Investor Sentiment and Asset Valuation
Author(s): Gregory W. Brown and Michael T. Cliff
Source: *The Journal of Business*, Vol. 78, No. 2 (March 2005), pp. 405-440
Published by: The University of Chicago Press
Stable URL: https://www.jstor.org/stable/10.1086/427633

Gregory W. Brown
University of North Carolina at Chapel Hill

Michael T. Cliff
Virginia Polytechnic Institute and State University

Investor Sentiment and Asset Valuation*

The link between asset valuation and investor sentiment is the subject of considerable debate in the profession. If excessive optimism drives prices above intrinsic values, periods of high sentiment should be followed by low returns, as market prices revert to fundamental values. Using survey data on investor sentiment, we provide evidence that sentiment affects asset valuation. Market pricing errors implied by an independent valuation model are positively related to sentiment. Future returns over multiyear horizons are negatively related to sentiment. These results are robust to the inclusion of other variables that have been shown to forecast stock returns.

I. Introduction

A long-running debate in financial economics concerns the possible effect of investor sentiment on asset prices. For example, respected researchers have entered on both sides of the argument as to whether the stock price run-up and subsequent market collapse of 1929 was rational or not (see DeLong and Shleifer 1991 and White 1990). Perhaps "irrational exuberance" (Shiller 2000) drove prices above fundamental values. More recently, some commentators have suggested the rapid rise and fall of technology stocks was due to excessively bullish sentiments that started returning to more typical levels in spring 2000. Indeed, even before the massive devaluation of technology stocks, Malkiel (1999) wrote, "The spreading philosophy of the day traders that 'fundamentals don't matter' may well have contributed to valuations

* Funding from the Cato Center for Applied Business Research is gratefully acknowledged. We thank Jennifer Conrad, Mike Cooper, Kent Daniel, Greg Kadlec, Adam Reed, Steve Slezak, Meir Statman, Kent Womack, an anonymous referee, and seminar participants at the Federal Reserve Board, the University of North Carolina, Virginia Tech, the 1999 Western Finance Association meeting (especially the discussant Bhaskaran Swaminathan), the 1999 Financial Management Association meeting, and the 2000 Batten Young Scholars Conference at William and Mary for their comments and suggestions. We also thank Gurdip Bakshi, Zhiwu Chen, Ken French, and Steve Sharpe for providing data. Contact the corresponding author, Gregory W. Brown, at gregwbrown@unc.edu.

(*Journal of Business*, 2005, vol. 78, no. 2)

0021-9398/2005/7802-0001$10.00

of Internet stocks that can only be described in terms of a financial bubble.''

The existence of systematic mispricing in the market remains contentious because of the difficulty in examining the issue empirically.[1] The absence of precise valuation models for stocks makes it difficult to measure deviations from theoretical prices. Similar problems arise from the difficulty in measuring investor sentiment.

Several recent papers identify examples that are difficult to reconcile with rational pricing. Lamont and Thaler (2001) examine 3Com's spin-off of Palm. In this transaction, the market valuations of the companies implied the 3Com ''stub'' (a 3Com share less the claim on Palm) was −\$63, a violation of the law of one price. Rashes (2001) discusses the heavy trading in Massmutual Corporate Investors (ticker MCI) around Worldcom's acquisition of MCI Communications (ticker MCIC). Apparently, these investors were unable to determine the correct ticker symbol and mistakenly placed trades in Massmutual, and their trading drove the price away from its intrinsic value. Ofek and Richardson (2001) discuss a variety of examples from the Internet sector in developing a ''strong, circumstantial case against market efficiency.'' While these papers provide interesting evidence of misvaluation, they are specific to either a few companies or an industry that is admittedly difficult to value. Our focus in this paper is at the level of broad market indices.

We shed new light on the issue of investor rationality by bringing new data and techniques to the question. Specifically, we use a direct survey measure of investor sentiment instead of relying on other proxies such as closed-end fund discounts (see Lee, Shleifer, and Thaler 1991; Swaminathan 1996; and Neal and Wheatley 1998). Perhaps the most important distinction between our paper and most prior research is that our empirical tests concentrate on relating sentiment levels directly to stock price deviations from fundamental value and on the long-run effects of sentiment on stock returns.[2] Shanken and Tamayo (2001) provide evidence of expected return predictability that is unrelated to risk. They characterize this predictability as being due to mispricing. Their analysis uses a slow-moving variable (the dividend yield) as the conditioning information, implying that mispricing may be a long-horizon phenomenon.

Intuitively, looking at the relation between long-run market returns and sentiment is appealing for two reasons. First, it seems natural to

1. See Fama (1998) for a recent review from the rational camp or Hirshleifer (2001) for a review from the behavioral side.

2. Solt and Statman (1988), Clarke and Statman (1998), Otoo (1999), Fisher and Statman (2000), and Brown and Cliff (2004) also use survey data but focus on the short-run implications. Neal and Wheatley (1998) do long-horizon regressions but use proxies for sentiment.

view sentiment as a persistent variable. People become more optimistic as they are reinforced by others joining on the bandwagon. Thus, the importance of sentiment may build over time. Second, arbitrage forces are likely to eliminate short-run mispricing but may break down at longer horizons. Two examples of the limits to arbitrage are the noise trader risk of DeLong, Shleifer, Summers, and Waldmann (1990; hereafter DSSW) and the interaction of agency costs and capital constraints in Shleifer and Vishny (1997). Indeed, sentiment appears to have little predictive power for subsequent near-term returns (see the papers cited in note 2). This result is not surprising, since short-run predictability would lead to a simple trading strategy generating abnormal returns. However, the lack of predictability in the short run does not imply that sentiment has no affect on prices. Sentiment may drive asset prices away from intrinsic value for extended periods of time, yet this would be difficult to detect over short horizons (see Summers 1986). Whether sentiment does affect asset valuations in this way is an important open question we address in this paper.

We test two main hypotheses. The first is that excessive optimism leads to periods of market overvaluation. If so, this leads to the second hypothesis, that high current sentiment is followed by low cumulative long-run returns as the market price reverts to its intrinsic value. If the price correction were quick and predictable, there would be a potentially profitable trading strategy. However, a gradual correction over some unknown horizon is possible if there are limits to arbitrage. For example, an arbitrageur may believe with high confidence that the market is overvalued but be unwilling to take a short position for fear the market may become more overvalued before reverting to its intrinsic value. Frequent performance evaluation exacerbates this problem, since investors may withdraw capital precisely at the time when it is needed to meet margin calls.

To test our first hypothesis, we relate the level of sentiment to market mispricing as proxied by the Dow Jones Industrial Average pricing errors from Bakshi and Chen (2001). Estimating the relation between sentiment and mispricing is econometrically challenging because both time series are highly persistent. We undertake two types of tests, each of which yields similar conclusions in support of the hypothesis that sentiment is related to market mispricing. The first test uses the level of sentiment to explain pricing errors. We find a significantly positive coefficient on the sentiment variable, indicating the market is overvalued during periods of optimism. A second test treats the market and model valuations as a cointegrated system. Sentiment is a significant explanatory variable in the error correction version of the cointegrating regression. That is, after controlling for changes in fundamental value and the error-correction adjustment toward equilibrium, sentiment is positively related to changes in market valuations. Both sets of tests are robust to controlling for other factors predicting market returns (e.g., past returns, dividend yield, Fama-French factors).

In the test of our second hypothesis, we find that sentiment is in fact significantly related to long-run stock returns in the manner predicted. Specifically, high levels of sentiment result in significantly lower returns over the next 2 or 3 years. While this effect is present for the aggregate stock market, it is concentrated in large-capitalization growth stocks. The economic significance of this result is also plausible. For example, a one standard deviation (bullish) shock to sentiment results in a predicted 7% underperformance of the market over the next 3 years. These results are robust to inclusion of the control variables and the econometric issues related to overlapping observations.

Together these three tests provide strong and consistent support for the hypothesis that asset values are affected by investor sentiment. Each of the tests provides both statistically and economically significant results, all pointing in the same direction. Namely, overly optimistic (pessimistic) investors drive prices above (below) fundamental values and these pricing errors tend to revert over a multi-year horizon. This pattern is consistent with the predictions of many behavioral models that prices underreact in the short run and overreact in the long run.

The remainder of the paper is organized as follows. Section II provides some motivation for the paper and our approach. Section III discusses the data. The next two sections contain our main analysis. Section IV covers the long-horizon regressions and section V examines the relation between sentiment and the mispricing implied by the Bakshi and Chen (2001) model. In section VI, we present a number of robustness checks and comparisons to other empirical findings using alternative sentiment proxies. Section VII concludes the paper, and an appendix contains the details of the simulation used with the long-horizon regressions.

II. Motivation

In this section, we develop our hypotheses by discussing an environment where sentiment can affect asset valuation. Our approach makes three main assumptions. First, we assume that a subset of investors makes biased asset valuations. Second, we assume that this bias is persistent. Finally, we assume that limits to arbitrage hinder the exploitation of asset mispricing. These assumptions lead to an environment where market prices can differ from intrinsic value for protracted periods of time. Excessive optimism leads to overvaluation of assets. Arbitrage forces can eliminate profitable short-term trading strategies but not longer run mispricing. Over longer horizons, the high sentiment that leads to overvaluation would be associated with low long-run returns as asset prices revert to intrinsic values.

We begin by partitioning the universe of investors into two groups. The first group, which we can think of as rational investors, we refer to as *fundamentalists*. This group has the property that they make an unbiased assessment of an asset's intrinsic value. The second group of investors are swayed by episodes of excessive optimism or pessimism. For convenience, we refer to this group as *speculators*. As a group, they tend to overvalue assets during times of extreme optimism or high sentiment. When their sentiment is especially low, the group tends to undervalue assets.

Intuitively, we can think of the market price of an asset as reflecting a weighted average of the valuations of these two groups. More formally, this can come from a model such as Litner (1969). In that model, the heterogeneous judgments of investors aggregate to form market prices as we describe. Our setup amounts to a way of partitioning the investor universe in the Lintner model. In particular, for a set of assets in unit supply, the vector of asset prices is

$$\mathbf{P} = \frac{1}{R_f}[w_S(\boldsymbol{\mu}_S - \alpha\boldsymbol{\Omega}\mathbf{i}) + w_F(\boldsymbol{\mu}_F - \alpha\boldsymbol{\Omega}\mathbf{i})] = w_S\mathbf{P}_S + w_F\mathbf{P}_F \quad (1)$$

where R_f is the gross riskless return, α is the coefficient of absolute risk aversion, $\boldsymbol{\Omega}$ is the belief of the variance-covariance matrix of asset returns, and $\boldsymbol{\mu}_i$ is the belief of the vector of asset payoffs ($i = S$ for the speculators and F for the fundamentalists). We focus on the simplest case, where both types of investors have identical risk aversion and estimates of $\boldsymbol{\Omega}$ but different expectations of asset payoffs.[3]

Under these assumptions, the weights w_S and w_F are determined by the fraction of investors that are speculators versus fundamentalists. Clearly, if there are any speculators in the market and $\boldsymbol{\mu}_S > \boldsymbol{\mu}_F$, then $\mathbf{P} > \mathbf{P_F}$. Interpreting the fundamentalists as properly valuing the assets, this says the market price is above fundamentals when the speculators are overoptimistic. Our point in introducing this model is not to take it literally but just to show that the simple idea we posit is supported by formal models.

One question that arises is how the speculators can survive if they systematically misvalue assets. Here, we appeal to the limits to arbitrage. In the context of DSSW, our fundamentalists may recognize the market is overvalued but still be unwilling to try to exploit the mispricing. For example, portfolio managers may be evaluated annually, so they would be unwilling to take a position that may take longer to pay off. There is no pure arbitrage opportunity, since it is unknown when the market prices will converge back to the intrinsic value. Shleifer and Vishny (1997) show how agency problems between an arbitrageur and

3. Generalizing to allow differences in risk aversion or covariance estimates complicates the algebra but the important features remain.

his or her source of capital can also hinder arbitrage. In their model, if a position moves against the arbitrageur, the investors withdraw some capital. But this is precisely when the arbitrageur needs capital to meet margin calls. The arbitrageur is potentially forced to liquidate the position at a loss, even though the expected return is even more attractive than when the position was initiated.

