Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Empirical Finance 11 (2004) 1–27
Journal of EMPIRICAL FINANCE
www.elsevier.com/locate/econbase

Investor sentiment and the near-term stock market

Gregory W. Brown[a,*], Michael T. Cliff[b]

[a] Kenan-Flagler Business School, Campus Box 3490, The University of North Carolina at Chapel Hill, CB 3490, McColl Building, Chapel Hill, NC 27599-3490, USA
[b] Krannert Graduate School of Management, Purdue University, 1310 Krannert Building, West Lafayette, IN 47907, USA

Accepted 20 December 2002

Abstract

We investigate investor sentiment and its relation to near-term stock market returns. We find that many commonly cited indirect measures of sentiment are related to direct measures (surveys) of investor sentiment. However, past market returns are also an important determinant of sentiment. Although sentiment levels and changes are strongly correlated with contemporaneous market returns, our tests show that sentiment has little predictive power for near-term future stock returns. Finally, our evidence does not support the conventional wisdom that sentiment primarily affects individual investors and small stocks.

JEL classification: G12; G14
Keywords: Investor sentiment; Return predictability

1. Introduction

For decades, an obvious schism has divided the field of finance: The success of the efficient market hypothesis (EMH) in explaining the lack of predictability in liquid asset returns has clashed with the traditional "search for value" undertaken by many finance practitioners. Only since the mid-1980s has there been a serious attempt to explore the possibility that liquid financial markets are not always as orderly as might be suggested by the efficient market advocates. For example, as the "noise trader" theories of Black (1986) and DeLong et al. (1990a) suggest, if some investors trade on a "noisy" signal that is unrelated to fundamentals, then asset prices will deviate from their intrinsic value.

* Corresponding author. Tel.: +1-919-962-9250; fax: +1-919-962-2068.
E-mail address: gregwbrown@unc.edu (G.W. Brown).

0927-5398/$ - see front matter
doi:10.1016/j.jempfin.2002.12.001

Unfortunately, most evidence supporting the noise trader theory is, at best, controversial.[1] Other recent empirical work also documents stock market anomalies such as market under- and over-reaction.[2]

More recently, attempts have been made to measure to what extent historical stock prices have reflected underlying fundamentals. Lee et al. (1999) calculate the intrinsic value for the Dow Jones Industrial Average (DJIA) from 1963 to 1996 and are able to predict subsequent market returns. Bakshi and Chen (2001) propose a model for the price of stocks (and stock indices) and present estimates showing persistent deviations from fair value for Dow 30 stocks, several technology stocks, and the Standard and Poor's 500 Index.

It is well known that market imperfections lead to observed prices deviating from underlying value (see Grossman and Stiglitz, 1980). Perhaps, the most important of these frictions for large US stocks is the costly nature of value-relevant information. Because so much of the information regarding the distribution of future stock returns is costly (or even impossible) to obtain and even then subject to interpretation, prices almost certainly trade inside a band around their value given perfect information. This judgment is supported by many empirical findings. For example, Seyhun (1986) shows that corporate insiders trading in their own stock earn superior returns. Donaldson and Kim (1993) find the DJIA is subject to "support" and "resistance" levels at multiples of 100. Siegel (1992) shows that neither changes in interest rates nor changes in future earnings account for the dramatic valuation changes around the October 1987 crash. He concludes that shifts in investor sentiment are correlated with market returns around the crash.

None of this is news to finance practitioners. It is hard to open a financial paper or turn on a market news program without hearing of an analyst having turned bullish or bearish on the market. In fact, for as long as markets have existed, pundits have had opinions concerning their relative value. Included in this group are esteemed economists and respected government officials. For example, Adam Smith (1776, p. 24) noted that, "The actual price at which any commodity is sold is called its market price. It may be either above, or below, or exactly the same as its natural price." More recently, Alan Greenspan (who spent 30 years on Wall Street) made news with his "irrational exuberance" speech, voicing his opinion that stock prices were too high.

