not only in differences. Even in levels (both series are non-stationary), the two are different. (Correlations in levels are a necessary, but not sufficient validation of sentiment measure identification.)

Not reported in the table, we also correlated other sentiment proxies. Under the hypothesis that investor sentiment is not idiosyncratic, we can validate different measures of investor sentiment against one another. The Conference Board and Michigan consumer survey are from different respondents, yet their changes correlate positively with one another, too. In contrast, changes in the CEFD do not correlate at all with changes in either measure of consumer confidence. Therefore, if a researcher wants to continue to maintain that the CEFD has a component that measures investor sentiment, such a researcher must posit that it is a component that would not be (strongly) reflected in the UBS/Gallup investor sentiment poll or in consumer confidence.

It is worthwhile to discuss potential shortcomings of our direct measure procedure:

[6] In addition to its better sampling method (because it measures changes for the same individual over time), consumer confidence measures are available in many countries, while the UBS/Gallup survey is also available only in the United States. And only the consumer confidence survey is available with a long, stable, regular, and professionally executed history, suitable for financial research. Note: As of April 2006, the 2003 disaggregated data is not yet available from the Roper center.

[7] The questions in the Conference Board measure are also less focused on consumer's financial situations than those in the Michigan measure. This is a good ex-ante reason to focus on the latter in our context, rather than the former.

Bias: We rely on changes in sentiment measures, which means that our concern would not be systematic bias, but only time-variation in the systematic bias—and such time-varying bias changes would have to be correlated across different respondents and surveys. Moreover, Gallup and the Michigan Center are professional pollsters, who select respondents appropriately randomly. Therefore, time-varying bias from these two sources is unlikely to be an important problem.

Noise: There are at least three sources of noise. First, UBS/Gallup may sample the wrong type of individuals. Second, respondents may not pay much attention to answering the survey. Third, the identities of respondents changes, and respondents are not asked about their own changing attitudes.

Therefore, a finding of no correlation could not only have meant that the Michigan Consumer Confidence Index is not a proxy for investor sentiment, but also that either the UBS/Gallup measure or the Michigan Consumer Confidence Measure or both are just noisy and thus so flawed as to be unusable. However, the two measures can validate one another. If either were not correlated with an underlying "generic sentiment" factor, then they should not correlate with one another. After all, they come from different respondents. Therefore, our finding of a strong correlation of changes in the Michigan Consumer Confidence index with changes in the UBS/Gallup poll is difficult to explain away as noise. Semantically, one could even define the component that is joint to the UBS/Gallup investor sentiment and the Michigan Consumer Confidence Index as generic "sentiment."[8]

Investor Identification: A reasonable critique of survey sentiment indices is that they may unduly measure the optimism of small investors, which are not of importance to the stock market (although the DeLong, Shleifer, Summers, and Waldmann (1990) and Lee, Shleifer, and Thaler (1991) rely on identification of noise traders with small investors). Put differently, why should we believe that how the retiree in Mississippi changes her views should matter in any way to how the wealthy New York city investors change their perspectives?

Question (S5) of the UBS/Gallup survey provides a classification of investors with more or less than $100,000 in stock and bond investments (henceforth termed "wealthy" and "poor"). Thus, the UBS/Gallup data allows us to split the sample between wealthier investors, who declare themselves with assets above $100,000, and poorer investors. The total numbers are 22,687 "wealthy" investor-months and 29,987 "poor" investor-months. With the UBS/Gallup investor wealth classification, we could determine one index based on wealthy investors only, and another index based on poor investors only. The data suggests that wealthy investors tend to be more optimistic

[8] It is of course another question whether this is the sentiment that is priced in the financial markets. However, the theory does not suggest that such investor sentiment should be orthogonal to generic sentiment. Moreover, if there are multiple dimensions of investor sentiment, then we need new theories. Moreover, if there is not a single factor but heterogeneity in multiple sentiment measures, then overall sentiment would likely wash out and thus play only a very small role in financial markets.

(mean 0.41, median 0.51) than poor investors (mean 0.30, median 0.34). This difference is statistically highly significant. However, this difference is not important within our context.

More importantly, the correlation between changes in the survey answers across the two investor groups is 80.4%, again indicating that the indices can measure some underlying factor, and not just respondent-specific noise.

The lower part of Table 2 brings more empirical evidence to bear on the relevancy of our sampled consumers. From the perspective of pricing in financial markets, it would be the wealthier investors in whom we may be most interested. Table 2 shows that the changes in the Michigan Consumer Confidence Index are as highly correlated with changes in sentiment in the wealthy subsample as they are with changes in sentiment in the poor subsample. (Again, changes in the CEFD have no correlation with either investor subset.)

A related reasonable critique of survey sentiment indices is whether it is primarily idiosyncratic and investor specific, or whether there is truly a measurable common component that transcends individual investors. We already have some of this evidence, because we know that measures from different surveys contain a common factor. We can further confirm this among individual investors, too. We proceeded to bootstrap 10,000 random distributions of investors, in which we randomly identified 29,987 and 22,687 investors as poor and wealthy respectively, recomputed two indices, their first changes, and their correlation. The mean (median) correlation was 83.2% (83.5%). The evidence thus suggests that changes in sentiment are not exclusively investor specific, but systematic. In the bootstrap, the empirically observed 80.4% correlation between the wealthy and poor investors sits at the 23rd percentile. Therefore, we can conclude that *there is no difference in how poor and wealthy investor sentiment changes month-to-month. Wealth is not a determinant of sentiment changes.* The critique that investor sentiment indices based on poor investors would not accurately reflect the investor sentiment of wealthier investors can thus be rejected. More broadly, it is unlikely that investor sentiment changes of a different group of unsampled investors would not be picked up by our sampled investors.

In sum, we have established that the CC correlates well with the UBS/Gallup investor sentiment measure, while the CEFD-based sentiment proxy does not. In order to continue considering the CEFD as an investor sentiment measure, an auxiliary assumption must be that the (relevant) noise investors are not only systematically different from those sampled by UBS/Gallup, but that the common factor, which appears in consumer confidence, the UBS/Gallup investor sentiment, and the UBS/Gallup poor and wealthy subset investor sentiments, is not sufficiently strongly related to the true measure of investor sentiment—or it should have shown up in CEFD changes. We view this assumption as unlikely to be true.

Direct Validation vs. Joint Hypotheses: Although we will also look at the Lee, Shleifer, and Thaler (1991) main theory-based validation test in Section III—the performance in

predicting small firm returns[9]—our paper considers the correlation of different proxies for sentiment to be a more powerful validation test (i.e., relative to the correlation of aproxy of sentiment with the size decile return spread.) The LST test is a joint test that must rely strongly on the theoretical conjecture itself. Without direct validation, a proxy must lean only on theoretical conjecture. Any test relying on such a proxy is a joint one. Not finding an effect could be attributed to a fact that the sentiment measure is poor, or to a fact that sentiment plays no role. The theory becomes practically unfalsifiable. Even though a correlation with sentiment is a necessary condition in DSSW, lack of finding a pricing correlation could be rationalized away as due to a bad sentiment proxy. Moreover, a positive finding is easier to be doubted by classical finance proponents—the sentiment proxy may measure something different.

With more direct validation for the measure of investor sentiment, subsequent tests become more informative. Even a finding that a sentiment measure plays no role in financial markets can now become interesting and interpretable. This is why we deem subsequent tests for the role of sentiment that use a directly validated sentiment measure to be fundamentally better—though they are definitely not perfect.

C Other Sentiment Measures

The common use of the CEFD, both in itself and as a component in other sentiment measures, explains why our paper is especially interested in the CEFD as the right "bogey" to begin its exploration of the literature on investor sentiment. Other sentiment measures proposed to-date in other papers have remained "boutique measures," which have not found wide usage beyond authors' own papers. (Our consumer confidence currently falls into this class, too.) It is of course impossible to compare consumer confidence to all of these series. Many are not easily available, either, and few have been validated as measures of sentiment, rather than just hypothesized and used.

There are however a number of other measures of investor sentiment in the literature that deserve mention. We explored some alternative measures, such as the happiness index of the General Social Science Survey, the Luxury Consumption Retail Sales Growth index (Ait-Sahalia, Parker, and Yogo (2004)), and Shiller's Investor 1-Year Confidence Index. None of these were available regularly enough or showed good validation characteristics.

Fisher and Statman (2003) explore the AAII survey of individual investors. Like the UBS/Gallup poll (and unlike the consumer confidence poll), it does not track changes in the same individuals over time. Unlike the UBS/Gallup poll, it allows members to vote whenever they feel like it. Thus, the AAII survey has an unusually severe self-selection problem. Further inspection reveals that the series is very volatile. Over the 47 months in

[9]We do not consider the CEFD correlations among different CEFs. Multiple factors could create a positive correlation among the two CEFDs. We also ignore CEF startup validation. As noted, our earlier draft could not detect an influence here, although it did find mild positive influence of the consumer confidence index.

which we had data for both, the AAII index has a −12.5% correlation with the UBS/Gallup survey. In its 190-plus months of overlap, changes in the AAII sentiment measure are not significantly correlated with changes in the CEFD, the Michigan Consumer Confidence Index, or the Conference Board consumer confidence index. There is a common factor among the UBS/Gallup indeics and the two consumer confidence surveys, but it is not shared by the AAII index. In sum, the AAII index likely measures changing heterogeneity in reporters more than changes in investor sentiment.

Baker and Wurgler (2004) construct an annual sentiment index based on a number of series hypothesized to relate to investor sentiment. Unfortunately, because of its periodicity (UBS/Gallup has only 3 years of monthly data, the BW index is annual), the UBS/Gallup survey cannot be used for validation. However, we can look at the correlation between 36 years of consumer confidence, the CEFD, and the BW index. Even levels of the BW sentiment do not correlate with consumer confidence levels. (The same applies to changes.) Therefore, whatever common sentiment factor appears in the consumer confidence index, it is not picked up by BW. The CEFD is a part of the BW index, and therefore has a 50% correlation in levels with the BW index, and a 26% correlation in changes. Therefore, some common factor of CEFD—and prior evidence leads us to believe that it does not contain sentiment—explains part of the results of BW. Other constituents of BW can be related to sentiment, or to other factors. (For example, IPO activity is closely related to recent stock performance, especially within specific industries. It thus picks up a momentum effect.)[10]

III The Role of Sentiment in the Financial Markets

In light of the evidence above, we believe that it is relatively uncontroversial that

1. the closed-end fund discount is not a proxy for investor sentiment;
2. the consumer confidence index contains a component related to investor sentiment.

The rest of our paper explores the link between sentiment and financial prices. The DeLong, Shleifer, Summers, and Waldmann (1990) prediction is that stocks disproportionally held by noise traders are disproportionally subject to investor sentiment. We therefore explore the following hypotheses:

1. Because small stocks are disproportionally held by noise traders, sentiment changes should change the spread between small decile firms and large decile firms. This is the main test in Lee, Shleifer, and Thaler (1991)

[10] A referee suggested measures based on financial discrepancies, such as the Palm/3COM split. If such a violations series were possible to construct, it would be less a proxy for common investor sentiment (a model input), but more a measure of irrational pricing in financial markets (a model output). It would also likely contain variations in agency or variation in trading costs that are less related to investor sentiment, which would have to be controlled for. (Moreover, the CEFD itself does not seem to be a violation of the law of one price; on the contrary, it may even exactly measure fund fees.) Most importantly, we do not believe that it is feasible to construct such a "financial irrationality index" on a regular and reliable basis.