There are two main implications of the environment we describe. First, we should see market overvaluation when sentiment is high. The second implication is that, as the market price reverts to its intrinsic value, long-horizon returns following periods of high sentiment should be abnormally low. These predictions are a common thread in many models. DeLong, Shleifer, Summers, and Waldmann (1990) and Shleifer and Vishny (1997) are two examples, although they have exogenous noise or sentiment traders. Several recent behavioral models model the source of noise more formally. In Daniel, Hirshleifer, and Subrahmanyam (1998), investors are overconfident about their private signals and they incorrectly attribute successful outcomes to their own abilities while blaming bad outcomes on chance (“biased self-attribution”). Barberis, Shleifer, and Vishny (1998) have a model where earnings are a random walk, but investors mistakenly believe earnings switch between a mean-reverting regime and a growth regime. The investors are slow to update their beliefs about the regime in the face of new evidence (“conservatism”). At the same time, these investors think they see patterns where none exist (“representativeness”). Hong and Stein (1999) posit two groups of boundedly rational agents, “newswatchers” and “momentum traders.” Newswatchers get private signals but ignore information in market prices. Slow transmittal of information causes prices to underreact in the short run. This underreaction leads to trading by the momentum investors. Overreaction results when the momentum investors have gone too far.

The primary aim of our paper is not to distinguish one of these theories from another. Instead, our goal is to evaluate the broad predictions of behavioral theories relative to the null of rational pricing. Where appropriate, we make connections to specific behavioral models. In the analyses to follow we present three main tests of these implications. In each case, we find strong support of these hypotheses. The tests are robust to controlling for rational factors that may be correlated with sentiment and to modifications to the methodology.

III. Data

The results in the paper are based on a sample of monthly data. For the long-horizon regressions, the data start in January 1963 and end in December 2000 for a total of 456 observations. The analysis involving pricing errors is limited to the shorter sample of 235 observations from January 1979 to July 1998.

A. Sentiment

A key variable used in the analysis is survey data on investor sentiment, denoted S_t. These data are from Investor's Intelligence[4] (II), which tracks the number of market newsletters that are bullish, bearish, or neutral. Currently, II categorizes approximately 130 market newsletters, although this number has apparently changed somewhat through time as market newsletters have come into or gone out of favor. Each newsletter is read and marked as bullish, bearish, or neutral based on the expectation of future market movements. Since the newsletters are not written with the survey in mind, they may differ somewhat in forecast horizon and thus may require interpretation in the categorization. Investors Intelligence indicates that a relatively small number of people are involved in categorizing the newsletters, so there should not be a large problem associated with differing interpretations. The sentiment data begin in 1963 with biweekly surveys. In 1969, the survey became weekly. We sample the series at month end, because a majority of the newsletters surveyed are published only monthly. On average, 43.8% of newsletters surveyed are bullish and 32.9% are bearish. The typical interpretation of the sentiment survey is as a contrarian indicator: When sentiment is unusually bullish, the market is predicted to experience below-average subsequent returns. For example, in a recent Forbes interview[5] with Michael Burke, editor of II, he stated:

"Most [newsletters] are trend followers. . . . If you go back to the end of 1994, when the Dow was at 3,700, people were very pessimistic . We had two weeks in a row with 59% bears on our chart, the most bears in 12 and a half years. Meanwhile, the market was just getting ready to take off on a five-year rise. In August 1987, however, two months before the crash, the sentiment was overly optimistic, with bulls over 61% and the bears at 19%. So we look at this indicator in a contrarian way."

Our preferred sentiment variable is the bull-bear spread, a common measure of sentiment in the popular press.[6] It is defined as the percentage of newsletters bullish minus the percentage bearish. Summary statistics for the bull-bear spread (S_t) over the two relevant sample periods are in table 1 and the series is plotted in figure 1. Our choice of using the bull-bear spread as a measure of sentiment does not drive our results. As we discuss in section VI, our results are robust to alternative variables using the Bull/Bear/Neutral data.

Inspection of the sentiment data reveals some interesting features. First, sentiment is quite variable: Readings greater or less than 20 in magnitude make up 43% of observations, although the most extreme

4. Chartcraft (www.chartcraft.com) provides the data on a subscription basis.
5. See Forbes.com, March 19, 2002.
6. For example, the bull-bear spread is published weekly in *Barron's* and is often mentioned in financial press articles.

TABLE 1 Summary Statistics

	Mean	Std Dev	Skewness	Ex Kurt	ρ_1	$\rho_{s,i}$
A. 456 Observations from 1/1963 to 12/2000						
S	10.8890	22.0475	.0428	−.3606	.7110	1.0000
RFx	.0037	.0989	−.1284	3.2788	.7755	−.2406
HB3	.0717	.1013	2.3729	10.1190	.3265	−.1202
TS	1.2339	1.2952	−.1419	−.3162	.9316	.2485
DS	1.0073	.4529	1.1951	1.2728	.9711	−.0302
DY	3.2349	.9970	−.0703	−.2224	.9865	−.2579
Infl	.3845	.3041	.9932	1.6812	.6468	−.2815
ExMkt	.5280	4.3997	−.5107	2.3572	.0416	.2828
SMB	.1596	3.2425	.5493	5.8565	.0956	.1822
HML	.4210	2.8878	−.0149	2.0385	.1657	−.0126
UMD	.9276	3.6184	−.0375	3.3080	.0392	.0039
B. 235 Observations from 1/1979 to 7/1998						
S	7.1694	18.4963	−.0171	−.4220	.7059	1.0000
RFx	−.0055	.1189	.0562	2.2936	.7570	−.2634
HB3	.0856	.1186	2.2201	8.0713	.3666	−.1006
TS	1.6073	1.4289	−.6903	.1779	.9272	.2104
DS	1.1466	.4912	1.0046	.5091	.9656	−.1016
DY	3.3521	1.0210	.0320	−.6550	.9873	−.2909
Infl	.3743	.3130	1.0573	1.3300	.7631	−.3221
ExMkt	.8067	4.2627	−.8068	3.9041	.0270	.3068
SMB	−.0008	2.5002	.0369	.6324	.1441	.1402
HML	.3631	2.5691	.1598	.3153	.1930	−.0599
UMD	.9203	3.1796	−.1080	2.0910	.0989	.0173
$p - p^*$	−.0040	.1117	−.1062	−.3174	.8834	.2044
R_{DOW}	1.7008	4.2046	−.4424	2.7785	−.0279	.2429

NOTE.—Panel A contains summary statistics for sentiment and the control variables used in the long-horizon regressions. Panel B contains the summary statistics for sentiment, the log pricing error on the "quasi-Dow" ($p - p^*$), and the monthly return on the quasi-Dow (R_{DOW}) for a shorter subsample. Panel B data are used in the pricing error regressions and cointegration regressions. The quasi-Dow is a value-weighted index we create using all available firms for which Bakshi and Chen (2001) have pricing errors. The variables are sentiment (S), the detrended interest rate (RFx), the difference in returns on 3- and 1-month T-bills (HB3), term spread (TS), default spread (DS), inflation (Infl), the Fama and French (1993) factors (ExMkt, SMB, and HML), and a momentum factor (UMD). See the text for additional variable descriptions.

readings occur in the first half of the sample (the standard deviation decreases from 25.7 in the first half of the sample to 17.0 in the second half). The second important feature is sentiment's strong persistence, with first-order autocorrelation of about 0.7. This is an intuitively pleasing characteristic, insofar as it suggests investors are not too fickle and bouts of optimism or pessimism may reinforce themselves. Finally, sentiment is strongly correlated with contemporaneous market returns but is not useful in predicting subsequent near-term returns.[7]

7. Brown and Cliff (2004) find that both recent market returns and prior sentiment readings are important predictors of sentiment levels and changes. While they document a strong positive contemporaneous correlation between sentiment and market returns, they find no evidence of profitable short-run trading strategies.

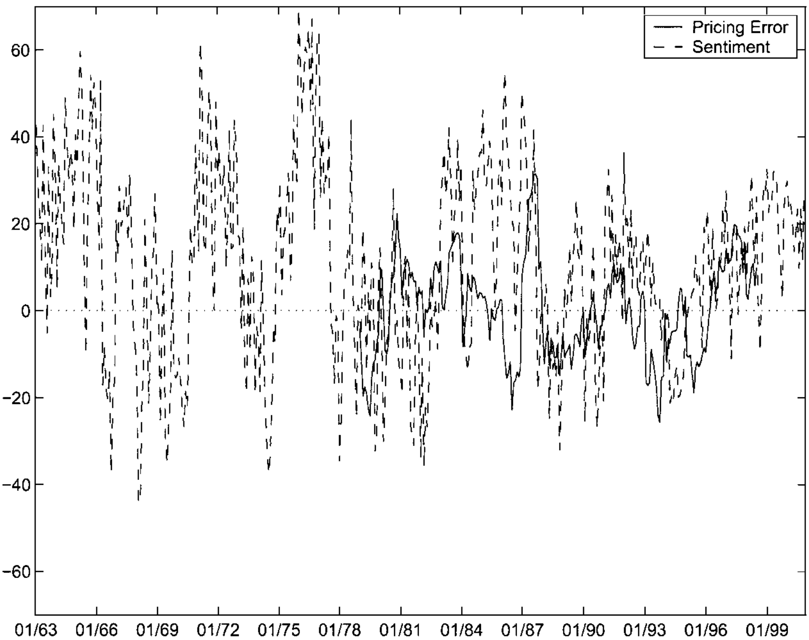

FIG. 1.—Pricing errors and sentiment. This figure shows the pricing error on the "quasi-Dow" index and investor sentiment, measured as percentages. The quasi-Dow is a value-weighted index we create using all available firms for which Bakshi and Chen (2001) have pricing errors. Sentiment data are monthly from 1963 through 2000. Pricing errors start in January 1979 and end in July 1998.

B. Dependent Variables

For the long-horizon regressions, we cumulate monthly log returns over various horizons. We use the 25 Fama and French (1993) portfolios formed on the basis of size and book/market sorts, in addition to the 5 portfolios from univariate size sorts, the 5 portfolios from univariate book/market sorts, and the overall market portfolio.[8] This collection of 36 portfolios allows an opportunity to see if predictability due to sentiment is affected by the size or value effects. Table 2 contains summary statistics for the 36 portfolios. In our sample, the value (high book/market) stocks had higher average returns and lower standard deviations than growth (low book/market) stocks. Small stocks generally had high average returns, high standard deviations, and large positive autocorrelations.

In the pricing error analysis, we have data from Bakshi and Chen (2001), who derive and estimate a discounted cash flow model with mean-reverting processes for earnings growth and interest rates. We use their pricing errors for each of the 30 stocks in the Dow Jones Industrial

8. We thank Ken French for providing these data and the SMB, HML, and UMD factors.

TABLE 2 Returns Summary Statistics

	Low	BM 2	BM 3	BM 4	High	All
A. Means						
Small	.7807	1.2724	1.3053	1.5025	1.6061	1.1995
Size 2	.8984	1.1530	1.3927	1.4536	1.5312	1.2288
Size 3	.9240	1.2557	1.2338	1.3820	1.5423	1.1982
Size 4	1.0742	1.0184	1.2598	1.4314	1.4682	1.1646
Large	1.0390	1.0043	1.0402	1.1479	1.2169	1.0252
All	1.0096	1.0181	1.1026	1.2738	1.4126	1.0352
B. Standard Deviations						
Small	8.0397	6.9855	6.0499	5.6533	5.8785	6.2894
Size 2	7.3342	5.9717	5.3282	5.0690	5.6048	5.8295
Size 3	6.7469	5.4185	4.9084	4.6638	5.2517	5.3003
Size 4	5.9466	5.1468	4.8414	4.5864	5.3428	4.9721
Large	4.7450	4.5314	4.3135	4.2734	4.5935	4.2223
All	4.9110	4.5888	4.2527	4.2147	4.7004	4.3795
C. Autocorrelations						
Small	.2128	.1813	.1861	.2000	.2359	.2179
Size 2	.1399	.1559	.1658	.1514	.1360	.1513
Size 3	.1055	.1424	.1384	.1455	.1403	.1283
Size 4	.0757	.1031	.0668	.0665	.0375	.0752
Large	.0256	.0077	−.0566	−.0556	.0136	−.0175
All	.0487	.0487	−.0033	.0120	.0746	.0352

Note.—This table shows the summary statistics for monthly returns on portfolios formed on size and book/market values. Portfolio formation follows the Fama and French (1993) procedure. Portfolios in the row labeled All are univariate book/market sorts. Portfolios in the All column are univariate size sorts. The portfolio in the All row and All column contains all available firms. These portfolios are used in the long-horizon regression by converting to multiperiod log returns. There are 456 observations from January 1963 through December 2000.