In short, market watchers and participants seem to believe in "sentiment." But, what exactly is sentiment? Intuitively, sentiment represents the expectations of market participants relative to a norm: a bullish (bearish) investor expects returns to be above (below) average, whatever "average" may be. While there is no doubt such expectations exist,[3] the primary goal of this research is to investigate the role (if any) of sentiment in the price formation process for stock portfolios.[4] We conduct our analysis in two main steps.

First, we determine the relations between survey data for investor sentiment and other commonly cited "sentiment measures" such as the advance-decline ratio, short interest, and closed-end fund discounts. Many of these measures contain related information. Conse-

[1] See, among others, Lee et al. (1991) and Chen et al. (1993).

[2] See Daniel et al. (1998, Appendix A) or Hirshleifer (2001) for a summary of this literature.

[3] There are even services that collect the sentiments of market soothsayers and resell the data as a market indicator (favorable or contrarian depending on your view).

[4] Avery and Chevalier (1999) show that sentiment affects point spreads in the football betting market and that betting against these point movements is marginally profitable.

quently, we employ the Kalman filter and principal component analysis as means of extracting composite unobserved sentiment measures. We also separate potential sentiment measures into two groups, institutional and individual, to see if there are different classes of investor sentiment.

Next, we examine the ability of these sentiment measures to predict returns. Specifically, we explore the bi-directional relation between investor sentiment and near-term stock returns in a vector autoregression (VAR) framework using the composite sentiment measures discussed above.

Overall, our findings are consistent with many of the anecdotes regarding investor sentiment. Specifically, we document strong relations between many disparate measures of investor sentiment. This probably explains why certain indicators have developed a reputation as sentiment measures. Furthermore, we document that changes in the survey and our composite measures of investor sentiment are highly correlated with contemporaneous market returns. The empirical sticking point is that this correlation does not directly reveal the causal relation between sentiment and the market.[5] The more sophisticated VAR analysis reveals that market returns clearly cause future changes in sentiment. However, very little evidence suggests sentiment causes subsequent market returns. This is bad news for investors trying to use sentiment measures for short-term market timing. It appears that such strategies are not profitable during our reasonably long sample period.

Our paper is not the first to explore the role of investor sentiment in the stock market, but it is the most comprehensive study to date. A series of papers has debated about closed-end fund discounts as a measure of sentiment. Lee et al. (1991), Swaminathan (1996), and Neal and Wheatley (1998) claim closed-end fund discounts measure investor sentiment, while Chen et al. (1993) and Elton et al. (1998) provide evidence to the contrary. Neal and Wheatley (1998) also find that net mutual fund redemptions are useful in predicting the size premium. Brown (1999) shows that sentiment is closely related to closed-end fund price volatility. Barber (1994) shows that the behavior of Prime and Score premiums is consistent with a noise trader model. Finally, Clarke and Statman (1998), Otoo (1999), Simon and Wiggins (1999), and Fisher and Statman (2000) examine the usefulness of a variety of technical variables in predicting short-horizon market returns. Our analysis expands upon these and other papers by considering a more comprehensive set of sentiment proxies, using direct (survey) data on sentiment, examining the (statistical) causal relation among the variables, and finally by investigating the relation between sentiment and subsequent market returns. In a companion paper, Brown and Cliff (2004) explore long-run effects and find that optimism is associated with overvaluation and low subsequent returns as the valuation level returns to its intrinsic value.

2. Motivation

We start our analysis by providing an informal discussion of the ways in which we expect sentiment to affect market valuations and returns. These predictions provide a framework for the empirical analysis and interpretation of the results. A number of

[5] By causality, we mean in the statistical sense of Granger.

researchers, such as Grossman and Stiglitz (1980), Black (1986), DeLong et al. (1990a), Campbell and Kyle (1993), Barberis et al. (1998), Daniel et al. (1998), and Hong and Stein (1999) have more formally modeled the role of sentiment or investor behavior. However, these models are difficult to test directly since they typically involve sources of noise (sentiment), which are difficult to measure. Since the overall goal of the paper is an empirical examination of the importance of sentiment, our approach focuses more on practical implementation rather than on testing formal economic models.