We try to refine the LST test in two dimensions. First, we try to bring some evidence to bear on the question of whether the sentiment pathway works primarily through the financial markets or primarily works through the real markets. Therefore, we try to separate out the component of consumer confidence that is likely to be reflected in future corporate earnings. We extract a component of changes in consumer confidence that is related to current and future underlying economic conditions, specifically, aggregate consumption and aggregate corporate profits, and one that is not. We also have a contemporaneous measure of business confidence from a survey of CEOs. Managers' confidence should be related more to their perspective of future corporate performance than consumers' confidence. If the pathway works primarily through the future profits of firms, we would expect CEO business confidence to pick up more of this effect than consumer confidence.

2. Sentiment changes disproportionally do (or do not) influence small stocks through their effect on real consumption, profits, or corporate performance.

Second, we try to improve on the LST measure of which stocks are held by noise traders. We distinguish not only between stocks that are small vs. large firms, but also between stocks that are liquid and held by institutional investors vs. stocks that are illiquid and held by retail investors. A real-markets pathway perspective would suggest that we should find bigger effects across firm-sizes than across financial holdings/tradings characteristics. A financial-markets pathway perspective would suggest the opposite.

3. Sentiment changes disproportionally influence stocks held not by institutional but by retail investors, especially if these stocks have insufficient liquidity to allow arbitrageurs to impose more rational pricing.

This section ends with a discussion of what we can and what we cannot claim in light of our evidence: Finding a link between sentiment and financial pricing is a necessary condition in the context of a behavioral (DeLong, Shleifer, Summers, and Waldmann (1990)) view of the world, but not a sufficient one.

A The LST Validation Metric: Excess Small Stock Returns

[Table 3]

The small stock return spread is the difference between the rates of return on the smallest-capitalization stocks and the largest capitalization stocks, based on the well-known CRSP decile portfolios. This variable is called smallstocks.retspread. The panel at the bottom of Table 3 shows that small stocks outperformed large stocks in our sample period on average, but not in the median month. The typical month-to-month standard deviation was around 7%.

Table 3 explains this small-firm return spread. In the full-sample regression, changes in both the Michigan and the CEFD series have statistically significant contemporaneous explanatory power. (The two measures are almost uncorrelated, which means that the

coefficients and significance levels on one remain practically the same if we exclude the other.) Over the full sample period, the CEFD slightly edges out the Michigan CC in terms of significance. A one standard deviation change in CEFD changes (2.18%) associates with a $0.734 \cdot 2.18\% \approx 1.6\%$ change in the small firm excess spread. A one standard deviation change in CC changes (3.27%) associates with a 1.3% change, just a little less. Both numbers are an economically meaningful one-fifth of the small firm spread's standard deviation of 7.1%.

However, the subsamples show that the CEFD has lost its contemporaneous explanatory power for the small stock return spread after 1985. The CEFD's sign remains positive, but the significance drops dramatically. In contrast, the Michigan Consumer Confidence Index survey actually improved in its ability to explain the small stock return spread.

The bottom half of Table 3 excludes January observations, long known to be peculiar. The closed-end fund discount seems to have explanatory power only if January observations are included, even prior to 1985. Without January observations, only the Michigan Consumer Confidence Index remains statistically significant at the 5% level. Not in the table, the January-only CEFD coefficient is economically and statistically significantly higher than the non-January CEFD coefficient. Consequently, even the LST sample itself suggests theories in which the CEFD would act as a sentiment proxy *primarily* in Januaries, and less during the rest of the year.

In sum, we interpret our evidence to suggest that the Michigan Consumer Confidence Index exerts an influence on the small firm stock spread. The closed-end fund discount does not.

B The Real Economy Pathway: Consumer Confidence Decomposed

We now want to determine the pathway through which the consumer confidence index influences financial markets. We decompose consumer confidence into a part that correlates with real activity, and a part that is orthogonal thereto. Although there is no generally accepted model of how consumer confidence changes would translate into differential returns on small vs. large firms, it is reasonable to presume that consumer confidence would work through a chain that would begin with an influence on future consumer expenditures (explored reversely in Carroll, Dynan, and Krane (1994), Acemoglu and Scott (1994), and Souleles (2004)), and such a chain would end with an influence on future corporate profits. We therefore obtained from the *Bureau of Economic Analysis* (National Income and Product Accounts, NIPA) the seasonally-adjusted inflation-adjusted series on U.S. consumption and U.S. corporate profits.[11] Unfortunately, macroeconomic data is available only quarterly, which requires us to run regressions with quarterly re-sampled data (e.g., the consumer confidence change is now computed relative to confidence three months earlier). [Table 4]

[11] We also tried a number of other models, e.g., using the three components of aggregate consumption. None had an important effect on our inference.

Panel A of Table 4 explores how changes in the Michigan Consumer Confidence Index correlate with contemporary and future corporate profits and aggregate consumption. Models 1 and 2 suggest that, although there is a strong positive correlation between contemporary corporate profit changes and consumer confidence changes, consumer confidence changes do not appear to predict future corporate profit changes. Models 3 to 5 suggest that changes in consumer confidence predict consumption changes for up to two quarters ahead. Finally, Model 6 incorporates both profit and consumption effects. (Not reported, further leads or lags on changes in consumption or changes in corporate profits are insignificant.)

We can consider the fitted regression values from Panel A as the component of the Michigan Consumer Confidence Index changes that is related to real economic activity, and the regression residuals as the component that is not related.

Panel B of Table 4 predicts the small firm decile excess spread, beginning with a base regression in which we merely include quarterly changes in the undecomposed consumer confidence. The base regression shows that we lose almost no power going from monthly to quarterly data. The CEFD T-statistic drops from 4.85 to 3.47, the CC T-statistic increases from 3.89 to 4.09. (Our results are similar to those in Table 3 if the sample starts in 1985 or if Januaries are excluded—the CEFD becomes insignificant.) Of primary interest to us are the differential correlations of sentiment changes that influenced financial prices through their influence on the underlying real economy, vs. those sentiment changes that had no such link. The table shows that both pathways play a role. The coefficient estimates are generally similar, suggesting that it makes relatively little difference in pricing whether sentiment changes are attributable to the underlying real economic situation or not. The exception is Model 3, which controls only for contemporaneous consumption changes. In this case, the real-economic sentiment effect is stronger. However, this barely changes the significance of the non-real-economy related consumer confidence component, and the stronger effect disappears if we add future consumption changes.

In sum, Table 4 suggests that both the component related to the underlying real economy and the component not related to the underlying real economy in the Michigan Consumer Confidence correlate with the decile return spread. Thus, it appears that there is both a real pathway and a financial pathway. The influence of consumer confidence on stock returns does not appear to be exclusively one or the other.

C The Real Economy Pathway: Competing with CEO Confidence

The aggregate consumption and profit link is reasonable, but it is not a generally accepted model of how increased consumer confidence is likely to translate into differential corporate present values, based on future earnings. Under one additional assumption, we can consider an alternative method. The *Conference Board* publishes a quarterly measure of CEO confidence. If we assume that CEOs are relatively more attuned to the real business conditions that apply to the profits of publicly traded companies in our sample, then we

can control for real changes in implied future profits by including CEO confidence in the regression.

More precisely, our measure of CEO Confidence is the quarterly survey of U.S. chief executives. The *Conference Board* began this program in 1976. Questionnaires (along with results of the prior quarter's survey) are mailed out in the second month of the quarter to about 800-1,000 chief executives of U.S. corporations, of which 100 typically respond. These companies are associates of the Conference Board. They are asked four questions: two on their assessment of the U.S. economy generally, and two on their own industry specifically. In addition to these regular questions, one or two special supplementary questions are asked each quarter. The conference board reports for each question a weighted sum of the percentage of respondents checking each response, as follows:

% saying SUBSTANTIALLY BETTER	x	1.00
% saying MODERATELY BETTER	x	0.75
% saying SAME	x	0.50
% saying MODERATELY WORSE	x	0.25
% saying SUBSTANTIALLY WORSE	x	0.00

The main index reported by the Conference Board is the artithmetic average of three (later four) questions ([a] current general economic conditions vs. six months ago; [b] current economic conditions vs. six months ago for the CEO's own industry, expectations for the economy six months ahead; and [c] expectations for own industry six months ahead). We use first differences in the main index (d.ceo.conf), and first differences in the expectations of economic conditions six months ahead (d.ceo.exp). [Table 5]

Panel A of Table 5 again explores simple correlations, primarily for validation of the measures. The UBS/Gallup survey still correlates strongly with the CC and not at all with the CEFD on a quarterly basis. There is also a strong correlation between changes in CEO confidence and changes in the Michigan CC—again, a common factor in participants of different surveys. Unfortunately, we only have 15 quarterly observations with data on both the UBS/Gallup measure and the CEO confidence series, so this validation cannot be done. It is positive, but not statistically significant.[12] By now not surprisingly, there is no positive correlation between changes in CEO confidence and changes in CEFD.

In sum, there is some evidence of a common factor in investor sentiment, consumer confidence, and CEO confidence, all from independent sources; and no evidence of a common factor in CEFD changes and various survey sentiment measures.

Panel B again predicts the extreme size decile return-spread. Controlling for changes in CEO confidence, changes in the Michigan CC retain the same effect on the small firm return spread. This suggests that the pathway through which the Michigan CC influences the return spread is not its correlation with the CEO corporate outlook.

[12] In 2006, we obtained aggregated series from Gallup up to 2005, which does not help with other results (due to lack of series overlap). Here, however, it does help. The correlation with d.ceo.exp increases to 29% and is highly statistically significant.

D Retail-Held and Illiquid Stocks or Small Stocks?

The prediction of DeLong, Shleifer, Summers, and Waldmann (1990) is that investor sentiment influences stocks held disproportionately by noise traders, which cannot easily be arbitraged by more sophisticated investors. From the theory's perspective, the firm-size decile spread is a noisy proxy for this type of stock. After 1980, we can use 13(f) data to identify stocks that are not held by institutional investors. These stocks should fit the theory better.

Our first attempt was to construct portfolios as follows. In the retail portfolio, we included only firms for which we could not identify holdings by an institutional investor who had filed a 13(f), and which ranked in the bottom half of dollar trading volume within this zero-holdings group. (This is required in DSSW to make inefficiencies more difficult to arbitrage.) In the institutionally-held portfolio, we allowed not only the two top institutionally held deciles, but also only those firms that rank in their decile's top half of dollar trading volume. (The appendix describes the construction of our retail-held illiquid stock vs. institutionally-held liquid spread portfolios.)