Average (DJIA) as of July 1998.[9] As the composition of the Dow has changed over time, the stocks for which we have pricing errors do not exactly represent the actual DJIA. In addition, not all the pricing errors are available back to January 1979. Consequently, we form an index, which we call the *quasi-Dow,* that uses the available stocks from this group. In particular, for each month, we value-weight the pricing errors on the stocks with pricing errors in that month. We also calculate the returns on this value-weighted index. The returns and pricing error series are then used to determine the log level of the market's valuation (p) and the log valuation level according to the Bakshi and Chen (2001) model (p^*). The working assumption is that the model valuation is a proxy for the intrinsic value of the index, and the market's deviation from this intrinsic value is a pricing error. It is possible, of course, that the "pricing error" $p - p^*$ is actually model misspecification, not

9. We thank Gurdip Bakshi and Zhiwu Chen for providing these data.

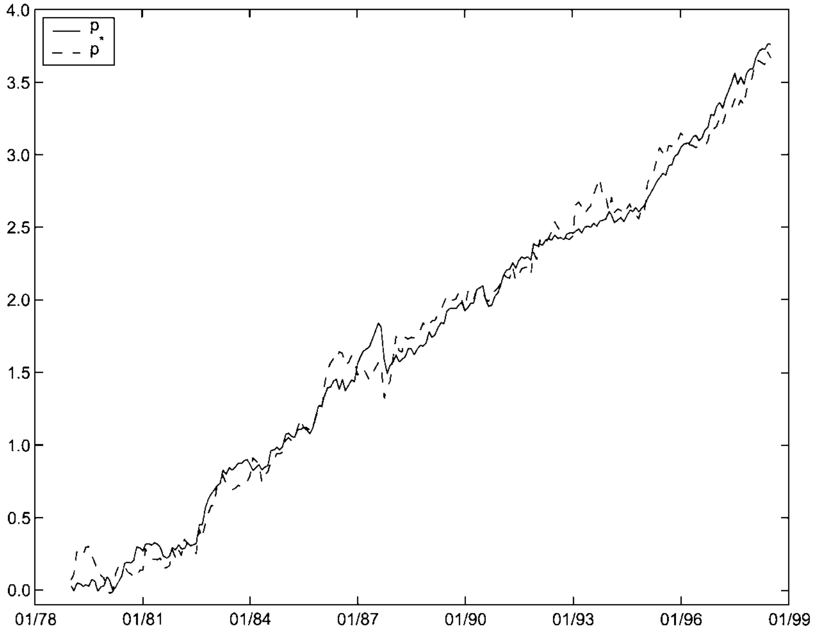

FIG. 2.—Market and model valuations. This figure shows the logs of the market valuation for the quasi-Dow (p) and the corresponding intrinsic value from the Bakshi and Chen (2001) model (p^*). The quasi-Dow is a value-weighted index we create using all available firms for which Bakshi and Chen (2001) have pricing errors. Data are monthly from January 1979 through July 1998.

mispricing. However, we present evidence that this does not seem to be the predominant explanation for our results.[10]

Figure 2 shows the (log) valuations for the market and the model. Both clearly move together over the long run, but there are substantial deviations for a year or more. Figure 1 shows the percentage pricing error along with the level of sentiment. The pricing errors make persistent swings around zero. The average pricing error is near zero, and the summary statistics in panel B of table 1 show that it is volatile and persistent. Figure 3 shows some properties of the pricing error and its relation to sentiment in the frequency domain. Panel A is the spectral

10. We choose the Bakshi and Chen (2001) pricing errors because we feel they provide the best measure, considering the available sample size and reasonableness of the pricing errors. Lee, Myers, and Swaminathan (1999) also model the intrinsic value of the Dow. However, their model assumes constant expected excess returns, which opens the possibility that measured pricing errors are due to true time-variation in risk premia. Since expected returns enter into the valuation equation nonlinearly, simply adding control variables in a linear regression may not be adequate. Also, their model does not pass our diagnostics tests. Specifically, we find that their intrinsic value measure reacts to sentiment. For our purposes, a good measure of intrinsic value should not respond to sentiment. Sharpe (2002) values the S&P 500 beginning in 1983. Results using his pricing error series are very similar to our main results. See section VI. D for further details.

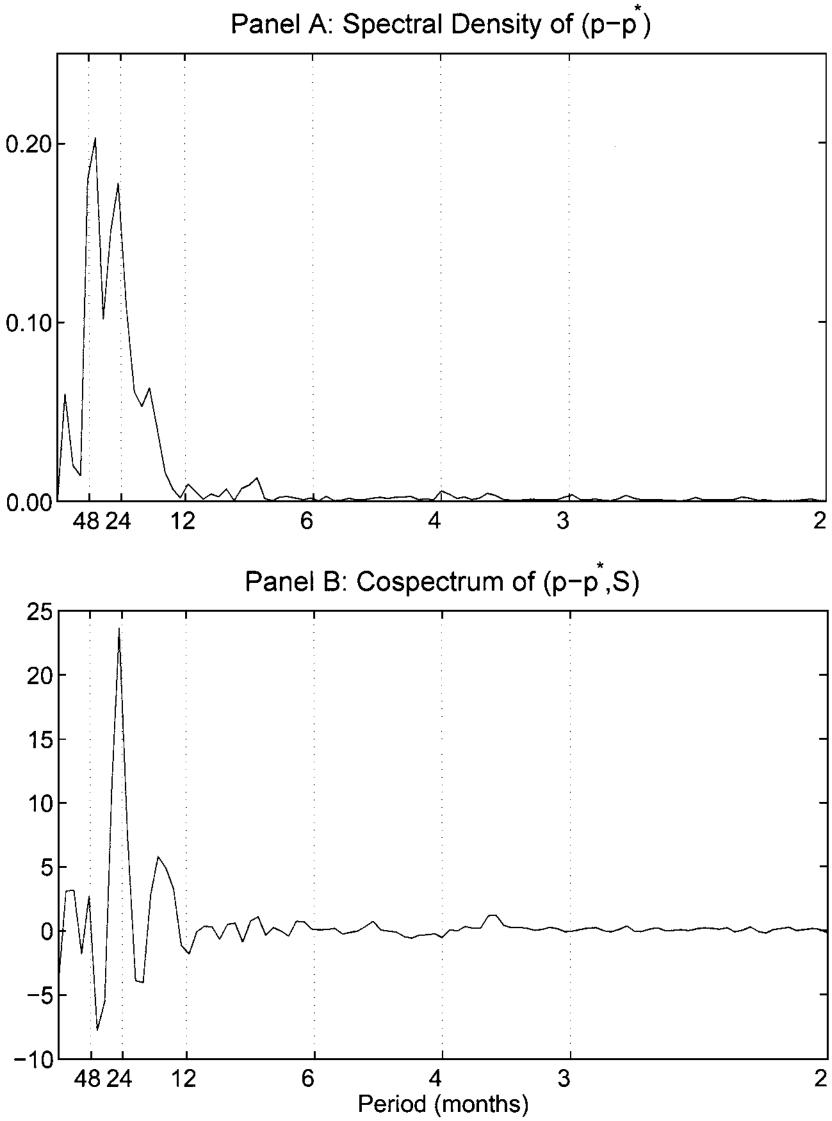

FIG. 3.—Frequency domain representation. Panel A plots the spectral density of the pricing error $p - p^*$ for the quasi-Dow index. It shows the importance of various frequencies in the variation of $p - p^*$. Panel B shows the cospectrum of $p - p^*$ and sentiment, which indicates the importance of various frequencies in the co-variation between the two series.

density of $p - p^*$ over various frequencies. This shows that most of the variability in the pricing errors is due to movements over the 2- to 4-year horizon, consistent with the swings in the pricing errors shown in figure 1 lasting for several years. Panel B of figure 3 shows the cospectrum of $p - p^*$ and S_t. This can be thought of as showing the comovement between the two series at various horizons. The plot indicates that most of the comovement comes from the 2-year horizon. The negative values around the 4-year horizon are consistent with a reversal correcting prior valuation errors. Although the figures do not provide formal tests that optimism (pessimism) drives market overvalutions (undervaluations), they are consistent with this interpretation.

C. Control Variables

To interpret our results as sentiment influencing future market valuations, we need to control for the information our sentiment variable may contain about rational factors. Indeed, our sentiment variable partially contains rational expectations based on risk factors and other variables that have been shown to predict future performance. When people say they are bullish on the market, this can be a rational reflection of prosperous times to come, an irrational hope for the future, or some combination of the two. We want to focus on the irrational part of sentiment, so we include a set of control variables designed to capture this rational predictability. We acknowledge that we might be missing some important rational factor, but we feel our set of control variables is a reasonable effort in mitigating this problem.

Our set of control variables are motivated by the conditional asset pricing literature. We use the stochastically detrended 1-month U.S. Treasury bill return (RFx; Campbell 1991; Hodrick 1992), the difference in monthly returns on 3-month and 1-month T-bills (HB3; Campbell 1987; Ferson and Harvey 1991), the term spread as measured by the spread in yields on the 10-year U.S. Treasury bond vs. the 3-month T-bill (TS; Fama and French 1989), the default spread measured as the difference in yields on Baa and Aaa corporate bonds (DS; Keim and Stambaugh 1986 or Fama 1990), the dividend yield for the value-weighted CRSP index over the past twelve months (DY; Fama and French 1988; Campbell and Shiler 1988a, 1988b), and the rate of inflation (Infl; Fama and Schwert 1977; Sharpe 2002).

In addition to these variables, we include several of the popular factors in asset pricing models. Specifically, we also include the excess return on the value-weighted market portfolio, the premium on a portfolio of small stocks relative to large stocks (SMB), the premium on a portfolio of high book/market stocks relative to low book/market stocks (HML), and the premium on a portfolio of past winners relative to losers (UMD). The market factor is based on the Capital Asset Pricing Model, while Fama and French (1993) show the SMB and HML factors are incrementally useful. The momentum factor is based on Jegadeesh and Titman (1993) and Carhart (1997).

Table 1 presents summary statistics for the control variables. The last column of the table shows the correlation with sentiment. Several of the variables have correlations of 0.2 or higher (in magnitude) with sentiment, but none is larger than 0.33. Therefore, it appears that sentiment does share common information with the control variables but may also contain incremental information.

IV. Sentiment and Long-Horizon Returns

Our first step in examining the impact of investor sentiment on asset valuations is to regress future k-period log returns on sentiment and the control variables ($\mathbf{z}_t$),

$$(r_{t+1} + \cdots + r_{t+k})/k = \alpha(k) + \mathbf{\Theta}'(k)\mathbf{z}_t + \beta(k)S_t + \varepsilon_t^{(k)}. \qquad (2)$$

As the length of the horizon increases, the number of observations in the regression drops, as does the number of independent observations.

The term $\beta(k)$ indicates the sensitivity of expected monthly returns over the horizon to investor sentiment. Under the null hypothesis that asset valuations are not influenced by behavioral forces, sentiment should not enter the regression significantly. Under the alternative hypothesis that optimism drives asset values above fundamental values and prices subsequently correct, β should be negative. Current optimism results in overvaluation, so future returns over the horizon would be lower than normal as the market valuation returns to its intrinsic value.

It turns out that careful implementation of this simple test is complicated. Long-horizon regressions are plagued by several econometric problems. The process of cumulating monthly returns, then running a regression with overlapping observations generates strong correlation in the residuals. Hansen and Hodrick (1980) provide standard errors that correct for this problem, but this correction does not perform well in finite samples (see Richardson and Stock 1989; Richardson and Smith 1991; Hodrick 1992; or Boudoukh and Richardson 1994). A second problem is that the inclusion of persistent independent variables can bias the coefficient estimates since they are predetermined but not strictly exogenous (see Stambaugh 1999). To circumvent these problems we use a bootstrap simulation. Details of the simulation are provided in the appendix. In short, the simulation allows us to correct for the bias and use the appropriate critical values for inferences.