Suppose there are two types of traders, fundamentalists, and speculators. The fundamentalists are characterized by the fact that, as a group, they have unbiased expectations of an asset's value. The group of speculators has a bias in its valuation of the asset. When the speculators believe that intrinsic value is greater (less) than the current price, they are bullish (bearish). From a theoretical viewpoint, bullishness measures the extent of the discount from intrinsic value. A bullish investor would expect not only a positive return, but also a return greater than the fundamentalists' required rate of return. In other words, an investor expecting a small return on the market, say one percent, would be viewed as bearish, even though they expect a slight gain. Unfortunately, measuring the deviations from intrinsic value or the return in excess of the required return are difficult tasks. Therefore, in our empirical work, we follow the conventional usage where bullish means an expected price increase and bearish means an expected price decline.

Conceptually, the speculators' bias, or excessive optimism or pessimism, is what we think of as sentiment. Since the two groups attach different values to the asset, the market price will be a weighted average differing from the intrinsic value. To have an equilibrium, we can appeal to a number of frictions, such as transactions costs, arbitrageur risk-aversion, or learning.[6] Lintner (1969) presents a model which formally describes the weighted-average valuation we have in mind.

So far, we have assumed that there is only one type of speculator. Anecdotal evidence suggests that market participants believe there are multiple types. For example, Solt and Statman (1988) describe the contrarian nature of the Investors Intelligence Bearish Sentiment Index (which measures the percentage of market newsletters that are bearish). By the nature of a contrarian strategy, the contrarians must believe that another group of investors is swayed by sentiment in the opposite (wrong) direction. In terms of our framework, we can describe two distinct types of speculators, institutions and individuals, that may have different systematic misvaluations.[7] For convenience, we can think of institutions as speculators who participate in the market for a living, and individuals as speculators whose primary line of business is outside the stock market.[8]

[6] Note that, because the relative holdings of investors and speculators are indeterminant and investors need not know their type, speculators need not lose their wealth to fundamentalists (see DeLong et al. (1989, 1990b), for a more thorough analysis). Since we have not specified the source of the bias for speculators, they can be viewed as acting rationally given their information set.

[7] We do not take a stand on which of the two make more serious errors. In this context neither institutions nor individuals are viewed as having superior information. Indeed, both are speculators in our framework, meaning they have incorrect assessments of the risky asset's value. To include informed agents into our framework, you could consider them as a subset of the fundamentalists with low (or no) variability in their (unbiased) valuation errors.

[8] This is consistent with many of the interpretations of contrarian indicators. See, for example, Fosback (1993).

To the extent that institutions and individuals respond differently to signals in the formation of their sentiments, we may be able to identify them (and any effects they have on prices) empirically. For example, it may be the case that both institutions and individuals exhibit significant bouts of sentiment but that only institutions have enough market power to affect prices. Similarly, since institutional investors make up a disproportionately larger part of the market for large stocks relative to individuals, perhaps institutional sentiment primarily affects large stocks and individual sentiment affects small stocks.

The intuition developed above implies that sentiment and contemporaneous returns should be positively related. Within a period, if investors are particularly optimistic, that should drive up prices. Of course, it is also possible that good returns during the period drive optimism, a "bandwagon" effect. Another interpretation is that the speculators do not affect prices. When they see stocks becoming a bargain (proxied in our empirical work as a negative return), they see a buying opportunity and become bullish. Thus, the "bargain shopper" effect predicts a negative relation between sentiment and contemporaneous returns.[9] Of these three alternatives, our evidence supports the "bandwagon" hypothesis: Returns and contemporaneous sentiment are strongly positively related, returns predict future sentiment, but sentiment does not predict future returns.