The cleanest way to attribute differential effects to the three portfolios is to run a reverse regression, in which the sentiment proxies become our dependent variables. We also entertain the unexpected component of consumer confidence from the final model from Table 4, Panel A, as a sentiment proxy. The rates of return on three portfolios—the S&P500, the small-firm spread portfolio, and the retail-held spread portfolio—are our independent variables. The results are

	Consumer Confidence				Closed-End Fund Discount	
Dependent:	d.bullish.mich		unexp.bullish.mich		d.bullish.cefd	
constant	−0.227	−1.03	−0.130	−0.60	−0.074	−0.61
smallstocks.xret	0.113	2.80**	0.102	2.55*	0.025	1.14
retailstocks.xret	7.948	1.87+	8.247	1.96+	2.043	0.86
sp500.pchg	0.225	4.42***	0.220	4.37***	0.105	3.71***
$\overline{R}^2$	9.8%		9.2%		4.5%	
df	281		281		270	

This evidence suggests that Michigan Consumer Confidence Index changes are correlated on the margin primarily with the S&P500, then the small-firm stock spread, and finally the retail-stock spread.

However, this evidence is deliberately not in a proper table, because we do not deem it to be robust. When we subsequently explored various dependent dual-sorts, in which we tried to enhance the distinction between small/large firms and retail-held illiquid/institutionally-held liquid firms, we generally find no significance at ordinary levels. There is a good reason. Almost all large firms are held primarily by institutional investors, and very few small firms are held primarily by institutional investors. Therefore, our tests here are fragile. The

correlation between the excess rate of return between the small-firm spread portfolio and the retail-held spread portfolio in the above table is +51%. (Moreover, whereas the small stock spread portfolio and the S&P500 had almost no mutual correlation, the retail-held spread portfolio and the S&P500 had a −40% correlation.) Thus, although our inability to improve our sort may not be surprising in light of the high correlations in portfolio returns, it leads us to conclude that we do not have good power to disentangle the influence of (real-economy unrelated) consumer confidence on small stocks vs. retail-held stocks.

E CAUTION

The evidence in the above sections suggest that consumer confidence changes, both those related and those not related to underlying economic conditions, correlate contemporaneously with the return spread between small firms and large firms. The interpretation of this evidence requires extra care, because there is considerable academic sentiment when it comes to interpretations of investor sentiment.

Rationality: It is necessary to explain why neither our (nor any other existing) paper can resolve whether this effect is rational or irrational.

Even if financial markets are hyper-efficient in terms of securities pricing, (irrational) sentiment could play a role. It is not difficult to write down a model in which underlying economic factors influence agents, which are both investors and consumers, and turn them more bullish or bearish both about the real and the financial economy. Bullish sentiment could induce investors/consumers to make real choices (e.g., different purchasing/saving decisions). Such agent behavior must have implications in both the real and the financial markets. Even if sentiment were entirely irrational, in a financial market in which prices are set perfectly rationally, consumer sentiment should therefore have rational pricing consequences. In sum, attributing the effect of consumer sentiment on future consumption and some correlation between stock prices and future consumption to irrational financial markets could be unwarranted.

Conversely, it would not be difficult to write down a model in which it is the irrational behavior and instincts that drive the correlation between the real economy and the financial economy. Sentiment could lead both consumers and investors to overreact, inducing a correlation between consumption, consumer sentiment, and stock prices, even if the present value of corporate profits were to remain mostly unchanged. Attributing even the part of the effect of consumer sentiment on stock prices that comes from future consumption changes to rational financial market pricing could be equally unwarranted.

Some distinctions may be outright philosophical. Sentiment could change investors' effective risk-aversion, too. This should also change the discount rate, and whether such changes should be considered rational or irrational is a philosophical question.

The only definitive way to use the evidence to discriminate between the classical rational and the behavioral irrational perspective is to have either perspective provide specific

predictions of exactly how much investor sentiment (or a proxy thereof) should influence financial prices.[13] It requires not just directions, but quantitative predictions. In the absence of quantitative theories, proponents of either perspective can claim for themselves the large part of any effect that overlaps with both perspectives. Until we have such theories, no paper can fully confirm or reject one of the two perspectives in favor of the other.

The fact that we cannot resolve whether the sentiment influence is rational or irrational does not mean that it is uninformative. Our paper did establish a necessary, but not a sufficient condition for the perspective advanced in DeLong, Shleifer, Summers, and Waldmann (1990). The fact that the pathway is not exclusively a real economy phenomenon may further mildly favor the behavioral perspective.

Causality: Our paper has no evidence on causality. A third factor, e.g., a component of GDP, could influence both financial returns and sentiment indices. On some level, this is semantics—this component could be called (derived) generic sentiment, and it would be picked up by our proxies. The direction of the link is also not determinable. For example, if good portfolio performance, especially of the small-firm components within the broader portfolio, cause investors to be more exuberant, then the small firm return itself becomes a proxy of sentiment—a different mechanism through which survey measures and small firms could be correlated.[14]

IV Other Findings

Three further empirical findings are worth reporting.

A The Cross-Section of Stock Returns

We experimented with whether the sentiment factor is priced in cross-section. That is, we wanted to see whether loading on the sentiment factors can explain the cross-section of returns. To explore this, we downloaded the Fama-French factors and the 25 Fama-French portfolios from Ken French's website.

[13] If researchers could identify the component of stock prices that was irrational, it would solve this dilemma. Alas, this is an even taller order. Stock returns are too noisy to allow us to reliably and convincingly measure "overshooting," at least as of yet.

[14] The financial markets correlations hypothesized in DSSW and LST are contemporaneous, and the link could even be so direct that it is almost mechanical: For example, if small decile stocks go up and increase the wealth of small investors, retail investors could use the returns to purchase more CEFs and thereby narrow the CEFD. Or, if the CEFD narrows, retail investors' portfolio would do relatively well, and could buy more small firms. However, this is not the case in the data. Moreover, even among retail investors, few portfolios are likely to be so small firm biased that the small firm effect would overwhelm the market rate of return in terms of wealth significance. Moreover, a referee wondered whether the wealth effects could be strong enough to explain all the correlation. This is unlikely, because we explain the zero-investment decile net spread, not the positive-investment market return. Even though retail investors do have a small firm bias relative to institutional investors, the spread is likely to be the a less important determinant of most retail investors' portfolio performance than the market rate of return. (The possible exception may be the Tech bubble in the late 1990s.) We do not have individual investor data to test this further.

Changes in the Michigan Consumer Confidence Index have a 16.6% correlation with SMB, no correlation with HML, and a 17.8% correlation with the FF excess rate of return on the market. (This is why it is necessary to control for the market rate of return in Table 3.) In contrast, the CEFD has 13.3% correlation with SMB, 8.1% correlation with HML, and 11% correlation with the market. The Michigan correlations remain the same in the subsample post 1985, but the correlation of the CEFD index change. The CEFD loses its correlation with both SMB and HML, and retains only its correlation with the excess market rate of return (which increases to 25%).

We then followed a Fama-MacBeth type procedure. We formed exposures to the factors for each of the 25 Fama-French portfolio over 5 years, and then correlated these with the cross-section of portfolio returns net of the risk-free rate in the following month. We ran this test for 390 months (1965/08 to 2002/12), and report average "gamma's" and their time-series T-statistic. This sample period is known to not have been overly favorable to size, and includes the Tech bubble period, which was not favorable to value firms.

Dependent Variables: 25 Portfolio Returns in Each X-Sectional Regression

Time-Series Statistics of Gamma's from 360 Monthly Regressions

		constant	$M - r_F$	SMB	HML	Mich	CEFD
Model 1	Sentiment Only	0.73				0.55	
	T-stat	(2.6)				(1.1)	
Model 2	Sentiments Only	0.63				0.14	0.96
	T-stat	(2.0)				(0.3)	(2.5)
Model 3	All Factors	1.39	−0.76	0.52	0.37	0.39	−0.22
	T-stat	(5.0)	(−1.3)	(1.8)	(1.0)	(0.8)	(−0.9)
Model 4	Fama-French Factors	1.49	−1.14	0.41	0.17		
	T-stat	(5.1)	(−2.2)	(1.6)	(0.5)		

The results suggest that portfolios that correlate more with changes in consumer confidence do not earn a higher rate of return than stocks that correlate less. Interestingly, although the closed-end fund discount is not a proxy for investor sentiment, portfolio exposure to changes in the CEFD do appear to be a priced factor if SMB and HML are not included. This may, however, partly explain why the CEFD remains a popular factor in academic research.

B The Time-Structure of Correlation of Sentiment and the Overall Stock Market

It is also reasonable to ask not just whether sentiment correlates with the stock spread, but also whether it can correlate with the overall stock market rate of return at different leads or lags. The channel may be behavioral or classical (and, unlike for the spread portfolios, wealth effects can play a good role in this correlation). Indeed, sentiment should not fall like manna from heaven. Instead, any reasonable sentiment index should be influenced by

recent positive stock returns—and especially recent high overall stock market (portfolio) returns, and have (mild) persistent effects on return spreads. The following are the t-statistics on the correlations between changes in our sentiment measures, and two monthly rate of return series:

	Size-Decile Based Small Stock Return Spread										
Lead/Lag of	−5	−4	−3	−2	−1	**0**	1	2	3	4	5
d.bullish.mich	−1.7	−1.3	−1.5	0.0	1.6	3.9	1.7	1.8	−0.7	1.0	0.4
d.bullish.cefd	−0.5	0.9	−0.9	−0.5	1.0	4.7	−0.7	−1.3	−1.6	1.1	2.1
	S&P500 Percent Change										
Lead/Lag of	−5	−4	−3	−2	−1	**0**	1	2	3	4	5
d.bullish.mich	0.6	−1.5	−0.9	0.5	0.5	3.8	5.2	3.6	−1.2	−1.1	1.3
d.bullish.cefd	1.4	−0.6	−0.2	−1.0	0.5	2.4	0.2	−1.9	−0.9	0.6	0.3
	(sentiment anticipates return)						(return anticipates sentiment)				

In a Granger causality sense, significant numbers on the left imply that the sentiment index predicts (influences) the return, numbers on the right imply that the sentiment index is predicted (influenced) by the return.[15] The market-adjustment in the middle panel is done by hedging out the in-sample S&P500 return via regression. (The numbers are similar for other retail spreads.)

The CEFD series' inflluence has very little persistence. The effects of changes in the Michigan Consumer Confidence Index seem to both be influenced by the lagged small stock return spread and influence the future small stock return spread. The relationship is even stronger for the influence of past market-wide rate of returns. We would argue that the consumer confidence correlation patterns are desirable characteristics for an investor sentiment index.

Durell (2001) explored similar questions, primarily related to issues of long-term correlations of overall stock market returns with consumer confidence indices. The paper is different from our own not only in emphasis (he explores the relation between the market and the consumer confidence index in more detail than our one subsection here), but also in some findings—not all his findings are similar to our own. From our perspective, most importantly, he finds a theoretically reasonable correlation between one component of the Conference Board consumer confidence index and the CEFD. This—and some other mild differences in results—may be partly due to differences in specifications (he tends to use longer-term returns), partly due to data (he uses only the Conference Board index and only one component thereof), and partly due to sample period (he has overlapping data for the CEFD and the Conference Board for only seven years, 1978–1985).