The results from the long-horizon regressions are collected in tables 3 through 6. Each table contains four panels, corresponding to horizons of 6, 12, 24, or 36 months. Within each panel are the results for each of

TABLE 3 Sentiment Coefficient in Long-Horizon Regressions

	Low	BM 2	BM 3	BM 4	High	All
			A. 6-Month Horizon			
Small	−.0042	−.0115	−.0114	−.0069	−.0088	−.0073
Sz 2	−.0158	−.0126	−.0141	−.0110	−.0079	−.0170
Sz 3	−.0197	−.0149	−.0085	−.0094	−.0080	−.0159
Sz 4	−.0144	−.0124	−.0096	−.0058	−.0057	−.0123
Big	−.0125	−.0127	−.0066	−.0057	−.0071	−.0083
All	−.0130	−.0135	−.0076	−.0052	−.0060	−.0067
			B. 12-Month Horizon			
Small	−.0053	−.0117	−.0113	−.0083	−.0093	−.0081
Sz 2	−.0137	−.0112	−.0125	−.0119	−.0087	−.0129
Sz 3	−.0157	−.0115	−.0090	−.0107	−.0102	−.0127
Sz 4	−.0142	−.0118	−.0105	−.0087	−.0082	−.0122
Big	−.0155	−.0143	−.0089	−.0109	−.0096	−.0130
All	−.0153	−.0143	−.0093	−.0094	−.0083	−.0110
			C. 24-Month Horizon			
Small	.0033	−.0035	−.0029	−.0022	−.0016	−.0004
Sz 2	−.0066	−.0031	−.0037	−.0046	−.0017	−.0039
Sz 3	−.0076	−.0053	−.0027	−.0032	−.0043	−.0048
Sz 4	−.0102	−.0081	−.0053	−.0034	−.0027	−.0069
Big	−.0134	−.0133	−.0070	−.0066	−.0094	−.0114
All	−.0122	−.0119	−.0063	−.0045	−.0050	−.0088
			D. 36-Month Horizon			
Small	.0073	.0009	−.0010	−.0006	.0005	.0020
Sz 2	−.0029	−.0013	−.0020	−.0038	−.0005	−.0018
Sz 3	−.0056	−.0044	−.0029	−.0023	−.0012	−.0035
Sz 4	−.0095	−.0074	−.0060	−.0032	−.0015	−.0065
Big	−.0145	−.0136	−.0079	−.0069	−.0082	−.0122
All	−.0129	−.0118	−.0070	−.0046	−.0037	−.0092

NOTE.—This table shows the coefficients on sentiment from regression of k-period log returns on lagged predictive variables. Explanatory variables are a constant, RFx, HB3, TS, DS, DY, Infl, ExMkt, SMB, HML, UMD, and sentiment. The full data set is 456 observations from 1/1963 to 12/2000. Each panel has k fewer observations due to construction of long-horizon returns. The reported coefficients are bias-adjusted from a simulation. See the appendix for additional details.

$$(r_{t+1} + \cdots + r_{t+k})/k = \alpha(k) + \mathbf{\Theta}'(k)\mathbf{z}_t + \beta(k)S_t + \varepsilon_t^{(k)}$$

the 36 portfolios, estimated separately. Table 3 reports the bias-adjusted coefficient estimates. They are almost universally negative as predicted by the alternative hypothesis and tend to be most negative for the larger firms and low book/market (growth) firms. A comparison of the panels shows that coefficients for horizons of a year or more are almost always more negative than those for the 6-month horizon. This pattern is consistent with limits to arbitrage hindering the ability of investors from profiting on mispricing that might persist for a significant amount of time.

TABLE 4 Significance of Sentiment in Long-Horizon Regressions

	Low	BM 2	BM 3	BM 4	High	All
A. 6-Month Horizon						
Small	.7873	.3910	.3125	.5169	.4372	.5476
Sz 2	.2522	.2552	.1601	.2429	.4347	.1223
Sz 3	.1105	.1432	.3641	.2820	.4158	.1022
Sz 4	.1866	.1969	.2661	.4786	.5314	.1655
Big	.1349	.1034	.3311	.4140	.3628	.2260
All	.1414	.0936	.2910	.4636	.4732	.3595
B. 12-Month Horizon						
Small	.6514	.2371	.1843	.2905	.2650	.3755
Sz 2	.1857	.1670	.0989	.0927	.2474	.1234
Sz 3	.0853	.1323	.1922	.1089	.1711	.0841
Sz 4	.0864	.1094	.1083	.1668	.2288	.0735
Big	.0195	.0188	.0867	.0393	.1080	.0112
All	.0274	.0207	.0804	.0818	.1826	.0422
C. 24-Month Horizon						
Small	.6764	.5883	.6050	.6827	.7787	.9512
Sz 2	.3462	.5611	.4607	.3235	.7311	.4942
Sz 3	.2127	.2906	.5654	.4657	.3912	.3251
Sz 4	.0777	.1044	.2335	.4412	.5656	.1367
Big	.0068	.0029	.0567	.0711	.0294	.0020
All	.0133	.0069	.0879	.2232	.2373	.0238
D. 36-Month Horizon						
Small	.2377	.8484	.8126	.8790	.9045	.6604
Sz 2	.5900	.7436	.5915	.2873	.9004	.6723
Sz 3	.2228	.2365	.4054	.5021	.7577	.3385
Sz 4	.0365	.0586	.0880	.3266	.6654	.0697
Big	.0009	.0004	.0077	.0166	.0210	.0001
All	.0019	.0013	.0192	.1132	.2718	.0035

Note.—This table shows the significance of sentiment from regression of k-period log returns on lagged predictive variables. Explanatory variables are a constant, RFx, HB3, TS, DS, DY, Infl, ExMkt, SMB, HML, UMD, and sentiment. The full data set is 456 observations from 1/1963 to 12/2000. Each panel has k fewer observations due to construction of long-horizon returns. The p-values are constructed from the distribution of the bias-adjusted coefficient estimates obtained by simulation to correct for problems associated with overlapping observations. See the appendix for additional details.

$$(r_{t+1} + \cdots + r_{t+k})/k = \alpha(k) + \mathbf{\Theta}'(k)\mathbf{z}_t + \beta(k)S_t + \varepsilon_t^{(k)}$$

Table 4 gauges the statistical significance of the sentiment variable.[11] These p-values are based on the empirical distributions obtained from the simulation. The test statistic is constructed as the bias-adjusted coefficient divided by the standard deviation (across simulations) of the coefficient estimate. We find these statistics adhere very closely to the standard normal distribution (panel A of figure 4 provides an

11. In unreported tables, we also examine the regression R^2 and conventional t-statistics for the long-horizon regression. For R^2, the general patterns are that predictability increases

TABLE 5 Joint Significance of Sentiment in Long-Horizon Regressions

	Low	BM 2	BM 3	BM 4	High	All
A. Wald Test for Joint Significance						
Small	2.5332	2.0355	1.9559	1.2605	1.7064	1.6520
Sz 2	1.8588	2.0181	3.0433	3.3161	1.6441	2.7889
Sz 3	3.6232	3.0979	2.3201	2.9193	2.1697	3.4967
Sz 4	5.3559	4.4025	4.4228	2.6852	1.6655	4.8433
Big	15.0748	17.7213	8.4798	8.8208	5.9554	20.8386
All	11.9338	13.7749	7.0637	5.0150	2.2370	11.3850
B. *p*-value for Wald Test						
Small	.6387	.7292	.7439	.8680	.7896	.7994
Sz 2	.7617	.7324	.5506	.5064	.8008	.5937
Sz 3	.4594	.5416	.6771	.5714	.7046	.4784
Sz 4	.2527	.3543	.3518	.6118	.7970	.3038
Big	.0045	.0014	.0755	.0657	.2025	.0003
All	.0179	.0080	.1326	.2858	.6923	.0226

NOTE.—This table shows the joint significance of sentiment from regression of k-period log returns on lagged predictive variables. Explanatory variables are a constant, RFx, HB3, TS, DS, DY, Infl, ExMkt, SMB, HML, UMD, and sentiment. The full data set is 456 observations from 1/1963 to 12/2000. Each panel has k fewer observations due to construction of long-horizon returns. The joint test is across horizons for a given portfolio. The p-values are constructed from the distribution of the bias-adjusted coefficient estimates obtained by simulation to correct for problems associated with overlapping observations. See the appendix for additional details.

$$(r_{t+1} + \Lambda + r_{t+k})/k = \alpha(k) + \mathbf{\Theta}'(k)\mathbf{z}_t + \beta(k)S_t + \varepsilon_t^{(k)}$$

illustration). Note that we use two-tailed tests, although the alternative hypothesis justifies one-sided tests. Significance using one-sided tests are twice as strong as what is reported in the tables.

The pattern of significance basically matches the magnitude of the coefficients. For larger firms or low book/market firms, sentiment is a significant predictor of future returns at the 1-, 2-, and 3-year horizon. Since these portfolios represent most of the market capitalization,[12] the market portfolio is also significantly affected by sentiment (significant at the 5% level for 1 and 2 years and 1% level for 3 years). That these portfolios are the ones most influenced by sentiment is interesting in light of the conventional wisdom that sentiment would most affect

with horizon, although most of the increase comes in moving from 6 to 12 months. Predictability tends to be lowest for small and low book/market firms, although there is no monotonic size or book/market relation. If we use Hansen and Hodrick (1980) standard errors to form conventional t-statistics, the overall patterns of significance are similar to those reported in table 4. We find these t-statistics to be less reliable since the standard errors often behave poorly. This results in many extreme statistics in the simulation, which make the empirical critical values large in magnitude (panel A of figure 4). Still, the large size, low book/market firms, and the overall market portfolio have significant coefficients at the 5–10% level. These tables are not reported to conserve space but are available upon request.

12. Fama and French (1993, table 1) report that the two portfolios with the stocks in largest capitalization quintile and lowest two book/market quintiles account for 46% of the overall market capitalization.

TABLE 6 Economic Magnitude of Sentiment in Long-Horizon Regressions

	Low	BM 2	BM 3	BM 4	High	All
A. 6-Month Horizon						
Small	−.5580	−1.5166	−1.5084	−0.9147	−1.1597	−.9698
Sz 2	−2.0871	−1.6685	−1.8598	−1.4494	−1.0447	−2.2535
Sz 3	−2.6035	−1.9689	−1.1192	−1.2469	−1.0568	−2.1000
Sz 4	−1.9075	−1.6364	−1.2719	−.7728	−.7550	−1.6297
Big	−1.6593	−1.6735	−.8795	−.7510	−.9372	−1.0936
All	−1.7255	−1.7904	−1.0025	−.6868	−.7904	−.8842
B. 12-Month Horizon						
Small	−1.4137	−3.0923	−2.9851	−2.1865	−2.4722	−2.1549
Sz 2	−3.6243	−2.9668	−3.3072	−3.1484	−2.3091	−3.4177
Sz 3	−4.1656	−3.0315	−2.3690	−2.8422	−2.6994	−3.3520
Sz 4	−3.7554	−3.1089	−2.7753	−2.2902	−2.1602	−3.2161
Big	−4.0967	−3.7819	−2.3628	−2.8943	−2.5385	−3.4354
All	−4.0467	−3.7774	−2.4615	−2.4782	−2.1979	−2.9053
C. 24-Month Horizon						
Small	1.7297	−1.8664	−1.5230	−1.1494	−.8305	−.1937
Sz 2	−3.4836	−1.6374	−1.9368	−2.4505	−.9027	−2.0582
Sz 3	−4.0465	−2.7858	−1.4034	−1.7016	−2.2740	−2.5554
Sz 4	−5.3837	−4.2942	−2.8157	−1.7756	−1.4065	−3.6736
Big	−7.0829	−7.0181	−3.6970	−3.4924	−4.9492	−6.0557
All	−6.4738	−6.2960	−3.3395	−2.3852	−2.6400	−4.6365
D. 36-Month Horizon						
Small	5.8000	.7355	−.7855	−.4652	.3890	1.6052
Sz 2	−2.2803	−1.0270	−1.5739	−3.0425	−.3713	−1.4286
Sz 3	−4.4351	−3.4877	−2.3215	−1.7958	−.9309	−2.7973
Sz 4	−7.5741	−5.8783	−4.7302	−2.5601	−1.2012	−5.1381
Big	−11.4787	−10.7907	−6.2390	−5.4918	−6.4988	−9.6607
All	−10.2450	−9.3471	−5.5354	−3.6213	−2.9056	−7.3289

NOTE.—This table shows the economic magnitude of a one standard deviation shock to sentiment. Results are based on bias-adjusted coefficients from regression of k-period log returns on lagged predictive variables. Explanatory variables are a constant, RFx, HB3, TS, DS, DY, Infl, ExMkt, SMB, HML, UMD, and sentiment. The full data set is 456 observations from 1/1963 to 12/2000. Each panel has k fewer observations due to construction of long-horizon returns.

$$(r_{t+1} + \cdots + r_{t+k})/k = \alpha(k) + \mathbf{\Theta}'(k)\mathbf{z}_t + \beta(k)S_t + \varepsilon_t^{(k)}$$

smaller stocks. Perhaps one reason for the lack of significance for smaller stocks is that the sentiment variable we use applies to the market as a whole: The newsletters on which our variable is based are forecasts for the overall market. Therefore, sentiment may affect small stocks more strongly, but our sentiment data do not allow us to address this. It seems plausible that growth firms are more influenced by sentiment than value firms since they typically are more difficult to value. If there are not well-established valuation benchmarks, then investors may be more likely to follow the pack.