3. Data

In implementing our analysis, we have chosen to concentrate on market aggregates instead of individual stocks. While the effects of sentiment will aggregate across stocks to the market level, we run the risk that roughly as many stocks are affected by bullish sentiment as bearish sentiment and thus the aggregate affect is negligible. Our choice of market aggregates is primarily driven by data limitations. Many of the measures we examine are available only for the market as a whole (i.e., the surveys, closed-end fund discounts, advance-decline ratio, etc.) and without the benefits of the subsequent results, there are no assurances that stock-specific indicators of sentiment actually contain information about views.[10]

The subsequent empirical analysis is performed at both the monthly and weekly frequencies. Most of the monthly data range from March 1965 through December 1998, giving a total of 406 observations. A few variables are not available for the full sample so some of the analysis is performed on slightly shorter subsamples. Weekly data consist of 596 observations spanning July 24, 1987 through December 18, 1998. As we discuss below, not all the same data items are available in both samples. Univariate summary statistics for the data are in Table 1 and important features are mentioned in the data description below. Table 2 shows contemporaneous pairwise correlations for our primary variables. In few cases are the correlations larger than 0.50. However, as will be discussed in Section 4.1, more complicated linear relations among the data are prevalent.

---

[9] We thank the referee for pointing out this issue.

[10] However, our results suggest it may be possible to construct good stock-specific sentiment measures, an idea we leave to future research.

Table 1
Summary statistics

	Mean	S.D.	Skewness	Kurtosis	ρ_1	N
Panel A: monthly data						
SENT^p	9.1426	22.3928	0.1609	− 0.2146	0.72	406
$\Delta\mathrm{SENT}^p$	− 0.0563	16.7228	− 0.1969	0.7562	− 0.22	405
R_{BIG}	1.0206	4.2960	0.3255	2.1417	− 0.00	406
R_{SML}	1.2413	5.9250	− 0.4145	2.9609	0.17	406
R_{SOB}	0.0000	3.2698	0.4127	1.4895	0.08	406
ADV/DEC	1.0166	0.1976	0.4878	1.1139	0.16	406
ARMS	0.9763	0.1353	2.6239	20.7747	0.41	406
HI/LO	6.7005	14.8590	5.3098	34.4822	0.55	406
ΔMARGIN	0.0093	0.0329	− 0.4671	0.9142	0.56	406
ΔSHORTIR	0.0065	0.0776	− 0.0588	0.2520	− 0.09	406
SPECIAL	0.4527	0.0860	0.6969	− 0.3538	0.87	406
ODDLOT	1.6063	0.5097	0.4647	− 0.4120	0.90	406
CEFD	− 16.8767	12.6714	− 0.3197	− 1.3447	0.99	406
IPORET	16.2448	18.6725	1.8828	5.2946	0.57	406
IPON	30.9606	25.7443	0.8295	0.0620	0.86	406
FUNDFLOW	0.2858	0.9426	0.1957	0.8810	0.67	406
FUNDCASH	7.4902	2.4287	1.3343	2.4067	0.94	336
Panel B: weekly data						
SENT^p	5.3027	14.3575	− 0.3154	− 0.3614	0.94	596
SENT^a	7.3767	17.1369	− 0.3606	0.2252	0.69	596
$\Delta\mathrm{SENT}^p$	− 0.0007	4.9402	− 0.3730	1.5728	0.31	595
$\Delta\mathrm{SENT}^a$	0.0034	13.5715	− 0.1672	0.6207	− 0.32	595
R_{BIG}	0.2428	1.9750	− 0.6565	3.7154	− 0.05	596
R_{SML}	0.1669	2.1298	− 1.9438	16.6542	0.23	596
R_{SOB}	− 0.0000	1.3593	− 0.8087	7.6520	0.09	596
ADV/DEC	1.3108	0.7632	1.6875	5.9121	0.16	596
ARMS	0.9830	0.2321	4.7191	40.9120	0.34	596
HI/LO	4.0786	4.8566	2.2420	6.4354	0.80	596
SPECIAL	0.3939	0.0482	0.5743	1.6205	0.64	596
ODDLOT	1.3535	0.3617	0.8556	0.4958	0.85	596
SHORTSLS	0.4123	0.0983	0.5398	0.8286	0.38	596
MKTVANE	49.9413	9.6534	0.2428	− 0.2983	0.80	596
PUT/CALL	1.1171	0.1915	0.5461	2.7002	0.54	596
VOL	1.4327	0.3406	0.4971	0.1911	0.29	596
FUT^p	3.7768	4027.6334	0.3672	1.6122	0.14	596
FUT^a	126.8490	3989.1772	0.4295	3.0840	0.11	596
CEFD	− 6.5985	3.2134	0.1316	− 0.8184	0.99	596