[15] We have nothing to say about true causality—whether sentiment (in the form of the CEFD or in the form of consumer confidence) "drives" financial markets or vice-versa. If anything, both may be driven by an underlying unmeasured variable, an abstract sentiment, or something even deeper. And we would also expect feedback loops: when the stock market drops, it may cause a drop in investor and consumer confidence, which in turn can cause a further drop in the stock market, and so on.

We also present some evidence of mild feedback effects between consumer confidence and stock returns—a desirable feature of a sentiment index. In sum, our evidence suggests that sentiment (as measured by the consumer confidence) plays a role in financial markets, but that the CEFD may be the wrong measure of sentiment.

C Out-of-Sample Data: Sentiment in the United Kingdom

Consumer confidence has the advantage that it is available in many other countries, which can serve as an out-of-sample test. Doukas and Milonas (2004) report that the closed-end fund discount fails to explain small stock excess returns in Greece. Although we do not have access to Greek data, we are able to do some preliminary exploration of our relationship in the United Kingdom, because Dimson, Nagel, and Quigley (2004) kindly made their U.K. decile portfolio return data (1955-2001) available to us. As with U.S. data, we compute an excess rate of return of small firms over large firms. There are 4 months in which the excess return exceeded 20%, among them 35% in November 1999, and 24% in January 2000. There was only 1 month in which the return was significantly below −10%, which was −19% in December 1999. (The next smallest excess returns were −10.6%.) These returns were about four (!) standard deviations off the series mean,[16] which leads us to believe that the 11/99 to 01/00 period was highly unusual and perhaps not representative. Therefore, it is probably appropriate to exclude these three months or winsorize them. As our proxy for the market rate of return, we use the rate of return on the FTSE index.

The European Commission publishes consumer confidence data, beginning in January 1985. This means that we only have 204 data points with both consumer confidence and stock return data to work with. The EC data contains not only the general consumer confidence indicator (CC, series 99), but also a financial situation indicator (FSI, series 01), and a general economic situation indicator (ESI, series 03). The change in CC has a mean of 0.04 and a standard deviation of 3.2; the change in FSI (ESI) has a mean of 0.10 (0.10) and a standard deviation of 2.6 (4.9).[17]

Reporting all coefficients in percent, explaining contemporaneous small firm excess returns, we find that

Explaining U.K. Small Firm Excess Returns With Changes in Consumer Confidence

Full Sample, 1955-2001	0.628	+	0.158 · ΔCC	+	(−5.76) · R_{FTSE}	+	ϵ
t-stat	(1.35)		(1.08)	+	(−0.59)		
excl. 11/99-01/00	0.498	+	0.308 · ΔCC	+	(−13.55) · R_{FTSE}	+	ϵ
t-stat	(1.23)		(2.43)**	+	(−1.61)		
winsorized at ±2σ	0.360	+	0.219 · ΔCC	+	(−7.10) · R_{FTSE}	+	ϵ
t-stat	(0.95)		(1.84)	+	(−0.90)		

[16] The mean excess return was 0.6% per month with a standard deviation of 6.5%.

[17] The EC also publishes forward looking statistics, and the results are reasonably similar to those reported below.

The t-statistics on changes in confidence improve to 2.00, 2.83, and 2.60 if we replace changes in consumer confidence (CC) with changes in the economic situation indicator (ESI). Changes in the FSI component are however less significant, contrary to our expectation. In sum, the U.K. "out-of-sample" evidence seems generally supportive of a similar influence of sentiment—with the exception of three months at the end of the Tech bubble period during which excess returns were four standard deviations out.

D Unreported Findings

We also explored a number of other relations, and some are worth a brief mention.

First, unlike LST, we found no correlation between the CEFD and CEF startups in the SDC data base. On an annual basis, there is some mild evidence that more CEFs start up when consumer confidence is bullish. Second, changes in GDP have only modest correlation, as do changes in the unemployment rates. CC Sentiment seems to turn bullish when unemployment decreases—a sensible correlation. Controlling for GDP and unemployment does not change the reported results. Third, the CEFD is really like a book-market ratio. We can ascertain that when the Fama-French book-market factor changes, so does the CEFD. (The CC does not correlate therewith.) Fourth, there is no interest rate or IPO startup correlation with the CEFD. However, consumers turned more bullish at times when interest rates ros.

V Summary of Findings and Interpretation

Our paper distinguishes sharply between validation of investor sentiment proxies through the UBS/Gallup investor sentiment survey, and such measures' correlations with financial market series.

1a. The closed-end fund discount (CEFD) cannot be validated as a proxy for investor sentiment, using direct UBS/Gallup investor sentiment survey data. In light of our evidence, we believe it to be inadmissible as a reasonable proxy for investor sentiment.

1b. The Consumer Confidence can be validated as a proxy for investor sentiment, using direct UBS/Gallup investor sentiment survey data. (The CEFD has no correlation with consumer confidence, either.)

The fact that respondents in different surveys—wealthy and poor investors sampled by UBS/Gallup, consumers sampled by the Michigan Consumer Research Center, consumers sampled by the Conference Board, and CEOs sampled by the Conference Board—answer in a positively correlated fashion would be hard to reconcile with the notion that these measures are noisy or irrelevant to some true proxy for generic sentiment. More likely, any true investor sentiment measure would correlate positively with the consumer confidence.

2a. The CEFD does not robustly correlate with the excess rate of return on small firms, either in non-Januaries pre-1985, or in any months post-1985. It thus fails the Lee, Shleifer, and Thaler (1991) validation. Thus, without even their own validation, use of the CEFD is based only on a theoretical conjecture, which we found to be inconsistent with the data.

2b. The consumer confidence index correlates with the excess rate of return on small firms. It thus passes the Lee, Shleifer, and Thaler (1991) indirect validation.

2c. The mechanism through which the consumer confidence influences financial returns operates neither exclusively through the real underlying economic markets, nor exclusively through the financial markets. Instead, both pathways seem to matter.

In sum, we interpret our evidence to suggest that it should be agreed that only the consumer confidence and not the CEFD qualifies as a proxy for investor sentiment. It should further be agreed that such investor sentiment does have a contemporaneous correlation with certain financial market prices, specifically the size decile spread. The evidence suggests that this effect is likely to come from both a financial and a real channel. This evidence should not be interpreted as being necessarily due to rational or irrational financial markets. Until either of these perspectives can offer exact quantitative predictions, it is unlikely that any paper can separate the two.

References

Acemoglu, Daron, and Andrew Scott, 1994, Consumer Confidence and Rational Expectations: Are Agents' Beliefs Consistent with the Theory?, *Economic Journal* 104, 1–19.

Ait-Sahalia, Yacine, Jonathan Parker, and Motohiro Yogo, 2004, Luxury Goods and the Equity Premium, *The Journal of Finance* 59, forthcoming.

Baker, Malcolm, and Jeffrey Wurgler, 2004, Investor Sentiment and the Cross Section of Stock Returns, Working paper, Harvard Business School and New York University.

Berk, Jonathan B., and Richard Stanton, 2004, A Rational Model of the Closed-End Fund Discount, Working paper, Haas School at the University of California at Berkeley.

Carroll, Christopher D., Karen E. Dynan, and Spencer D. Krane, 1994, Does Consumer Sentiment Forecast Household Spending? If So, Why?, *American Economic Review* 84, 1397–1408.

Chen, Nai-Fu, Raymond Kan, and Merton H. Miller, 1993, Are the discounts on Closed-end Funds a Sentiment Index, *The Journal of Finance* 48, 795–800.

Chopra, Navin, Charles M.C. Lee, Andrei Shleifer, and Richard H. Thaler, 1993, Yes, Discounts on Closed-End Funds are a Sentiment Index, *The Journal of Finance* 48, 801–808.

Curtin, Richard T., 2000, Surveys of Consumers, Working paper, Survey Research Center at the University of Michigan date obtained via email.

DeLong, J.B., Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann, 1990, Noise Trader Risk in Financial Markets, *Journal of Political Economy* 98, 703–738.

Dimson, Elroy, Stefan Nagel, and Garrett Quigley, 2004, Capturing the Value Premium in the U.K. 1955-2001, *Financial Analysts Journal* ?, forthcoming.

Doukas, John A., and Nikolaos T. Milonas, 2004, Investor Sentiment and the Closed-end Fund Puzzle: Out-of-sample Evidence, *European Financial Management* 10, 235–266.

Durell, Alan, 2001, Stock Market Expectations and Stock Market Returns, Working paper, Dartmouth College.

Fisher, Kenneth L., and Meir Statman, 2003, Consumer Confidence and Stock Returns, *Journal of Portfolio Management* pp. 115–127.

Lee, Charles M.C., Andrei Shleifer, and Richard H. Thaler, 1991, Investor Sentiment and the Closed-End Fund Puzzle, *The Journal of Finance* 46, 75–109.

Lemmon, Michael L., and Evgenia V. Portniaguina, 2004, Consumer Confidence and Asset Prices: Some Empirical Evidence, Working paper, University of Utah.

Ludvigson, Sydney C., 2004, Consumer Confidence and Consumer Spending, *Journal of Economic Literature* 18, 29–50.

Ross, Stephen A., 2005, *Neoclassical Finance*. (Princeton University Press Princeton, NJ).

Souleles, Nicholas S., 2004, Expectations, Heterogeneous Forecast Errors and Consumption: Micro Evidence from the Michigan Consumer Sentiment Surveys, *Journal of Money, Credit, and Banking* 36, 39–72.

Spiegel, Matthew, 1997, Closed-End Fund Discounts in a Rational Agent Economy, Working paper, Haas School at the University of California at Berkeley.

A Frequently and Referee Asked Questions (FAQ)

(This section is ultimately not for publication.)