Since the results across horizons are correlated, it is most appropriate to judge significance based on a joint test of the coefficients. Table 5 provides Wald tests of the null hypothesis that $\beta(6) = \beta(12) = \beta(24) = \beta(36) = 0$ for a given portfolio. These tests confirm that large and value stocks are significantly related to sentiment, as is the market portfolio.

To understand the economic magnitude of the coefficients on sentiment, it is helpful to refer to table 6. This table takes the bias-adjusted coefficient and multiplies it by the horizon and the standard deviation of sentiment (22%). The values in the table indicate the effect of a one standard deviation increase in sentiment on the return over the indicated horizon (in percent). For example, a one standard deviation increase in sentiment is associated with a reduction in the return on the large size, low book/market portfolio of 7.1% over a 2-year period. The average (simple) 2-year return for this portfolio is about 25%, so this is an economically significant reduction. Moreover, it is not so severe as to be implausible. Even if institutional investors are aware that future returns may be reduced by 7%, they may not be willing to try to exploit this mispricing since this is not an arbitrage opportunity. There is a very real chance the position may move against them in the interim, and they are likely to have their performance evaluated before this "convergence" strategy is expected to pay off.

V. Sentiment and Pricing Errors

In this section, we perform two types of analyses to examine the impact of sentiment on asset valuation. Both analyses rely on the pricing errors from the Bakshi and Chen (2001) model, which we use to form a market valuation and model valuation for a quasi-Dow index. The first test is simply to regress the pricing errors $e_t = (P_t - P_t^*)/P_t^*$ (market minus model) on sentiment and the control variables. The second test exploits the fact that the market and model valuations should be cointegrated.

A. Pricing Errors Regressions

The simple question we have in mind is: Can sentiment explain the pricing errors? We start by regressing the pricing errors on a constant, sentiment, and the return on the quasi-Dow:

$$e_t = \alpha + \beta S_t + \delta R_{\mathrm{Dow},t} + \varepsilon_t \tag{3}$$

The last variable is used as a control to pick up misspecification of the model. For example, if the model inputs, such as earnings growth forecasts, are delayed due to reporting lags, the market price might react to that information before the model. This would show up in both the pricing error and the return.

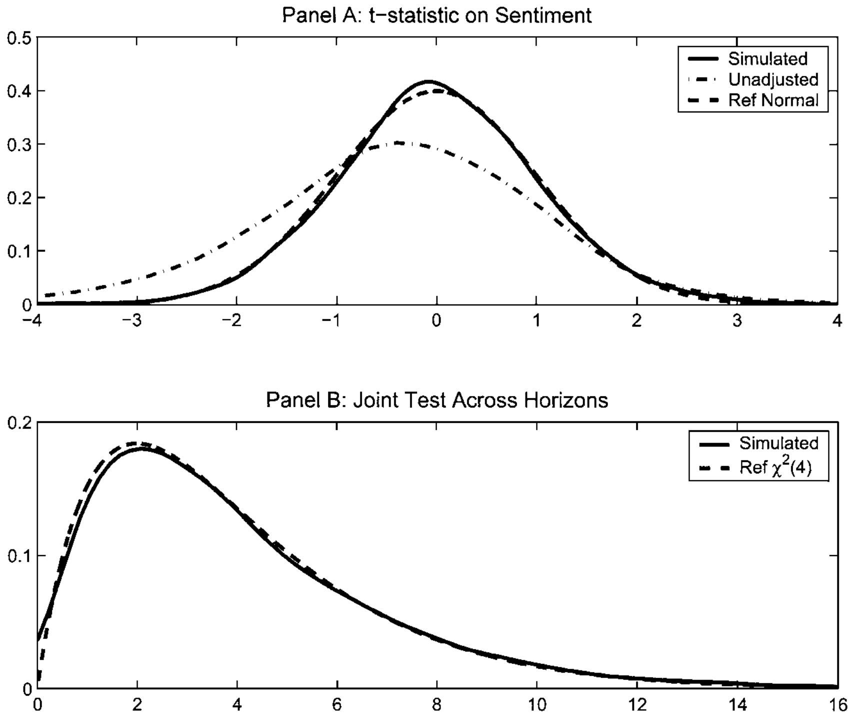

FIG. 4.—Example of distribution of test statistics. This figure shows the distribution of the test statistics from simulation for the value-weighted market portfolio. Panel A shows the distribution of the *t* test that the sentiment coefficient is zero for the 24-month horizon. The solid line is the adjusted statistic from the simulations, the dash-dotted line is the conventional (unadjusted) statistic, and the dashed line is a reference normal. Panel B shows the distribution of the joint test that coefficients on sentiment are zero for all four horizons. The reference χ^2 distribution with 4 degrees of freedom is also shown for comparison.

As the pricing errors are highly persistent (autocorrelation of 0.9), serial correlation in the regression residuals is a problem. We correct for this serial correlation in three ways. First, we use a Newey and West (1987) correction with 24 lags. This type of correction may not be adequate when the residuals are highly autocorrelated so we also use the Cochrane-Orcutt procedure and maximum likelihood with AR(1) errors.

The results of these regression are in table 7. The row labeled ρ indicates the autocorrelation of the residuals for the Newey-West regression and the estimated autocorrelation coefficient in the other two regressions. For all three of these regressions, ρ is 0.89, and the *t*-statistics of 29 indicate it is highly significant.[13] For simplicity, most

13. Additional tests estimating various ARMA(p, q) models for the residuals from ordinary least squares indicates the AR(1) specification is appropriate.

TABLE 7 Dow Pricing Error Analysis

	Newey–West (24 lags)		Cochrane–Orcutt		ML AR(1)	
Const	−.6299 (−.3155)	−992.0765 (−1.6475)	.0255 (.0085)	−62.8978 (−.3911)	−.2537 (−.0905)	−68.4954 (−.4264)
S	.1302 (1.8383)	.1338 (2.1532)	.0854 (3.4560)	.0767 (2.8437)	.0857 (3.4699)	.0780 (2.8993)
R_{Dow}	−.0504 (−.2550)	1.1872 (2.5469)	−.0026 (−.0445)	.1234 (.6911)	−.0028 (−.0486)	.1179 (.6614)
RFx		9.0546 (1.3752)		−2.3794 (−.5729)		−2.4019 (−.5785)
HB3		4.3642 (.5060)		1.4280 (.4145)		1.4827 (.4311)
TS		.0116 (.8165)		.0120 (1.4743)		.0115 (1.4154)
DS		.1746 (6.1171)		−.0576 (−2.0908)		−.0584 (−2.1343)
DY		−8.8921 (−3.3788)		−7.1731 (−2.6379)		−6.5361 (−2.5339)
CPI		9.9456 (1.6519)		.8963 (.5598)		.9455 (.5912)
ExMkt		−1.7101 (−4.4358)		−.3076 (−1.6101)		−.2896 (−1.5315)
SMB		.1795 (.4682)		−.1001 (−.7710)		−.0973 (−.7497)
HML		−.3505 (−1.2787)		−.1628 (−1.2870)		−.1593 (−1.2604)
UMD		.1611 (.6348)		−.0427 (−.5047)		−.0440 (−.5196)
ρ	.8863	.7530	.8896 (29.7311)	.9449 (44.0527)	.8864 (29.2298)	.9419 (42.7924)
$\bar{R}^2$	.0368	.2426	.0421	.0863	.0422	.0840

NOTE.—This table shows the results of regressing pricing errors for the “quasi-Dow” on sentiment and control variables. Three methods of correcting for the strong autocorrelation in pricing errors are used: Newey-West with 24 lags, Cochrane-Orcutt, and maximum likelihood with AR(1) residuals. Regressions use 235 observations from January 1979 through July 1998. The t-statistics are reported in parenthesis. For the Newey-West regressions, ρ is the autocorrelation of the residuals.

of the following discussion focuses on the Cochrane-Orcutt regression (the maximum likelihood regression is nearly identical). For all three regressions, the coefficient on sentiment is positive and significant (t-statistic of 3.5, or 1.8 for the Newey-West regression). The positive sign is as expected. When investors are optimistic, the market valuation is higher than the intrinsic value.

The first three regressions leave open the possibility that the sentiment variable simply picks up some other rational factors. We therefore include the set of control variables from the long-horizon regressions,

$$e_t = \alpha + \beta S_t + \delta R_{\text{Dow},t} + \mathbf{\Theta}'\mathbf{z}_t + \varepsilon_t \tag{4}$$

Even in the presence of the control variables, the coefficient on sentiment is positive and significant in all three regressions (the t-statistic is 2.8, or 2.2 for the Newey-West regression). Although we do not interpret the control variables individually, they do increase the adjusted R^2 and the default spread and dividend yield enter significantly.

The pricing error regressions show that, even controlling for common factors associated with rational asset pricing, some mispricing is explained by investor sentiment. This finding is robust to various forms of serial correlation correction and alternative regression specifications. We next explore this issue with another approach to see if we obtain the same conclusions.

B. Cointegration

In a cointegrated system, two (or more) variables are individually integrated but a linear combination of those variables is not integrated. We have a very natural environment for such a system. Since the market value of an index can be viewed as the sum of permanent shocks (and a drift), it is reasonable to expect that it is integrated. By the same logic, the intrinsic (model) value of the index should be integrated as well. Yet, the difference of the log series should not have permanent components. That is $p - p^*$ should be integrated of order zero, making p and p^* cointegrated with the cointegrating vector $[1\ -1]'$. The error correction interpretation of cointegration says that, when the market is overvalued, it will adjust toward its equilibrium value based on the sensitivity to the error $p - p^*$.

In this section, we explore this idea. We first establish that p and p^* are cointegrated, then estimate the error correction representation of the cointegrating regressions. In these regressions, we include sentiment and the control variables to see if sentiment can explain any of the deviations from intrinsic value.

To test for the integration of p and p^*, we estimate augmented Dickey-Fuller (ADF) regressions including a constant, a time trend,

TABLE 8 Tests for Integration and Cointegration

A. ADF Tests

Variable Z_t		Critical Values 1%	5%	10%
p	−2.1001	−3.9942	−3.4229	−3.1398
p^*	−3.2709	−3.9942	−3.4229	−3.1398
$p - p^*$	−3.8531	−3.4456	−2.8418	−2.5731
$p + \hat{\theta}p^*$	−3.6782	−3.9914	−3.4154	−3.1359

B. Johansen Tests $p - p^*$

	Lags				Critical Values		
	1	2	3	4	1%	5%	10%
Trace	15.0823	17.3196	16.6898	18.8019	19.9349	15.4943	13.4294
Eig	14.6439	16.7486	16.0755	18.1525	18.5200	14.2639	12.2971

NOTE.—Panel A reports augmented Dickey-Fuller tests of the hypothesis that the variable is integrated of order 1. The ADF regressions for $p - p^*$ contains a constant. The other ADF regressions include a constant and time trend. The ADF regressions for p and p^* include 12 lagged changes of the dependent variable to account for serial correlation. The $p - p^*$ and $p + \hat{\theta}p^*$ regressions include a single lag. The critical values shown for $p + \hat{\theta}p^*$ assume $\hat{\theta}$ is known. The asymptotic critical values accounting for the estimation of θ are −4.32, −3.78, and −3.50. Tests that fail to reject the null hypothesis indicate the variable is integrated, so for $p - p^*$ and $p + \hat{\theta}p^*$, rejection of the null hypothesis indicates cointegration. Panel B reports the trace and eigenvalue statistics for Johansen's tests of cointegration using one to four lags. LR and AIC tests indicate one lag, which BIC suggests two lags are needed. The data are 234 observations from February 1979 through July 1998.

and 12 lags of changes to the dependent variable to correct for serial correlation.[14] Panel A of table 8 contains the results for this test. The null hypothesis of integration is not rejected for p. For p^*, the null hypothesis of integration is rejected at the 10% level, although integration is not rejected at alternative lag lengths in the ADF test. Given the strong a priori theoretical reasoning for integration of p^*, the fairly weak evidence against the null hypothesis, and the sensitivity to the choice of lags, we proceed assuming that the intrinsic value series is also integrated.