This table shows summary statistics for the data used in the analysis. The full monthly sample contains 406 observations from March 1965 through December 1998. The full weekly data consist of 596 observations from July 1987 through December 1998. See the text for variable descriptions.

3.1. Direct sentiment measures

To empirically test our hypotheses, we require measures of investor sentiment. Fortunately, there are two surveys that directly measure the sentiment of market participants. The first is a survey conducted by the American Association of Individual

Investors (AAII). The association polls a random sample of its members each week,[11] beginning in July 1987. The sample size of the survey has varied between 125 and 500 with the number of weekly respondents varying between 26 and 422. The average number of respondents is 137 with a standard deviation of 69. The average response rate is 51% with a standard deviation of 15%. The association asks each participant where they think the stock market will be in 6 months: up, down, or the same. AAII then labels these responses as bullish, bearish, or neutral, respectively. On average the responses are almost evenly split with 36% of all responses bullish, 28% bearish, and 36% neutral. Since this survey is targeted towards individuals, we interpret it as primarily a measure of individual investor sentiment.

Investors Intelligence (II) compiles another weekly bull-bear spread by categorizing approximately 150 market newsletters. Each week, the newsletters are read and marked as bullish, bearish, or neutral again 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.[12] Many of the newsletters are issued less than once a week, in which case the most recent edition of each newsletter is included. This causes the Investors Intelligence series to have high autocorrelation (0.94 at one lag for the weekly sample). Since many of the authors of these newsletters are current or retired market professionals, we interpret the Investors Intelligence data as a proxy for institutional sentiment. The II data are also available on a monthly basis starting in 1965.

We use the surveys measures of the percentage of bullish investors minus the percentage bearish (bull-bear spread) to measure investor sentiment. For notational convenience, let $SENT^a$ denote the sentiment of amateurs (individuals) as measured by the AAII bull-bear spread and let $SENT^p$ represent the sentiment of professionals (institutions) as proxied by the II data. Fig. 1 shows the Investors Intelligence series is much smoother than the AAII series in the weekly data. Inspection of the figure shows that the two series share many features, but they are not identical. Indeed, the correlation between the two is 0.43.

3.2. Indirect sentiment measures

Stock market commentators often refer to certain financial variables as market weather vanes. We examine a number of these indicators, which we categorize into four main groups. Although many of these variables can be interpreted as bullish or bearish signals depending on whether you believe in a momentum or contrarian strategy, our primary interest is in establishing any relation between these variables and sentiment. A few of the variables have implications for other research areas (e.g., closed-end funds, IPOs, and

[11] Surveys are mailed to investors and returned by mail to AAII. Each week, AAII compiles the responses received during the week. There is a potential timing issue due to the delay between respondent's filling out the form, mailing, and tabulation by AAII. These data are shifted by a week to accommodate this delay.

[12] Investors Intelligence indicates that there are a relatively small number of people involved in categorizing the newsletters so there should not be a large problem associated with differing interpretations.