Here are the earlier referee reports. Many of these questions are difficult to address in the main paper (or we had addressed them but the referee had overlooked them). The reports are almost unabridged:

A QJE (May 2005)

1. **(QJE #1): It seems to me that in order to evaluate the merits of a potential sentiment index one needs to have in mind the particular application for which the index is going to be used.**
   If every situation required its own choice of investor sentiment, how could one ever reject behavioral hypotheses? Hopefully, there is a meaningful universal sentiment, that does not wash out but has influence—and we think we do have some evidence that it exists.
2. **(QJE #1): For example, if one is interested in thinking about IPOs and SEOs as an arbitrage response to potential market inefficiencies, it might make sense to have an index based on obvious violations of the law of one price in new equity issues: e.g., one might try to count the number of instances in a year that there were flagrant violations like in the Palm/3Com situation, and use this as the right-hand-side variable in a regression to explain IPO or SEO volume.**
   We have sacrificed three chicken, and performed the rain-dance. Alas, we still cannot construct such a series. It still seems to us that this is infeasible. Moreover, we think it is the wrong measure. Therefore, we sacrificed three more chicken in our quest to obtain a model input measure (sentiment), not the suggested model output measure (financial mispricing). How does one test a theory that is about inputs→outputs with an output measure? See, it would not be absurd to believe that mispricing in financial markets is not driven by sentiment but by another phenomenon such as agency conflicts. (Moreover, the series of pricing violations would pick up changes in transaction costs more than changes in sentiment.) Therefore, the price violation series might not be a sentiment measure series, at all, regardless of whether sentiment matters or not. (Our revised draft also offers a better discussion now on Page 15 and its footnote.)
3. **(QJE #1 continued): From this perspective, a conceptual virtue of the CEFD measure is that, since it shares this law-of-one-price type of benchmarking, it is cleanly interpretable as being about obvious mistakes made by investors; that is, it is hard to interpret a 20% closed end-fund premium as anything but a mistake. Thus if it were the case that the CEFD measure were robustly correlated with both new fund formation as well as IPO and SEO activity more generally, we would have a pretty good story about these new issues being a response to clear-cut pricing errors. Of course, if this is not the case in the data, too bad, and the paper does an important service by pointing out that the CEFD is a less robust variable in this regard than might previously have been thought.**
   Totally false.
   - See preceding.
   - Changes in the CEFD are not a good proxy for changes in the violation of the law of one price. The differential is probably due mostly to a fund fee wedge (in addition to a liquidity issue wedge)—an agency issue. The CEFD is *not* a violation of the law of one price. Indeed, it may be exactly the opposite—confirmation of the law of one price, given closed-end fund fees.
   - We do not see a correlation between CEFD changes and IPO/SEO startups in the data. (We see no correlation between CEFD changes and other sentiment proxies, either.)
4. **(QJE #1): By contrast, the Michigan Consumer Confidence Index—as well as the apparently somewhat related UBS/Gallup index—cannot be so narrowly interpreted. As I read the Michigan questions, it seems like I could perfectly well interpret the resulting sentiment measure as capturing rational or quasi-rational time-variation in individual-investor risk tolerance, perhaps due to changes in expected future wealth or investment opportunities.**
   Serious?! If I understand this concern, it is "could the survey just measure *rational* monthly changes in derived risk-tolerance, which are induced by monthly wealth changes?" The argument has the right direction, but the order of magnitude would be totally wrong. It seems exceedingly unlikely that changes in a part of even a *semi-rational* investor's portfolios could be big enough to have a large influence on their derived risk aversion month-to-month.

- We predict the small firm zero-investment net spread, not the market rate of return. Indeed, we hold the market rate of return constant, and this is not what is loading! So, the usual overall portfolio wealth change is not what our spread variable can capture.
- Let's do a back-of-the-envelope calculation of what sort of implied risk tolerance changes one could get for a *rational* investor if our dependent variable were not a zero-investment return spread, but the positive-investment small firm return. The maintained assumption is a rational world. The stock portfolio would be a small part of life-time income and wealth (bonds, real estate, etc.). The decile-10 firms would be only a small part of our rational retail investor's portfolio (if economy-value-weighted, less than 1%). The month-to-month dollar return would be only a fraction of the value of the investor's small firms in the portfolio (say 5%). These fractions are multiplicative. The last multiplicative factor is how much the *rational*, derived risk-tolerance of an investor should *change*. In other words, how much more risk-seeking would a rational investor become, if lifetime net worth would increase in a month by even an astronomical 1%, say from $1 million to $1.01 million, and then back to $1 million? Such monthly changes in derived risk tolerance would seem trivial for anyone half-way rational with any reasonable utility function.

Of course, if investors are irrational and just feel wealthier or just feel luckier or just feel smarter, then an argument like this can work—*but this is the argument for investor sentiment*. If investors have a high risk tolerance and the next month a low risk tolerance, and this changes every month, philosophically this should be considered sentiment.

Moreover, our paper does not say that we can exclude a rational perspective. Sorry, in this context, that is without quantitative predictions by the behavioral or classical model as to exact return behavior, no paper can do this convincingly—although we concede that other papers set out using sentiment under the assumption that it measures irrational behavior.

5. **(QJE #1): The latter version of the story strikes me as particularly problematic for the authors' conclusions. It suggests a mechanical sort of hard-wiring-in the limit, that respondents to the survey are effectively just answering the question: "how did your personal stock portfolio do this month?". To the extent that this is the case, the Michigan index will of course appear to work well in explaining the contemporaneous returns of small retail-owned stocks. But it would be wrong to conclude, as the authors do, that this makes the Michigan index an appropriate measure of investor sentiment for other applications.**

   To the extent that this is the case, it is exactly what we consider to be sentiment! This is because immediate questions are not about the past portfolio performance, but about general outlook, and if the answers about outlook depended one-to-one on past portfolio performance, then past portfolio performance would be a measure of sentiment.

   More generally, we would expect that less unemployment and more wealth and a rising stock market would make naive retail investors more buoyant about the future. (In what [behavioral] theory would this not be the case?) Optimism and sentiment that is correlated must not fall like manna from heaven. If some proposed proxy does not correlate with any macroeconomic conditions, e.g., employment, recessions, and depressions, it would not exactly speak for this measure as a proxy for sentiment. And, again, sentiment can be rational or irrational; we have no single causal mechanism in mind. We further discuss this in Section III.E It specifically addresses the hard-wiring (which is unlikely to be a big factor).

6. **(QJE #2): There seems to be a wide variety of sentiment measures out there. A quick web search reveals many other "sentiment" measures, such as:**
   - **the ratio of odd-lot sales to purchases;**
   - **change in margin borrowing, from the Fed;**
   - **net mutual fund redemptions;**
   - **premia on "Primes" and "Scores" (used to trade on the American Stock Exchange, split the value of a stock into dividends and capital gains)**
   - **some notion of market liquidity, measured by lower bid-ask spreads, lower price impacts of trading, higher turnover, the number of advancing to declining issues;**
   - **each of the five non-CEFD measures in the "BW Sentiment Index" short interest;**
   - **implied volatility of S&P 500 index options (the "VIX") and the similar Nasdaq-100 VXN computer database counts of "bullish" and "bearish" words in the news media buy-sell indexes, stock by stock, using data on individual investor trades**

- **American Association of Individual Investors surveys**
- **Investors Intelligence surveys**
- **Surveys put out by Consensus Inc. and Market Vane**

So given what the literature has already looked at, the question is why study these particular measures.

- No alternative measures seem to have the popular following of the CEFD or enjoy even half the prominence of the CEFD. The CEFD still seems to be the right "bogey." Every theory in which investor sentiment appears as a concept seems to point to empirical papers that use the CEFD either as a primary sentiment measure, or as a component in constructing one. Perhaps the most prominent working papers right now circulating in this arena is Baker and Wurgler (2004)), and it does rely on the CEFD.
- It is correct that other sentiment measures have been proposed, but our impression is that, except for the CEFD, they are all "boutique." We could not find measures that seem to have been used by more than three authors/papers. Moreover, we could not even locate most of the papers (and therefore ask the authors for the variables). We cannot find reasonably common alternatives in published papers in good journals. We cannot even replicate the referee's search to locate all these variables—and it would be impossible to write a paper about sentiment if we have to deal with every sentiment measure that has appeared in some outlet somewhere. Still, we would love to look at other reasonably common sentiment proxies. **(We did ask the referee for a list of references, but did not receive an answer.)**
- In our revised version, we try to address one issue sharply, rather than write an encyclopedical paper about possible sentiment measures. We believe we offer one reasonable alternative sentiment measure, which does validate well against the UBS/Gallup poll. (We cannot possibly go after all possible measures of sentiment. We need some specific targets. We will be happy to try out other time-series measures that authors want to provide us with.) For a discussion of the AAII investor sentiment survey, please see faq question 9 below.
- Most other investor sentiment measures are just hypothesized constructs—proposed but never attempted to be validated. Consumer confidence at least validates against the direct UBS/Gallup sentiment survey. Do we know what these other proxies really measure?

7. **(QJE #2): Put differently, the literature has come a long way since Lee, Shleifer, Thaler (1991) looked at CEFD.**

   It is alive and well, but where has it gone to?

   - What, now, is a good measure of sentiment?
   - We know of no earlier paper that critiques it along the lines proposed in our own paper.
   - Which papers exactly have rejected the CEFD in a similar context to our's, i.e., the context in which the CEFD was originally validated (small firm returns)? Is this really a common finding? There were no earlier articles in publication that we are aware of. All we knew was the "Miller" dispute, which is very different from what we point out—actually, if Miller had looked at January separately, they would have had much more convincing evidence. We do. Also, Miller lost this argument in the opinion of many.
   - We have not seen LST write that the CEFD should not be used, nor does there seem to be a clear consensus about the fact that CEFD should be avoided in the behavioral finance literature. Instead, the CEFD lives on, though now often as a component of other ad-hoc, unvalidated proxies.
   - Attacking the CEFD is the focal point of Steve Ross' "crusade" against behavioral finance. (Ross' critique is based on the different transaction cost argument, though.) Ross is right. The CEFD is *not* a measure of investor sentiment. Ross will ultimately win this argument. In the long run, continued reliance on a faulty measure (the CEFD), or the hope of anonymous diffusion that the CEFD is faulty, can only end up discrediting a potentially useful concept (investor sentiment). Seriously, what is wrong with consumer confidence—it does work well, is stable, and validates directly.

8. **(QJE #2): The related question is why should we think (a priori) that consumer confidence surveys are better than CEFD (and all of the above measures). What is the motivation—given that many questions come to mind. The Michigan survey, which the paper favors, is a monthly telephone survey of 500 people, and only a fraction of them may have an intelligent opinion about the economy, and only a fraction of those will have a decent-sized portfolio.**

   These are the types of investors DSSW and behavioral finance is hypothesizing about. They are not hypothesizing about professional investors and economists. Wealthier investors and poorer investors respond alike. UBS/Gallup investors respond alike. CFOs respond alike. (Are they all conspiring to answer based on an irrelevant signal?!) Sentiment is a good name for this common factor—and it seems to be a good part of both investor sentiment and consumer confidence. See also Section II.B.

9. **(QJE #2): But the American Association of Individual Investors and Investors Intelligence surveys (the former goes back to the early 1960's, don't know about the former) are available and presumably better because they survey larger samples, and only investors.**

   See Page 14. Aside, it is not a universally accepted prominent standard to make it interesting enough to be a primary "bogey."

10. **(QJE #2): Also, the questions asked by the consumer confidence surveys (on page 8) don't have any direct bearing on "investor sentiment" as opposed to general feelings about the economy and personal financial conditions.**

    Yes. Measures need validation. How would one explain its correlation with investor sentiment? (And please read our discussion about the drawbacks of the consumer confidence as a measure.)