We next test for cointegration in two ways. Given our strong theoretical prior on the cointegrating vector, we follow the recommendation of Hamilton (1994, p. 582) and estimate the ADF regressions for $p - p^*$. In this regression, a time trend is not included, since the errors do not appear to drift up over time (see figure 1) and only one lagged change in the dependent variable is included based on specification tests (although the results are robust to other lag choices). Panel A shows that the null of integration is strongly rejected, indicating the

14. In testing for the number of lags, there is evidence that lags 12 and 24 are significant for p^*. We present the more parsimonious regression here, although the tests still fail to reject the null hypothesis with the additional lags.

TABLE 9 Cointegration Regressions

	Δp				Δp*			
Const	.0060	(1.3621)	.0201	(.0205)	.0052	(.8108)	3.4344	(2.0220)
S	.0006	(4.5498)	.0003	(2.1903)	−.0004	(−2.1045)	−.0003	(−1.1945)
RFx			−.0390	(−1.8184)			.0440	(1.1653)
HB3			.0512	(2.7996)			−.0114	(−.3496)
TS			−.0000	(−1.0419)			−.0000	(−.3458)
DS			.0003	(3.2170)			−.0002	(−1.4117)
DY			−.0314	(−4.0993)			.0171	(1.2321)
CPI			.0009	(.0924)			−.0345	(−2.0579)
L(ExMkt)			−.0004	(−.8421)			−.0010	(−1.0894)
SMB			−.0018	(−2.1155)			.0004	(.2577)
HML			−.0063	(−7.9258)			.0057	(3.7019)
UMD			.0004	(.6547)			.0003	(.2977)
Δ(p*)	.3563	(9.3280)	.3122	(9.6831)				
Δ(p)					.7728	(9.3280)	.9601	(9.6831)
$L(p)$	−.0726	(−3.4734)	−.0933	(−4.6891)	0.1181	(3.8581)	.1226	(3.4384)
$L(p^*)$	.0727	(3.4949)	.0775	(4.0623)	−.1174	(−3.8542)	−.1201	(−3.5635)
θ	−1.0008		−0.8300		−1.0061		−1.0207	
$\bar{R}^2$	.3166		.5500		.2818		.3296	

Note.—This table shows the results from estimating cointegration regressions with additional explanatory variables. The regressions use 234 observations from February 1979 through July 1998. The *t*-statistics are reported in parenthesis. The cointegrating vector is $[1\ \theta]'$. The coefficient on the own-lagged level is the error correction. $L()$ indicates the lag of a variable.

two series are cointegrated. We also check our assumption about the cointegrating vector by rerunning the ADF test using the cointegrating vector [1 −0.83]’ as estimated in the second column of table 9.[15] In this regression, a time trend is included since this linear combination of p and p^* does trend upward over time. Once again, the null hypothesis of integration is rejected at the 5% level, meaning the two series are cointegrated.

The second test of cointegration uses Johansen’s trace and eigenvalue statistics. The tests are of the null hypothesis that there are zero cointegrating relationships. A rejection of the null hypothesis is evidence that p and p^* are cointegrated. The results of these tests are reported in panel B of table 9. We report the test for VAR lags from 1 to 4 orders, although the conclusions are the same in all cases.[16] For each lag order and for either test, the null hypothesis is rejected at the 5% level in favor of the alternative that p and p^* are cointegrated.[17]

15. The cointegrating vectors for the other regressions in table 9 are not materially different from $[1\ -1]'$, so they are not reported.

16. The likelihood ratio or AIC tests for model order of the VAR indicate one lag is appropriate. The BIC test favors two lags.

17. We do not report the Johansen test for the null hypothesis of less than one cointegrating relationship. Rejection of this null hypothesis would mean both variables are I(0). In every instance, we fail to reject this hypothesis, supporting cointegration of p and p^*.

Having established the cointegration of these series, we then assess whether sentiment is marginally significant when added to the error correction version of the cointegrating regression. In particular, we estimate

$$\Delta p_t = \alpha_1 + \beta_1 S_t + \boldsymbol{\Theta}_1' \mathbf{z}_t + \varphi_1 \Delta p_t^* + \gamma_1 \left(p_{t-1} + \theta_1 p_{t-1}^* \right) + \varepsilon_{1,t} \quad (5)$$

and

$$\Delta p_t^* = \alpha_2 + \beta_2 S_t + \boldsymbol{\Theta}_2' \mathbf{z}_t + \varphi_2 \Delta p_t + \gamma_2 \left(p_{t-1}^* + \theta_2 p_{t-1} \right) + \varepsilon_{2,t} \quad (6)$$

The coefficient γ_1 represents the correction of p_{t+1} to the error $p_t + \theta_1 p_t^*$ and $[1\ \theta_1]'$ is the cointegrating vector.

The results of these regressions are in table 9. One pair of regressions excludes the control variables, the other pair includes them.[18] The first two columns are for the regressions with Δp as the dependent variable. In either case, sentiment is positive and highly significant. This indicates that, when investors are optimistic, the market valuation tends to be high, controlling for the intrinsic value and (possibly) other variables to proxy for rational factors.

The last two columns in the table are for the regressions with Δp^* as the dependent variable. Here, the question is whether the sentiment variable can explain the level of the model valuation. If we could perfectly measure sentiment and intrinsic value, we should not find a significant relation. Excessive optimism or pessimism should be unrelated to the intrinsic value. However, a significant relation may indicate our proxy for intrinsic value is plagued by misspecification, or that we have not adequately controlled for the rational part of sentiment. The regression without controls shows the sentiment coefficient is negative and significant, but we already have seen that sentiment captures some of the information in the control variables. Thus, the negative relation may simply be due to the information in these control variables, which should be related to the intrinsic value. When we include the controls in the regression, there is no longer a significant relation. Therefore, we conclude that the ability of sentiment to explain pricing errors is driven by its relation with p and not p^*.

In summary, this section attacks from two fronts the issue of whether sentiment affects the level of asset values. We find strong evidence that sentiment can explain deviations from intrinsic value even when controlling for rational factors. In addition, we show, in a cointegration framework, that sentiment is significantly related to the level of market valuation. In either approach, the results indicate that, when investors are optimistic, the market valuation often exceeds the intrinsic value.

18. Note that we lag the excess market return by one period, since it can explain almost all the variation in Δp by construction.

VI. Robustness

We undertake a host of robustness checks to examine if our results are driven by some admittedly arbitrary choices. We first consider a variety of alternative proxies for investor sentiment. Many of these have been examined previously in the literature, so we seek to show that our variable provides some incremental information. We then consider reasonable alternatives to the way in which we construct our sentiment variable. Finally, we discuss some other checks we have done to control for endogeneity and to examine robustness to the choice of rational asset pricing model used to measure pricing errors.

A. Alternative Sentiment Proxies

As discussed in the Introduction, a number of other researchers have considered alternative proxies for investor sentiment. Although few of these papers directly explore the valuation and long-horizon issues on which we concentrate, it is important to provide evidence that our measure of sentiment is not simply picking up previously documented results.

We consider seven alternative sentiment proxies, several of which have been used explicitly in this capacity in the literature.[19] First, we consider the discount on closed-end funds (CEFD). The role of this variable has been the source of considerable controversy, as witnessed by the paper by Lee, Shleifer, and Thaler (1991) and the response by Chen, Kan, and Miller (1993). Second, we examine the ratio of NYSE odd-lot sales to purchases (ODD). Third, we employ net mutual fund flows (FUNDFLOW), which measure the actual investment behavior of individual investors. These three variables have also been examined by Neal and Wheatley (1998) over long horizons, and we compare our results to theirs.[20] FUNDCASH measures the proportion of aggregate mutual fund assets held as cash. All else equal, bearish fund managers have large cash positions.[21] Next, we consider the ARMS Index, a popular measure of market sentiment among technical analysts. It is reported daily in the *Wall Street Journal,* which indicates that "Generally, an ARMS of less than 1.00 indicates buying demand; above 1.00 indicates selling pressure." Finally, we consider two IPO-related variables. IPON measures the number of IPOs during the month. Ljungqvist, Nanda, and Singh (2002) suggest firms time their issues during "hot markets," which are times of excessive optimism. IPORET is the monthly average of first-day IPO returns, which has similarly been suggested as a sentiment measure.

19. To conserve space, we define these variables only briefly here. Interested readers are referred to Brown and Cliff (2004) for a more complete discussion.

20. Neal and Wheatley (1998) use annual data from 1933 to 1993 for closed-end fund discounts and from 1960 to 1993 for net mutual fund flows.

21. FUNDCASH is available from only 1970 to 1997, so all analysis involving that variable has 336 observations.

Table 10 summarizes the three analyses for each of these variables. The table shows results where the proxies are included in the analysis one at a time. Results using all proxies simultaneously are not shown in the table but are discussed here. Panel A shows the economic significance of the sentiment proxies from the long-horizon regression analysis for the small- and large-capitalization portfolios (results for the value-weighted market are similar to the large stock portfolio and not included in the table). We find no evidence that the closed-end fund discount is related to subsequent stock returns. Neal and Wheatley (1998) find a significant relation for small stocks or a portfolio long in small stocks and short in large stocks. However, their results are weaker when also controlling for the average price of small stocks. Consistent with Neal and Wheatley (1998), we find little evidence that the odd-lot ratio can predict future returns. For FUNDFLOW, we find a significant positive relation to future returns on the large-size portfolio. The relation for the small-size portfolio is also positive but weaker statistically. The positive coefficient indicates high long-run returns following cash inflows. Neal and Wheatley (1998) find a significantly negative relation for the small-size portfolio and the size premium. The sign for their large-size portfolio is consistent with ours but insignificant. Perhaps our results differ from Neal and Wheatley (1998) due to the sample period or our inclusion of additional control variables. Overall, our results do not provide much support for the view that these variables do a good job of measuring sentiment.

Several of the other variables in the long-horizon regression are significant. FUNDCASH is positively related to the returns on large stocks. Apparently, as fund managers are holding large cash positions, the market performs well going forward. The results indicate the IPO variables are negatively related to small stock returns. IPON is jointly significant at the 10% level, while IPORET is significant at the 1% level. Small stocks appear to do poorly following these proxies for hot markets.

Panel B summarizes the results from the quasi-Dow pricing error regressions (based on the Cochrane-Orcutt regressions). The first row of the panel is for regressions without the control variables, the second row uses the control variables. In every instance, the coefficients are insignificant.

Panel C shows the sentiment coefficients from the Δp portion of the cointegration analysis. Again, the table shows the results with and without the control variables. In several cases, the sentiment proxies are related to pricing errors when the control variables are not included (ODD, FUNDFLOW, FUNDCASH, and ARMS). However, most of this evidence appears to be due to common variation with the control variables. After including the control variables, only FUNDFLOW and ARMS remain significant.