Table 2
Contemporaneous correlatations

Panel A: monthly data

		$SENT^p$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
2	$\Delta SENT^p$	0.36															
3	R_{BIG}	0.29	0.41														
4	R_{SML}	0.32	0.46	0.83													
5	R_{SOB}	0.15	0.20	0.00	0.55												
6	ADV/DEC	0.36	0.49	0.82	0.89	0.38											
7	ARMS	−0.36	−0.26	−0.54	−0.53	−0.15	−0.40										
8	HI/LO	0.31	0.11	0.24	0.36	0.29	0.39	−0.15									
9	ΔMARGIN	0.40	−0.16	0.02	0.02	0.02	0.00	−0.28	0.18								
10	ΔSHORTIR	−0.01	0.21	0.13	0.10	−0.01	0.07	−0.13	−0.02	−0.13							
11	SPECIAL	0.22	−0.00	−0.04	0.06	0.17	−0.04	−0.16	0.15	0.08	−0.16						
12	ODDLOT	0.17	0.09	0.12	0.13	0.06	0.15	−0.16	0.21	0.11	0.07	−0.24					
13	CEFD	−0.07	−0.01	0.09	−0.00	−0.14	0.08	−0.02	0.00	0.04	0.01	−0.61	0.02				
14	IPORET	−0.00	−0.06	0.14	0.21	0.18	0.10	−0.28	0.13	0.27	−0.07	0.20	−0.00	0.05			
15	IPON	0.12	−0.01	0.03	−0.08	−0.19	−0.03	−0.08	−0.03	0.27	−0.00	−0.22	−0.06	0.58	0.00		
16	FUNDFLOW	0.14	0.19	0.33	0.27	−0.02	0.27	−0.25	0.09	0.10	0.01	−0.19	−0.32	0.58	0.01	0.51	
17	FUNDCASH	−0.36	0.01	−0.02	0.02	0.06	−0.02	−0.07	−0.11	−0.20	−0.01	−0.14	0.40	−0.06	0.24	−0.33	−0.33

Panel B: weekly data

		$SENT^p$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
2	$SENT^a$	0.43																	
3	$\Delta SENT^p$	0.17	0.20																
4	$\Delta SENT^a$	− 0.08	0.40	0.03															
5	R_{BIG}	− 0.06	0.17	0.14	0.22														
6	R_{SML}	− 0.02	0.21	0.23	0.19	0.77													
7	R_{SOB}	0.04	0.13	0.18	0.03	− 0.00	0.64												
8	ADV/DEC	− 0.04	0.20	0.15	0.24	0.78	0.73	0.21											
9	ARMS	0.03	− 0.08	− 0.20	− 0.12	− 0.64	− 0.60	− 0.16	− 0.40										
10	HI/LO	0.44	0.44	0.22	0.07	0.18	0.29	0.23	0.32	− 0.14									
11	SPECIAL	0.28	0.12	0.12	− 0.09	− 0.05	− 0.00	0.06	− 0.03	0.07	0.09								
12	ODDLOT	− 0.47	− 0.22	− 0.02	− 0.01	− 0.04	0.03	0.09	− 0.04	− 0.11	− 0.24	− 0.13							
13	SHORTSLS	0.17	0.07	0.10	− 0.05	0.02	0.10	0.13	0.00	− 0.04	0.09	0.13	− 0.23						
14	MKTVANE	0.37	0.43	0.28	0.01	0.12	0.19	0.15	0.14	− 0.15	0.49	0.08	− 0.07	0.21					
15	PUT/CALL	0.06	0.08	0.04	− 0.04	− 0.09	− 0.08	− 0.02	− 0.17	− 0.10	− 0.02	− 0.06	− 0.20	0.10	0.16				
16	VOL	− 0.00	0.22	0.11	0.03	0.06	0.13	0.14	0.01	− 0.18	0.03	− 0.09	0.07	0.25	0.31	0.18			
17	FUT^p	0.09	0.13	0.14	− 0.03	0.01	0.06	0.09	0.05	− 0.03	0.14	− 0.04	− 0.01	− 0.02	0.10	− 0.07	0.07		
18	FUT^a	0.08	0.10	0.12	0.04	0.12	0.16	0.10	0.16	− 0.08	0.14	0.02	0.01	0.06	0.15	− 0.03	0.01	0.06	
19	CEFD	0.20	− 0.01	− 0.02	− 0.01	− 0.01	0.02	0.04	− 0.02	− 0.02	0.07	0.11	− 0.57	0.01	− 0.24	0.13	− 0.18	0.00	− 0.04