11. **(QJE #2): Also, the CC surveys are monthly, but big bubble-pop cycles seem to happen over several years, so I don't know whether there can even be 400 independent data points on sentiment, even if it was measurable, or something more like 10 or 20 over this sample period.**

    Good question. The evidence seems to suggest there is a more high-frequency effect here. (Of course, most papers in the sentiment literature would be affected, not just our's.) Annual sentiment measures in other academic papers rely on conjecture of their proxy, and not on any direct validation, in their use. They are often themselves financial variables (outcomes), rather than measures of individual's sentiment (model inputs). In our opinion, there is just not enough data right now one way or the other. (We did try some annual measures, and they produced similar results.) The only paper that we know that consider more long-term sentiment is Baker and Wurgler (2004)), but we find their sentiment proxy to not be one at all. (And, on a long-term basis, can one ever distinguish time-varying expected rates of return, perhaps risk-aversion based, from sentiment?) Is sentiment just a philosophical concept?

12. **(QJE #2): Another question is do we want a measure that isolates "sentiment" from macroeconomic conditions, or not. CEFD seems to have the separateness property, but these CC surveys do not.**

    No, we do not want isolated sentiment. The theories are about sentiment, not about sentiment orthogonal to macroeconomic conditions. In what theory would we expect sentiment not to be related to unemployment, GDP, portfolio returns, wealth changes, etc.? (Answer: None!) Sentiment does not drop like manna from heaven. We may need the decomposition to shed light on pathway, though. Also, please see our improved discussion of mechanisms and real vs financial pathways.

13. **(QJE #2): CEFD seems to have the separateness property, but these CC surveys do not.**

    Yes, the CEFD does have a separateness property to the economy, but *not* to the financial markets. (One has to have strong beliefs that the CEFD is driven primarily by sentiment, and not by, e.g., a book/market or other financial non-sentiment factor.)

14. **(QJE #2): I would say that the paper should address these questions. Given the many alternatives to CEFD, it is not surprising that one can find a measure that beats it on some metrics.**

    We did not try dozens of boutique measures to identify one that works. We only tried one, based on our priors. This has not been a data mining exercise. Also, you seem to suggest that the CEFD is still the measure to beat. Did our measure and tests not convincingly eliminate CEFD? What else is required to overturn the CEFD?

15. **(QJE #2): But, measuring "sentiment" is (apparently) quite difficult and controversial, and is always going to involve some subjectivity. We will always need to add some prior judgment before saying such and such is "the" better measure of sentiment. In other words, it is important to have measures with a strong a priori motivation as well as strong performance in various statistical respects. The consumer confidence surveys seem lacking in a priori motivation, at least in comparison to the many available alternative measures.**

    - Exactly, what many alternative common measures are readily available and generally accepted or at least commonly used?

- Did the inventors of other proxies just do what we are accused of—use an 'a priori' opinion?
- Have these inventors empirically validated their measures, not in terms of financial market correlations, but in terms of some form of "sentiment" interpretation?
- We are just stunned by the firm belief that the CEFD remains the right measure, despite our evidence. What evidence does it take to change it? How strong are the priors that the CEFD is the right measure?

Aside, we only had one idea for a proxy—consumer confidence.

16. **(QJE #2): The paper considers a couple of metrics: their correlation with "direct" UBS/Gallup survey of investor sentiment; correlation with the gap between small vs. large stocks and between individual vs. institutionally-owned stocks; correlation with closed-end fund IPOs. By each of these metrics, the consumer confidence measure seems to outperform the CEFD measure. The correlation with the UBS/Gallup survey is interesting, but if this type of "direct" survey is viewed as a superior measure in the first place, then why not use the American Association of Individual Investors survey, for which the data go way back, as noted above.**

    Please, not again. (The paper was not even published in an academic, but in a practitioner oriented journal.) See Page 14.

17. **(QJE #2): Let's take a look at the time series plots of the metrics. Something to like about the Michigan Consumer Confidence Index in Figure 1 is that it's high and rising during the Internet mania. (The negative of) CEFD, however, isn't doing much of anything over the last few years. But, look closer, and think about why Michigan CC jumps so dramatically in 1975, 1982/1983, and 1992. I'm not certain, but I don't believe those are stock market bubble years. Rather, the main factor is that the US was coming out of recessions. (Business conditions were also very strong in the late 1990s.) Business conditions are genuinely getting better, consumers are "more confident". Is it at all surprising that small stocks, with precariously undiversified product lines, would generally do better than large stocks as the economy improves? This doesn't say much of anything about investor sentiment, it merely corroborates CC as a business cycle indicator. Likewise, it's not a stretch to think that stocks owned by retail investors would also do better as CC improves. Retail investors tend to hold stocks in firms with which they are familiar, in particular those they buy products from. If those investors feel better about their own personal financial prospects, they will buy more-presumably from the places they shop at, which are overrepresented in their portfolios. Once again, what does investor sentiment, as opposed to a completely rational revaluation of stocks' prospects, have to do with why Michigan CC does better than CEFD on these metrics?**

    Ultimately, we will not be able to interpret sentiment as rational or irrational, because neither classical nor behavioral finance can tell us what part of the overlapping evidence they exactly allow for. Thus, either camp can claim for themselves any evidence that is overlapping (i.e., can be interpreted as either). This is the case for all sentiment measures. Our revised paper does better on distinguishing pathways. This is not perfect. Sentiment influence is a necessary, but not a sufficient condition for the DSSW behavioral hypothesis. See our extended discussion in Section II.B.

B RFS (Jan 2006)

1. **(RFS #1): "The paper offers little that is new. We know from Lemmon and Portniagunina (2004) that consumer confidence is related to the small-firm differential. We also know from them that there is an unstable relationship between consumer confidence and the sentiment measure of BW that included the CEFD among its components."**

    The L&P paper is independent and contemporaneous. It has the same ideas our paper has—consumer sentiment, comparison to B&W and the CEFD, a retail premium. It had better luck with referees, so it is now forthcoming in the RFS. There are some differences, though:

    - Most importantly, L&P work with *levels* of consumer confidence. This variable has an autocoefficient of 0.96 in our data—it is a non-stationary variable. (L&P run regressions on this level measure, and, even though they may be valid, the statistical properties are not straightforward.) In contrast, our paper works with differences in consumer sentiment.
    - For the most part, the theory relates prices to sentiment, and returns to sentiment changes. L&P relate sentiment levels to returns.
    - L&P are perhaps less critical of the CEFD.

- L&P do not consider validation, and do not use the UBS/Gallup measure. We believe that, though simple, this is important.
- We cannot compare our relative interpretations of the role of the retail spread (them: significant; us: not-robust), because we do not know how they computed their portfolios.

In their favor, L&P did try to control for more macroeconomic variables, such as Lettau/Ludvigson's CAY. (Of course, CAY has as two of its components future wealth and future income, which could relate to future returns. In this context, it should be harmless, though.)

2. **(RFS #1): The empirical work is poorly executed. For example, the returns of small stocks and those of the S&P 500 Index are highly correlated. Placing both as independent variables, as in the equation on p.22, is the likely cause of the negative coefficient on small stock returns and the positive coefficient on the S&P 500 returns.**

   It's just a common market model.

3. **(RFS #1): There are advantages and disadvantages to surveys of sentiment, such as consumer confidence, and inference of sentiment from market data, such as the CEFD. People do not always do what they say.**

   True. There is no particular reason here why respondents would lie, though. Also, the sampled individuals by UBS/Gallup are different from those sampled by Michigan (and the Conference Board). What factor would explain the correlation if one set of them lied? This would have to be strangely coordinated lies (or mistakes).

4. **(RFS #1): Moreover, it is likely that there is more than one sentiment at a given time. For example, Fisher and Statman (Financial Analyst Journal, 2000) found that the sentiment of Wall Street Strategists is uncorrelated with that of writers of investment newsletters but both are useful in the prediction of future returns.**

   If there is no strong common factor, what good is the concept of investor sentiment? The AAII has issues. Please see our discussion on Page 32.

C JFE (June 2006)

(Omitting paper summary. Rest is verbatim.)

1. **This paper has many problems. Perhaps the most severe is exposition. While reading it, I guessed the paper was written by a Ph.D. candidate who had been buffeted by a big and overactive dissertation committee. The four referee reports appended to the paper suggest I was not too far off. The result is a paper that lacks a simple, coherent structure or logic. At several points, I found myself wishing the authors would just get on with it.**

   The writing author has been trying, at least since 1987—apparently without much success.

2. **For example, why are we subjected to almost a full page (6) discussion of how the LST and Ross measures of CEFD are spliced? And despite the full page, we never learn why the two measures differ or, remarkably, how the two series are actually spliced! The next page (7) is a long discussion of consumer confidence that says little of consequence. A short paragraph of the new ideas would be sufficient. Pages 8 and 9 provide excruciating detail about the Michigan Consumer Confidence Index. Again, the important facts could be summarized in one or two brief paragraphs. These are four almost randomly chosen pages. I could make similar comments about many of the others.**

   This may be a matter of taste: We feel that explaining the details of splicing, and the surveys themselves is very important. The text is 27 pages, not overly long for a paper that has not converged with one referee/journal.

3. **More generally, the paper is repetitive, argumentative, and unfocused. The same ideas, arguments, and conclusions keep reappearing throughout the paper. It leaves me with the impression the authors think they are debating with a group particularly dull critics.**

   ...and, worse, apparently without much success, either.

4. **I do recognize the irony of criticizing the authors for reacting to referees, but in its current state the paper should not be published so more feedback cannot make the authors worse off. Their challenge is to constructively incorporate my comments and those of the other referees. The goal is a succinct, logical paper that makes sense to a reader who has not participated in prior debates about the subject with the authors.**

   (The other referees (and the next ones) are likely focused on the existing debate. They would probably complain if the paper was not a direct reaction thereto.) We also believe our paper is logical, and at 27 pages of text, double-spaced, not excessively long.

5. **I have at least two other substantive concerns. First, I think most behavioral finance types, including LST, interpret investor sentiment to mean irrational beliefs. The authors clearly recognize this, but they also recognize that investor beliefs about investment prospects can change for rational reasons. These rational reasons include changes in expected cashflows and changes in expected returns, which in turn might be caused by changes in risk or changes in risk aversion. (In the spirit of my criticism above, a clear, succinct, and better organized discussion of this issue would improve the paper.) The authors take a stab at disentangling the various "real" factors affecting consumer sentiment by projecting change in the Michigan Consumer Confidence index on contemporaneous and future corporate profits and aggregate consumption. They interpret the fitted values from these first pass regressions as the part of the change in consumer confidence due to the real economy. The residual is the part that is not due to the real economy. They then use the fitted values and the residuals from these first pass regressions to explain the small-big returns spread and conclude that the two parts of the change in consumer confidence contribute roughly the same amount to the prediction of the return spread.**
   This test does not seem robust. There are at least two potential (and possibly off-setting) problems. The first is functional form. If the relations between changes in consumer confidence and the macro variables are not linear, errors in the specifications will show up in the residuals and bias the tests toward too much importance on the component that is not related to the economy. The second problem arises because the first pass projects changes in consumer confidence on future variables that we think are related to the dependent variable in the second pass. Except for some pathological situations in the first pass, we expect to find a relation between the fitted values from the first pass and the return spread in the second. To see the point, suppose we run a univariate first pass regression in which the independent variable is highly correlated with the return spread, but essentially uncorrelated with changes in consumer confidence. As long as the first pass slope is not identically zero, we will find a strong relation in the second pass, even if the t-statistic in the first pass is trivially different from zero.