TABLE 10 Alternative Sentiment Proxies

	CEFD	ODD	FUNDFLOW	FUNDCASH	IPON	IPORET	ARMS
A. Long-Horizon Regressions							
6 mo (Small)	−1.4959	−2.4604	−2.7092	1.2966	−3.0890	1.9916	2.4601
12 mo (Small)	−2.4529	−4.9998	−1.6039	4.0478	−5.8148^{c}	2.3525	1.3366
24 mo (Small)	−2.1232	−10.9674^{b}	5.2196	1.3095	−5.5719	−5.8478^{b}	.9293
36 mo (Small)	−1.6373	−13.3465^{b}	6.6440^{c}	−.3601	−10.9886^{b}	−5.5024^{c}	−1.6855
Joint p (Small)	0.9944	.2136	.0884	.8971	.1010	.0058	.4353
6 mo (Large)	−.4713	.6749	−.8034	4.1447^{b}	−.4041	.9660	1.2202
12 mo (Large)	.0719	2.6242	.3372	9.2635^{a}	.6751	1.1662	1.2619
24 mo (Large)	3.6535	2.7463	6.7903^{a}	9.3707^{b}	5.4976^{c}	−1.0488	1.0057
36 mo (Large)	7.5271	1.5747	11.2679^{a}	9.9317^{c}	7.6727^{b}	.5063	−1.2261
Joint p (Large)	.8625	.8004	.0001	.0626	.2245	.2903	.4775
B. Pricing Error Analysis							
β (no **z**)	−.0610	.7745	−.2125	.2687	−.0191	.0298	.0649
β (with **z**)	−.1199	2.0569	−.0452	.7127	−.0293	.0130	−.6206
C. Cointegration Analysis							
β (no **z**)	.0003	.0158^{b}	.0172^{a}	−.0037^{b}	.0001	.0003	−.1735^{a}
β (with **z**)	−.0002	.0102	.0130^{a}	.0028	−.0002	.0000	−.1535^{a}

NOTE.—This table summarizes the key results using alternative measures of sentiment. CEFD is the closed-end fund discount, ODD is the ratio of odd-lot sales to purchases, FUNDFLOW is the net flow of funds into mutual funds, FUNDCASH is the percentage of mutual fund assets held as cash, IPON is the number of IPOs during the month, IPORET is the first-day return on IPOs during the month, and ARMS is the ARMS Index. Panel A shows the economic magnitude of a one standard deviation increase in sentiment on the cumulative return over the period for either the small- or large-size quintiles. The panel also shows the p-value for the test of joint significance across horizons. The long-horizon regressions include control variables, not shown to conserve space, and use the simulations discussed in Appendix A. Panel B shows the coefficient on the sentiment variable in the pricing error regressions for the quasi-Dow, as in table 7. Panel C reports the sentiment coefficient in the error correction cointegration regression, similar to the Δp columns of table 9. In panels B and C, the results are reported for regressions with and without the control variables (**z**). In all panels, a, b, and c superscripts indicate significance at the 1, 5, and 10% levels, respectively.

To summarize our findings on the robustness to alternative sentiment proxies, our survey data do withstand the test. Although several of the proxies are significant in the long-horizon regressions or the cointegration analysis, none of the proxies exhibits such strong and clear significance as our survey data. When including all variables together, our survey data retains its significance.[22] There is no denying some overlap in the information contained in our variable with these other proxies. However, it appears that our variable effectively captures much of the common information and also provides incremental explanatory power.

B. Construction of the Sentiment Variable

The next issue we address is whether our construction of the sentiment variable drives the results. Unfortunately, there is no theoretically correct way of formulating a sentiment index from the survey data on the fraction Bullish, Bearish, and Neutral. Our choice of the Bull-Bear spread is based solely on its popularity among technical analysts and the financial press. Consequently, we define four other measures based on the same underlying survey data. First, we consider the fraction of Bulls to those with an opinion, Bull/(Bull + Bear). Second, we consider a pair of measures, Neutral and Bull-Bear. By including Neutral as a second variable, we can distinguish between cases where there is indifference and cases where there is strong disagreement (e.g., the Bull-Bear spread is the same if all investors are Neutral or bulls and bears are split 50/50). Third, we consider using Neutral/(Bull + Bear) and the Bull-Bear spread. Scaling Neutral in this way puts more emphasis on the observations with extreme neutrality. Finally, we split the Bull-Bear spread into positive and negative values. We discuss this in the next subsection, so we skip it here.

The results of the three analyses are summarized in table 11. Panel A shows that all variants of bullish sentiment remain significant in the long-horizon regressions. For the two cases with a version of neutrality, there is little evidence that neutrality is significant. The Cochrane-Orcutt

22. We explore two ways of using the various proxies simultaneously. First, we repeat the three analyses by simply using the survey data and all proxies at once. In the long-horizon regression, our sentiment variable remains significant. In the pricing error analysis, the survey sentiment variable remains significant with or without the control variables. None of the other proxies is significant in either regression. The survey sentiment variable is also significant in the cointegration analysis, although several other proxies are significant as well (CEFD, ODD, FUNDFLOW, IPON, and ARMS). The second approach is to regress survey sentiment on the proxies and rerun each analysis first with the fitted value from the regression, then with the regression residuals. The fitted value picks up the common variation between the survey data and the other proxies, while the residual contains the information in the survey data incremental to the other proxies. The fitted value is significant in the long-horizon regression and the cointegration analysis but insignificant in the pricing error regressions. The residual is significant in the pricing error regressions but not the cointegration analysis or the long-horizon regressions.

TABLE 11 Robustness to Construction of Sentiment

A. Long-Horizon Regressions

Case	Variable	6 mos.	12 mos.	24 mos.	36 mos.
1	Bull/(Bull + Bear)	−.8726	−2.8862[b]	−4.6066[b]	−7.2759[a]
2	Neutral	−.8037	−1.6023	−3.6765[c]	−2.5574
2	Bull-Bear	−.7167	−2.5236[c]	−3.8625[c]	−6.7593[a]
3	Neutral/(Bull + Bear)	−.4146	−.8782	−2.5927	−2.4425
3	Bull-Bear	−.8361	−2.7316[c]	−4.2397[b]	−6.9181[a]
4	S^+	−1.0165	−2.1246	−4.7556[a]	−7.2758[a]
4	S^-	.1891	−1.1768	.0515	−.2703

B. Pricing Error Analysis

		No **z**		With **z**	
		Coeff.	*t*-Statistric	Coeff.	*t*-Statistic
1	Bull/(Bull + Bear)	12.5139	3.2206	11.0963	2.5997
2	Neutral	.0505	.8373	.0127	.1962
2	Bull-Bear	.0916	3.5492	.0788	2.7503
3	Neutral/(Bull + Bear)	2.9410	.9066	.7199	.2069
3	Bull-Bear	.0918	3.5706	.0790	2.7556
4	S^+	.0735	2.0503	.0787	2.1616
4	S^-	.1024	2.3001	.0733	1.5327

C. Cointegration Analysis

		No **z**		With **z**	
		Coeff.	*t*-Statistic	Coeff.	*t*-Statistic
1	Bull/(Bull + Bear)	.0820	4.3754	.0439	2.1262
2	Neutral	−.0005	−1.4914	−.0006	−2.0461
2	Bull-Bear	.0006	4.7976	.0003	2.2296
3	Neutral/(Bull + Bear)	−.0242	−1.3042	−.0282	−1.6814
3	Bull-Bear	.0006	4.7394	.0003	2.2237
4	S^+	.0004	2.0916	.0002	1.3436
4	S^-	.0008	2.5686	.0004	1.3925

Note.—This table summarizes the key results when using four methods of constructing sentiment. Case 1 uses the percentage Bullish divided by the percentage with a view (Bullish plus Bearish). Case 2 uses two variables, the percentage Neutral and the Bull-Bear spread. Case 3 uses the percentage Neutral divided by the percentage with a view, along with the Bull-Bear spread. Case 4 splits the sentiment measure into positive and negative values. Panel A shows the economic magnitude of a one standard deviation increase in sentiment on the cumulative return on the value-weighted market over the period. The *a*, *b*, and *c* superscripts indicate significance at the 1, 5, and 10% levels, respectively. The long-horizon regressions include control variables, not shown to conserve space, and use the simulations discussed in the appendix. Panel B shows the sentiment coefficients and *t*-statistics based on the Cochrane-Orcutt quasi-Dow pricing error regressions, similar to table 7. Panel C shows the sentiment coefficients and *t*-statistics from the cointegration analysis, as in the Δp columns of table 9. In panels B and C, the first two columns are without the control variables (**z**), the last two columns include the control variables.

regressions of quasi-Dow pricing errors are in panel B. Again, all versions of bullish sentiment remain significant and the neutrality variables are insignificant. Finally, panel C shows the results for the market mispricing part of the cointegration. The significance of the bullish sentiment variables remains intact, but there is some weak evidence that high neutrality is related to drops in the market. In particular, for the regression using Neutral and Bull-Bear and including the control variables the t-statistic on the neutrality coefficient is -2.04. Replacing Neutral with Neutral/(Bull + Bear) reduces the t-statistic to -1.68.

In summary, it does not appear that our choice of defining sentiment as the Bull-Bear spread drives our results. Any combination of two of the three categories of optimism produces similar results.[23] It is probably possible to construct variables using the Bull/Bear/Neutral information that performs "better" than our simple Bull-Bear spread. However, our goal is not to fish for the variable(s) that best explains the data (in the sample) but to show that our results are robust to this choice.

C. *Asymmetric Response to Optimism and Pessimism*

Empirical evidence on momentum and reversals (see, for example, Hong, Lim, and Stein 2000) as well as implications of the behavioral theories suggest that the importance of sentiment may be asymmetric. One reason for this has to do with limits to arbitrage. Practical limitations to short-selling activity may make it difficult for rational investors to prevent market prices from being pushed above their intrinsic value during periods of excessive optimism. On the other hand, when some investors are especially pessimistic, no similar frictions prevent arbitrageurs from taking the necessary long position.[24] A second source of asymmetry has to do with the direct implication of the DHS model that overconfidence drives investors to overreact to private information. Overvaluation (undervaluation) tends to follow a string of good (bad) news. Gervais and Odean (2001) show that, because investors are net long in equities, aggregate overconfidence tends to occur following gains. Taken together, these implications suggest an asymmetric relation between sentiment and valuations.[25]

Case 4 in table 11 provides some evidence on this issue by splitting S into positive (S^+) and negative (S^-) values. Panel A shows the long-horizon regression results for the value-weighted market portfolio. The negative part of sentiment is insignificant, but the positive part is significant at the 1% level at the 24- and 36-month horizons. The joint test of all eight coefficients (not reported in the table) is significant, with a p-value of 0.0368. Panel B shows that, in the pricing error regression,

23. We also consider the three pairwise combinations of Bull, Bear, and Neutral. Single variables at a time do not work as reliably and all three cannot be used simultaneously, since they are related by an identity.

24. Short-sale constraints are not a necessary condition for the DHS, BSV, or DSSW models, since risk alone may make investors unwilling to sell short even if they are able to do so.

25. Such an asymmetric relation is not necessarily evidence of a behavioral model. For example, Veronesi (1999) has a rational model where a piece of news has more price impact if it refutes inferences about the current state.

the positive part of sentiment is statistically significant with or without the control variables. The evidence on the negative part of sentiment is more questionable. Without the control variables, it is significant. However, when controlling for common information, negative sentiment is no longer significant. The evidence from the cointegration analysis (panel C) is weaker. Both variables are significant for the regression without controls, but become insignificant when including the control variables. Overall, these results are consistent with the asymmetric patterns predicted by the behavioral theories.

D. Other Robustness Checks

We offer brief comments on other robustness checks we have done. One concern is that the endogeneity between sentiment and market valuations may influence our pricing error regressions. For example, if bad news comes out during the month, prices may drop and investors may reduce their optimism. Both adjustments potentially come in response to an exogenous shock, and the new market price and new sentiment level are presumably determined together. We address this issue by simply modifying the orthogonality conditions in the regressions. Specifically, we replace the usual condition from the normal equations that says $E(\varepsilon_t \times S_t) = 0$ with $E(\varepsilon_t \times S_{t-1}) = 0$. The main results are insensitive to this change.

We also examine robustness to the valuation model used to form pricing errors. We have pricing errors on the S&P 500 Index from both the Bakshi and Chen (2001) and Sharpe (2002) models.[26] In each case, these data start in 1983. These tests (not reported in tables) provide evidence that our Dow results are not simply picking up misspecification of the valuation model. The pricing error regressions, as in table 7, have significant coefficients on sentiment for either set of S&P pricing errors. When repeating the cointegration analysis of tables 8 and 9, the sentiment coefficient in the market mispricing regressions is significant for either set of pricing errors when the control variables are not included. When adding the control variables, the Bakshi and Chen (2001) S&P errors are no longer significantly related to sentiment, although this result is hard to interpret since these pricing errors failed the cointegration test diagnostic.[27] Results from the Sharpe model remain significant. Since our results are largely robust to both an alternative index and an alternative valuation model, we feel our evidence is more suggestive of picking up market misvaluation rather than model misspecification.