   [1] Functional form specifications can indeed cause an omitted variables problem. Similar requirements are unfortunately universal in any kinds of decompositions. We are unaware of any methods to avoid this. [2] Your question is whether fitted consumption confidence changes are simply (linear) future actual consumption or profit changes (which therefore influence the return spread). If they are, then this is fine with us. It would be exactly the interpretation proposed by the rational camp.

6. **Second, the authors overstate their results. For example, they argue that the CEFD "has lost it contemporaneous explanatory power for the small stock return spread after 1985." The coefficient does fall by about two thirds, from 1.043 to 0.355, and the t-statistic falls from 5.21 to 1.58, but a Bayesian would probably not conclude the coefficient is now zero.**

   What can we say? Yes, we are interpreting a T-stat of 1.58 to be insignificant, even though a Bayesian may not.

7. **As another of many possible examples, on page 13 the authors write: "Therefore, we can conclude that there is no difference in how poor and wealthy investor sentiment changes month-to-month. Wealth is not a determinant of sentiment changes." (The italics, in the original, magnify the overstatement.) This does not pass the smell test, let alone the Classical Statistics 101 question about failure to reject versus accepting the null. Frequent overstatements such as these undermine the reader's faith in the authors' objectivity.**

   We have about 60,000 investors. Under a null, the correlation between the two has a bootstrapped 83%. The empirical estimate sits at 80.4%, which is the 23rd percentile. I am not clear about your critique: Should we believe that our failure to reject the NULL of generally equal perspectives to be primarily an issue of low power within this large a set of data? Moreover, the differences seem surely not related to wealth in an economically important fashion, either.

8. **Some smaller comments: 1. Why does the sample period stop in 2002? I realize the authors have been trying to place the paper in a journal for a while, but given the short sample period for some of the critical data, the extra three years available since 2002 could improve the empirical tests a lot.**
   We tried to get more, most recently in 2006, but this has not been possible.

D JF (< 2005) and General

1. **(JF:) Please revise and resubmit, but make this more of a paper on consumer confidence as a factor.**
   Thank you for the kind words and decision, but this is not the goal of our paper. We are interested in sentiment. Consumer confidence is a good-yet-noisy proxy from our perspective, and not the prime focus in its own right.

The remaining FAQs are not based on referee comments:

1. **Is this an economics or a finance paper?**
   Both. For economists, it is about the correlation between consumer confidence and the overall stock market. It also is related to DeLong, Shleifer, Summers, and Waldmann (1990), which is on noise trader risk in financial markets, published in the Journal of Political Economy.
2. **Is this a behavioral finance or a classical finance paper?**
   Neither and both. It seeks to offend equally.
3. **We already have discarded the CEFD. We already knew this.**
   Really? It continues to be used and LST continues to be referenced.
4. **Consumer confidence is not a measure of confidence. It is driven by macroeconomics.**
   CC is correlated with the UBS/Gallup direct measure. And, yes, investor sentiment *should* be driven by macroeconomics and other variables. Sentiment does not fall like manna from heaven.
5. **Your validation metrics are poor.**
   Our paper is not perfect. We agree that the validation metrics in Lee, Shleifer, and Thaler (1991) are problematic. We believe that our own validation metrics significantly improve thereon: We have direct validation with the UBS/Gallup investor sentiment measure.
6. **Would a non-finding for CC in pricing small stocks have mattered?**
   It would not have mattered if CC were only a theoretical hypothesized measure. If the CEFD had not worked in LST, a non-finding would have been uninteresting—we would have discounted the interpretation of the CEFD as a proxy for sentiment. However, we do use the UBS/Gallup more direct proxy for investor sentiment for validation. This makes even a finding of no measure and/or role of investor sentiment interesting.
7. **You are missing the common-factor cross-sectional tests in Lee, Shleifer, and Thaler (1991).**
   True. However, it is not clear what this adds. If rational discount factors have time-series variation, then the PV of transaction costs would likely make funds' CEFD covary, too.
8. **You are missing the CEF initiation tests in Lee, Shleifer, and Thaler (1991).**
   Although no longer in the paper, we could not replicate the LST CEF opening evidence in the broader SDC data base on the CEFD metric.

B Institutional vs. Retail Portfolio Construction

The institutional holding portfolios were formed from Thomson 13(f) data, starting in 1980. Stocks are ranked using the holding data at the end of prior quarter; for example, January, February, and March groups are formed according to holdings in the prior December last year. We do not have holding data for Dec. 1979, so we can not form deciles in Jan. and Feb. of 1980.

Firms with zero institutional holdings are grouped into their own category; the remaining stocks are grouped into deciles. The theory predicts higher explanatory power for stocks held primarily by retail investors and not very liquidly traded. Thus, retailstocks.xret is constructed as follows:

Within each institutional holdings decile portfolio and within the zero-institutional holdings decile portfolio, stocks are sorted by dollar trading volume. We then create one equal-weighted portfolio of the low-trading volume zero-institutional holding stocks and one high-trading volume high-institutional holding stock portfolio. The difference in returns is retailstocks.xret.[18]

[18]The unconditional spread is very high. This disappears if we value-weight the portfolios. However, our cross-sectional results remain robust.

Table 1: Sentiment Variables

Closed End Fund Discount, Extracted from Financial Series

Variable	Mean	Sdv	Min	Q1	Median	Q3	Max	ρ	Range	#obs
bullish.cefd.ew, in %	−8.489	7.29	−23.85	−13.81	−8.72	−4.07	14.37	96	1965/07- 2002/12	450
bullish.cefd.vw, in %	−8.488	7.28	−26.69	−14.04	−9.52	−3.27	13.05	95	1965/07- 2002/12	450
d.bullish.cefd.ew, in %	0.005	2.18	−9.36	−1.09	−0.05	0.91	8.39	−25	1965/08- 2002/12	449
d.bullish.cefd.vw, in %	0.043	2.34	−10.92	−1.25	0.00	1.11	10.28	−19	1965/08- 2002/12	449

Primary Sentiment Measures, Survey Based

Variable	Mean	Sdv	Min	Q1	Median	Q3	Max	ρ	Range	#obs
bullish.mich.cc	87.042	12.24	51.7	78.3	89.9	94.75	112	96	1965/01± 2004/03	471
d.bullish.mich.cc	−0.013	3.27	−12.2	−1.7	−0.22	1.90	17.3	1	1965/02–2004/03	470

Validation Sentiment Measures, Survey Based

Variable	Mean	Sdv	Min	Q1	Median	Q3	Max	ρ	Range	#obs
bullish.gallup.all	0.349	0.29	−0.35	0.13	0.43	0.58	0.78	86	1999/02± 2002/12	57
bullish.gallup.wealthy	0.407	0.31	−0.33	0.17	0.51	0.66	0.89	80	1999/02± 2002/12	57
bullish.gallup.poor	0.304	0.29	−0.36	0.06	0.34	0.55	0.76	88	1999/02± 2002/12	55
									–	
d.bullish.gallup.all	−0.010	0.16	−0.31	−0.13	−0.01	0.10	0.41	−24	1999/03- 2002/12	47
d.bullish.gallup.wealthy	−0.008	0.20	−0.48	−0.14	−0.03	0.11	0.52	−36	1999/03- 2002/12	47
d.bullish.gallup.poor	−0.012	0.14	−0.31	−0.12	−0.02	0.09	0.33	−17	1999/03- 2002/12	46

Description: All data are monthly. Prefix d denotes monthly differences. cef are closed-end funds, cefd is the closed-end fund discount, mich is the Michigan consumer confidence index. Prior to 1978, it was linearly interpolated from quarterly observations. gallup is the UBS/Gallup poll of investors. (From 1996 to 1999, it was available only irregularly, and thus ignored in changes.)

Table 2: Sentiment Measures Validations, Monthly Data, Nov 1996 To Dec 2002

Correlations with Changes in the UBS Gallup Poll, d.gallup.opt.all

Validation		Corr	T-stat
Changes in Consumer Confidence (Michigan)	d.bullish.mich.cc	55.2%	4.45***
Changes in Consumer Confidence (Conf Board)	d.bullish.cb.cc	38.1%	2.77***
Changes in Consumer's Business Confidence (Michigan)	d.bullish.mich.bc	53.0%	2.86***
Changes in the value-weighted negative CEFD	d.bullish.cefd.vw	−2.0%	−0.13
Changes in the equal-weighted negative CEFD	d.bullish.cefd.ew	18.0%	1.23

Correlations with Changes in the UBS Gallup Poll, by Respondent Self-Reported Wealth

		d.gallup.opt.wealthy		d.gallup.opt.poor	
		Corr	T-stat	Corr	T-stat
Changes in Consumer Confidence	d.bullish.mich.cc	56.0%	3.58***	47.5%	4.53***
Changes in CEFD	d.bullish.cefd.ew	14.5%	1.43	21.1%	0.98

Description: Data are 47 monthly observations for all correlations. d denotes the first difference. The theories suggest that d.bullish variables should be positive when investors become more optimistic. cefd is the closed-end fund discount based measure, ew denotes that it is equal-weighted, vw that it is value-weighted. mich is the Michigan consumer confidence index. gallup is the UBS/Gallup poll of investors, wealthy refers to investors with more than $100,000 in wealth. Significance Levels: * = 5%, ** = 1%, *** = 0.1%.

Table 3: The Lee, Shleifer, and Thaler (1991) Validation Metric: Sentiment Correlation With the Return Spread of Small Minus Large Firm Deciles

Data		constant	d.bullish.mich	d.bullish.cefd.ew	sp500.pctchg
full sample	Coefficient:	0.876	0.402	0.734	−0.059
df= 445, $\overline{R}^2$ = 7.7%	T-statistic:	2.66**	3.89***	4.85**	−0.78
pre-1985	Coefficient:	1.297	0.457	1.043	0.250
df= 229, $\overline{R}^2$ = 14.6%	T-statistic:	2.70**	2.72**	5.21**	2.21*
post-1985	Coefficient:	0.486	0.367	0.355	−0.297
df= 212, $\overline{R}^2$ = 6.5%	T-statistic:	1.15	3.06**	1.58	−3.14**

Excluding January Observations

Data		constant	d.bullish.mich	d.bullish.cefd.ew	sp500.pctchg
full sample	Coefficient:	−0.097	0.369	0.172	−0.105
df= 408, $\overline{R}^2$ = 4.1%	T-statistic:	−0.35	4.23***	1.28	−1.66
pre-1985	Coefficient:	0.213	0.418	0.359	0.127
df= 210, $\overline{R}^2$ = 5.0%	T-statistic:	0.51	2.80**	1.92	1.26
post-1985	Coefficient:	−0.368	0.333	0.049	−0.288
df= 194, $\overline{R}^2$ = 9.2%	T-statistic:	−1.08	3.45***	0.26	−3.76***

Description: The dependent variable, smallstocks.retspread, is the monthly rate of return on the smallest decile of firms minus that of the largest decile of firms. d.bullish.mich is the change in the Michigan consumer sentiment index. d.bullish.cefd.ew is the decrease in the equal-weighted closed-end fund discount. sp500.pctchg is the percent change in the S&P500 index, and basically is in the regression as a market-model control. Because the dependent variable is a spread, it is not highly correlated with the S&P500 change, either. Significance Levels: * = 5%, ** = 1%, *** = 0.1%.
The first row of each regression prints the plain OLS coefficient, the second row prints the standardized coefficient (both dependent and independent variables are normalized to a mean of 0 and a standard deviation of 1). The third row prints the t-statistic. One star (two stars) denote significance at the 5% (1%) level, *two-sided*. The return series are available from 1965/01 to 2003/12. Their univariate characteristics are:

Variable	Mean	Sdv	Min	Q1	Median	Q3	Max	ρ	#obs
smallstocks.xret, in %	0.885	7.10	−20.62	−3.18	−0.03	3.25	42.92	11	468
sp500.pctchg, in %	0.642	4.39	−21.76	−1.90	0.79	3.63	16.30	1	471

Table 4: Quarterly Data, Correlations with Real Economic Performance (SA), Full Sample Regression

Panel A: Decomposing Changes in Michigan Consumer Confidence

Model	Variable (Changes in)		Constant	Lags	Contemp	Lead-1Q	Lead-2Q	Lead-3Q	Lead-4Q
*	Corporate Profits			Lags generally not significant					
1	Corporate Profits		−0.353		38.091				
		T-statistic	−0.85		5.01***				
2	Corporate Profits		−0.533		36.611	10.999	8.503	5.047	
		T-statistic	−1.26		4.80***	1.44	1.12	0.67	
	Consumption			Lags generally not significant					
		T-statistic							
3	Consumption		−1.187		131.535				
		T-statistic	−1.69		2.08*				
4	Consumption		−2.653		82.812	128.605	147.949	−57.761	
		T-statistic	−2.87**		1.26	1.96	2.25*	−0.86*	
5	Consumption		−2.895			121.589	143.349	−64.530	132.392
		T-statistic	−3.13**			1.87	2.20*	−0.97	1.98*
6	Both	Consumption:	−1.462		−29.508	111.446	129.153	−94.873	
		Profits:	−1.35		−0.43	1.64	1.89	−1.34	
					36.304	5.072	1.215	8.145	
					4.37***	0.61	0.15	0.98	
*	Models With Both:		Close lags and next few leads are not significant.						

Panel B: Explaining the Quarterly Size-Decile Return Spread

	constant	d.bullish.cefd.ew	sp500.ret	d.mich.bus.umich	
Base	2.952	1.367	0.065	0.943	
T-statistic	2.52*	3.47***	0.43	4.09***	
CC Decomposition	constant	d.bullish.cefd.ew	sp500.ret	d.bullish.mich.cc due to economy	d.bullish.mich.cc not economy related
Model 1	2.962	1.362	0.070	1.191	0.899
T-statistic	2.52*	3.45***	0.46	2.16*	3.64***
Model 2	2.979	1.369	0.064	1.211	0.890
T-statistic	2.53*	3.47***	0.42	2.34*	3.59***
Model 3	3.110	1.446	0.072	3.840	0.849
T-statistic	2.69**	3.72***	0.48	3.14**	3.69***
Model 4	3.056	1.383	0.038	1.422	0.908
T-statistic	2.58*	3.50***	0.24	1.91	3.84***
Model 5	2.951	1.367	0.065	0.942	0.943
T-statistic	2.49*	3.46***	0.41	1.37	3.96***
Model 6	2.957	1.367	0.064	0.970	0.936
T-statistic	2.51*	3.46***	0.42	2.03*	3.72***

Description: The dependent variable in Panel A, the first stage regression, is the quarterly change in the Michigan Consumer Confidence Index, d.bullish.mich.cc. There are 147 degrees of freedom in these regressions. Model 6 reaches an $\overline{R}^2$ of 16.4%. The dependent variable in Panel B is the quarterly return spread between size decile 1 and size decile 10 stocks. d.bullish.cefd.ew is significant only because this is the full-sample regression; post-1985 or without Januaries, it is not. "Due to economy" is the fitted value from the Panel A model, as indicated in the left-most column, "not economy-related" is the residual. There are

Table 5: Adding CEO Confidence Measures with Quarterly Data

Panel A: Business Confidence Correlation

Corr d.bullish.gallup.all with d.bullish.mich.cc	73.0%	3.85**
Corr d.bullish.gallup.all with d.bullish.cefd.ew	−1.9%	−0.07
Corr d.ceo.conf with d.bullish.mich.cc	42.1%	4.85***
Corr d.ceo.conf with d.gallup.all	27.1%	1.02
Corr d.ceo.conf with d.bullish.cefd.ew	−26.9%	−2.85**
Corr d.ceo.exp with d.bullish.mich.cc	41.9%	4.82**
Corr d.ceo.exp with d.gallup.all	29.7%	1.12
Corr d.ceo.exp with d.bullish.cefd.ew	−29.0%	−3.10**

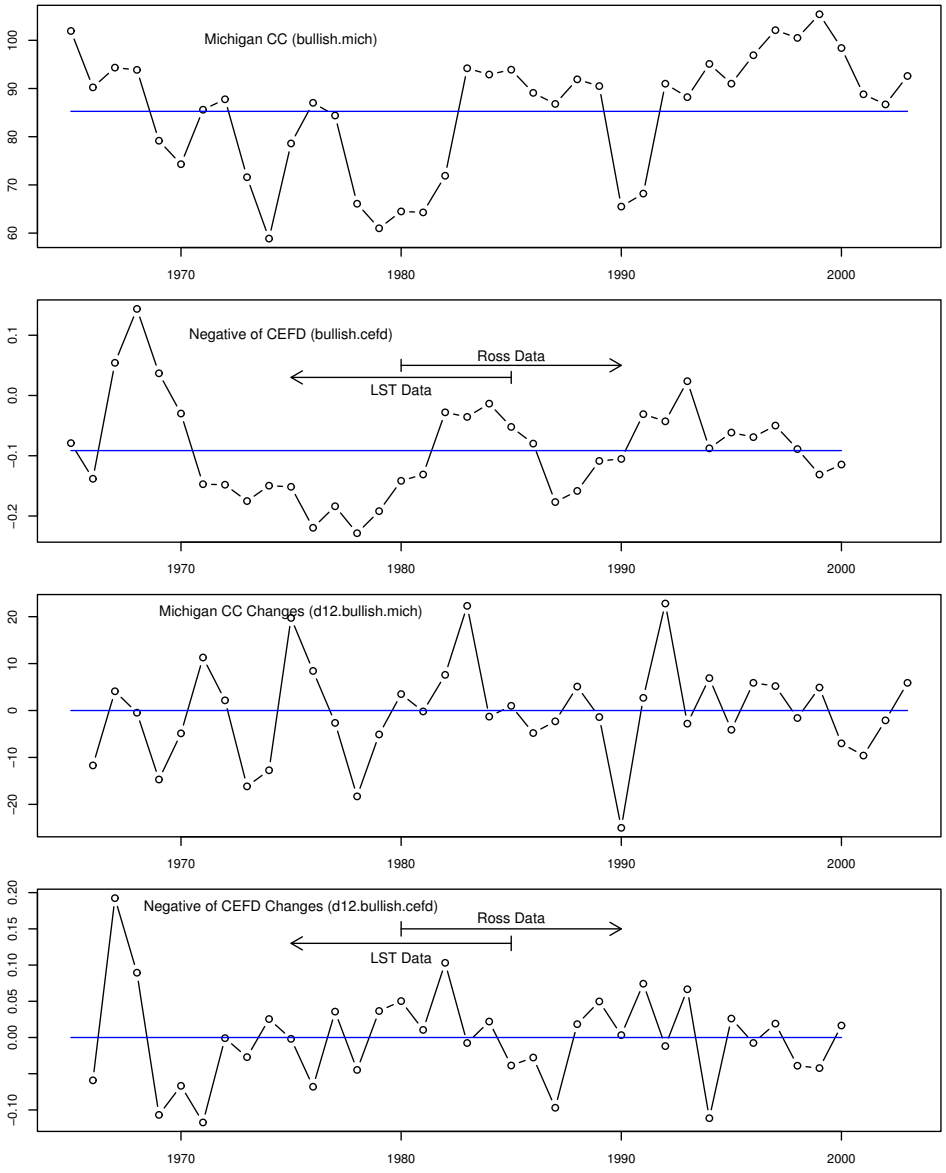
Panel B: The Sentiment Correlation With the Return Spread of Small Minus Large Firm Deciles, Adding Contemporaneous CEO Confidence.

			Full Sample				1985-	
Intercept	2.952	2.52*	2.618	2.45*	2.591	2.42*	1.881	1.44
d.bullish.mich.cc	0.943	4.09***	0.978	4.82***	0.979	4.80***	0.933	3.77***
d.bullish.cefd.ew	1.367	3.47***	0.206	0.48	0.207	0.48	0.445	0.88
sp500.ret	0.065	0.43	−0.309	−2.17*	−0.303	−2.11*	−0.397	−2.41*
d.ceo.exp			0.393	2.46*	0.597	1.41	1.296	2.31*
d.ceo.conf					−0.197	−0.48	−0.754	−1.34

Description: In Panel A, correlations are based on 15 quarterly observations when the Gallup variable is involved, and on 111 and on 106 quarterly observations otherwise (for the CC and CEFD, respectively). In Panel B, the dependent variable is the size extreme deciles return spread. The adjusted $\overline{R}^2$ is 28% in the two left panels, 31.3% in the right panel. Significance Levels: * = 5%, ** = 1%, *** = 0.1%. Univariate statistics are as follows:

Variable	Mean	Sdv	Min	Q1	Median	Q3	Max	ρ	Range		#obs
smallstocks.xret	2.375	13.1	−23.68	−6.18	−0.23	9.90	55.87	−13	1965/Q2	2003/Q4	112
d.bullish.mich.cc	0.152	5.98	−15.5	−3.80	0.80	3.20	22.20	−20	1965/Q4	2004/Q1	113
d.bullish.cefd.ew	−0.079	2.46	−6.20	−1.67	−0.18	1.39	8.18	−14	1965/Q4	2002/Q4	108
sp500.ret	2.567	7.86	−23.2	−1.26	2.49	7.54	20.87	0	1965/Q2	2004/Q1	113
d.ceo.conf	−0.162	7.00	−16.0	−5.00	−0.33	4.00	26	−7	1976/Q2	2005/Q4	118
d.exp.econ	−0.156	7.13	−14.0	−4.55	−1.00	4.00	29	−8	1976/Q2	2005/Q4	118

Figure 1: Time-Series of Sentiment Measures



Description: These figures are based on annual data. The levels are representative of the monthly levels, but changes are not. Our results are statistically less significant, but most of them tend to be similar, if we use annual data.