26. We thank Steve Sharpe for kindly sharing these data with us. The pricing errors in the published version of his paper are quarterly, but he provided a monthly sample.

27. The S&P pricing errors from the Bakshi and Chen (2001) model fail the test for cointegration. The Sharpe (2002) errors pass the cointegration tests.

VII. Conclusions

Our analysis shows that a direct survey measure of investor sentiment predicts market returns over the next 1–3 years and this measure has the ability to explain deviations from intrinsic value as measured by other researchers' models of stock prices. In all cases, the significance of our results is robust to controlling for rational factors and to changes in the methodology.

There are (at least) two ways to interpret these findings. The more conservative interpretation is that we have identified some new factor related to asset valuation. This factor may be derived from investors' rational outlook for the market or from some other origin altogether, for example, unidentified risk factors. Regardless of the interpretation, our sentiment variable forecasts market returns over the next several years and helps to explain mispricings from a rigorous valuation model.

The bolder interpretation is that we actually used an accurate measure of investor sentiment and this measure is related to the level of stock prices. This finding has several important implications. First and foremost, our results support the important yet controversial behavioral theories that predict the irrational sentiments of investors do in fact affect asset price levels. Second, this suggests asset pricing models should consider the role of investor sentiment. Third, market regulators and government officials should be concerned about the potential for market bubbles or "irrational exuberance" if a sudden change in sentiment translates into a negative wealth shock that depresses economic activity. Finally, individual investors should be aware of the impact sentiment can have on both their own and money managers' investment strategies.

Appendix

Simulation Details

We regress k-period returns on sentiment and control variables,

$$(r_{t+1} + \cdots + r_{t+k})/k = \alpha(k) + \beta(k)S_t + \mathbf{\Theta}'(k)\mathbf{z}_t + \varepsilon_{t+k}^{(k)}. \quad (7)$$

Here, r_{t+1} is the log return from month t to $t+1$ and $\mathbf{z}_t$ is the vector of predictive variables known at time t. The use of overlapping observations induces a MA$(k-1)$ structure in the residuals under the null that $\varepsilon^{(1)}$ is serially uncorrelated. Hansen and Hodrick (1980) propose a correction for this induced correlation.

Two main problems are associated with the regression in (7). First, there is a bias in the coefficient estimates, since they include persistent independent variables that are predetermined but not strictly exogenous. Second, Hansen and Hodrick (1980) standard errors do not perform well when the degree of overlap is "large" relative to the sample size. Therefore, we perform a simulation to account for the bias and adjust the critical values used in inference.

The simulation estimates a VAR(1) for $y_t = [r_t\, S_t \mathbf{z}_t']'$.[28] The null hypothesis that sentiment does not predict returns is imposed on the coefficient matrix by setting the appropriate element to zero. The constant in the constrained model is adjusted to restore the original mean, and the residuals from the constrained model are saved. We bootstrap the residuals from the calibration regressions to account for heteroscedasticity and generate and discard an additional 100 observations to remove any startup effects. At each replication, the 456 simulated observations are used to form long-horizon returns and estimate (7) at each of the horizons. The coefficients at iteration i are saved along with the corresponding Hansen and Hodrick (1980) standard errors. This process is repeated for 10,000 artificial data sets to get a distribution of the values of $\hat{\beta}(k)$.

We calculate t-statistics for the test that $\beta(k) = 0$ by subtracting the bias and dividing by the standard deviation of the coefficient estimates. That is, the t-statistic at the ith iteration for horizon k is

$$t^i(k) = \frac{\hat{\beta}^i(k) - \overline{\hat{\beta}^i(k)}}{\mathrm{std}\left[\hat{\beta}^i(k)\right]},$$

where $\overline{\hat{\beta}^i(k)}$ and std $[\hat{\beta}^i(k)]$ indicate the mean and standard deviation of the values of $\hat{\beta}^i(k)$ across the 10,000 simulations. We find that t-statistics constructed in this fashion perform much better than those obtained by using the standard error of the coefficient.

Figure 4 provides a graphical illustration for the value-weighted market return. Panel A shows the distribution of t-statistics for the sentiment coefficient over the 10,000 simulations under the null hypothesis. The solid line represents the simulation-adjusted t-statistics (used throughout the long-horizon regression analysis in the text). Test statistics formed in this way adhere very closely to the normal distribution (dashed line). Also shown in panel A are conventional t-statistics (i.e., using Hansen and Hodrick 1980 standard errors and without adjusting for bias). These t-statistics are obviously negatively biased and also have an extremely thick left tail. For example, the conventional t-statistics fall below -1.96 13.62% of the time, compared to 2.37% for the adjusted t-statistics and 2.50% for a standard normal variable. The empirical distribution of the t-statistics provides small sample critical values for inference. Each of the 36 portfolios has its own simulation for bias and size adjustments.

Panel B of the figure shows the distribution of the joint test of significance across the four horizons. These tests estimate the covariance matrix of the vector of coefficients $\hat{\beta} = \lfloor \hat{\beta}(6)\, \hat{\beta}(12)\, \hat{\beta}(24)\, \hat{\beta}(36) \rfloor'$ as the covariance (across the 10,000 simulations) of $\hat{\beta}^i$. The joint test conforms closely to the reference χ^2 distribution with 4 degrees of freedom.

References

Bakshi, Gurdip, and Zhiwu Chen. 2001. Stock valuation in dynamic economies. *Journal of Financial Markets,* in press.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny. 1998. A model of investor sentiment. *Journal of Financial Economics* 49:307–43.

28. AIC and BIC model-selection criteria indicate the VAR(1) specification is appropriate.

Boudoukh, Jacob, and Matthew Richardson. 1994. The statistics of long-horizon regressions revisited. *Mathematical Finance* 4:103–19.
Brown, Gregory W., and Michael T. Cliff. 2004. Investor sentiment and the near-term stock market. *Journal of Empirical Finance* 11:1–27.
Campbell, John Y. 1987. Stock returns and the term structure. *Journal of Financial Economics* 18:373–99.
———. 1991. A variance decomposition for stock returns. *Economic Journal* 101:157–79.
Campbell, John Y., and Robert J. Shiller. 1988a. The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies* 1:195–228.
———. 1988b. Stock prices, earnings, and expected dividends. *Journal of Finance* 43: 661–76.
Carhart, Mark M. 1997. On persistence in mutual fund performance. *Journal of Finance* 52:57–82.
Chen, Nai-fu, Raymond Kan, and Merton Miller. 1993. Are the discounts on closed-end funds a sentiment index? *Journal of Finance* 48:795–800.
Clarke, Roger G., and Meir Statman. 1998. Bullish or bearish? *Financial Analysts Journal* (May–June):63–72.
Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam. 1998. Investor psychology and security market under- and overreactions. *Journal of Finance* 53:1839–86.
DeLong, J. Bradford, and Andrei Shleifer. 1991. The bubble of 1929: Evidence from closed-end funds. *Journal of Economic History* 51:675–700.
DeLong, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. 1990. Noise trader risk in financial markets. *Journal of Political Economy* 98:703–38.
Fama, Eugene. 1990. Term-structure forecasts of interest rates, inflation, and real returns. *Journal of Monetary Economics* 25:59–76.
———. 1998. Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics* 49:283–306.
Fama, Eugene, and Kenneth French. 1988. Dividend yields and expected stock returns. *Journal of Financial Economics* 22:3–25.
———. 1989. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics* 25:23–49.
———. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33:3–56.
Fama, Eugene, and G. William Schwert. 1977. Asset returns and inflation. *Journal of Financial Economics* 5:115–46.
Ferson, Wayne E., and Campbell R. Harvey. 1991. The variation in economic risk premiums. *Journal of Political Economy* 99:385–415.
Fisher, Kenneth L., and Meir Statman. 2000. Investor sentiment and stock returns. *Financial Analysts Journal* (March–April):16–23.
Gervais, Simon, and Terrance Odean. 2001. Learning to be overconfident. *Review of Financial Studies* 14:1–28.
Hamilton, James D. 1994. *Time series analysis.* Princeton, NJ: Princeton University Press.
Hansen, Lars Peter, and Robert J. Hodrick. 1980. Forward exchange rates as optimal predictors of future spot rates: An empirical analysis. *Journal of Political Economy* 88:829–53.
Hirshleifer, David. 2001. Investor psychology and asset pricing. *Journal of Finance* 56: 1533–98.
Hodrick, Robert J. 1992. Dividend yields and expected stock returns: Alternative procedures for inference and measurement. *Review of Financial Studies* 5:357–86.
Hong, Harrison, Terence Lim, and Jeremy C. Stein. 2000. Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *Journal of Finance* 55:265–95.
Hong, Harrison, and Jeremy C. Stein. 1999. A unified theory of underreaction, momentum trading, and overreaction in asset markets. *Journal of Finance* 54:2143–84.
Jegadeesh, Narasimhan, and Sheridan Titman. 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance* 48:65–91.
Keim, Donald B., and Robert F. Stambaugh. 1986. Predicting returns in the bond and stock markets. *Journal of Financial Economics* 17:357–90.
Lamont, Owen A., and Richard H. Thaler. 2001. Can the market add and subtract? Mispricing in tech stock carve-outs. Working paper, University of Chicago.

Lee, Charles M. C., James Myers, and Bhaskaran Swaminathan. 1999. What is the intrinsic value of the Dow? *Journal of Finance* 54:1693–1741.
Lee, Charles, Andrei Shleifer, and Richard Thaler. 1991. Investor sentiment and the closed-end fund puzzle. *Journal of Finance* 46:75–109.
Lintner, John. 1969. The aggregation of investor's diverse judgments and preferences in purely competitive security markets. *Journal of Financial and Quantitative Analysis* 4: 347–400.
Ljungqvist, Alexander P., Vikram Nanda, and Rajdeep Singh. 2002. Hot markets, investor sentiment, and IPO pricing. Working paper. New York University.
Malkiel, Burton G. 1999. Day trading, and its dangers. *Wall Street Journal*:A22.
Neal, Robert, and Simon M. Wheatley. 1998. Do measures of investor sentiment predict returns? *Journal of Financial and Quantitative Analysis* 33:523–47.
Newey, Whitney, and Kenneth West. 1987. A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55:703–8.
Ofek, Eli, and Matthew Richardson. 2001. DotCom mania: A survey of market efficiency in the Internet sector. Working paper. New York University.
Otoo, Maria Ward. 1999. Consumer sentiment and the stock market. Working paper, Federal Reserve Board of Governors.
Rashes, Michael S. 2001. Massively confused investors making conspicuously ignorant choices (MCI-MCIC). *Journal of Finance* 56:1911–27.
Richardson, Matthew, and Tom Smith. 1991. Tests of financial models in the presence of overlapping observations. *Review of Financial Studies* 4:227–54.
Richardson, Matthew, and James H. Stock. 1989. Drawing inferences from statistics based on multiyear asset returns. *Journal of Financial Economics* 25:323–48.
Shanken, Jay, and Ane Tamayo. 2001. Risk, mispricing, and asset allocation: Conditioning on the dividend yield. Worjking paper, University of Rochester.
Sharpe, Steven A. 2002. Reexamining stock valuation and inflation: The implications of analysts' earnings forecasts. *Review of Economics and Statistics* 84:632–48.
Shiller, Robert J. 2000. *Irrational exuberance.* Princeton, NJ: Princeton University Press.
Shleifer, Andrei, and Robert Vishny. 1997. The limits of arbitrage. *Journal of Finance* 52:35–55.
Solt, Michael E., and Meir Statman. 1988. How useful is the sentiment index? *Financial Analysts Journal* (September–October): 45–55.
Stambaugh, Robert F. 1999. Predictive regressions. *Journal of Financial Economics* 54: 375–421.
Summers, Lawrence H. 1986. Does the stock market rationally reflect fundamental values? *Journal of Finance* 41:591–601.
Swaminathan, B. 1996. Time-varying expected small firm returns and closed-end fund discounts. *Review of Financial Studies* 9:845–87.
Veronesi, Pietro. 1999. Stock market overreaction to bad news in good times: A rational expectations equilibrium model. *Review of Financial Studies* 12:975–1007.
White, Eugene. 1990. The stock market boom and crash of 1929 revisited. *Journal of Economic Perspectives* 4:67–83.