Pairwise correlations for selected variables used in the analysis. Most correlations use 406 observations from March 1965 to December 1998 for the monthly sample and 596 observations from July 1987 through December 1998 for the weekly sample. Correlations involving $\Delta SENT^a$, $\Delta SENT^p$, or FUNDCASH use fewer observations as indicated in Table 1. See the text for variable descriptions.

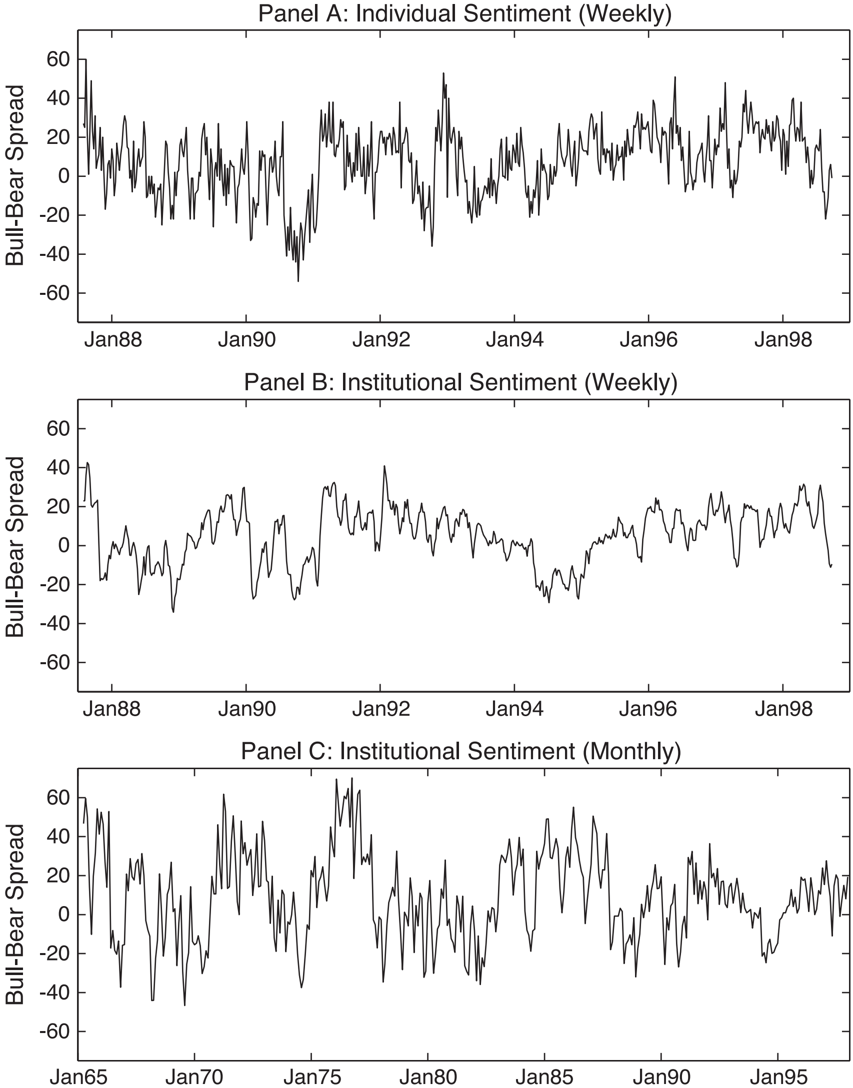

Fig. 1. The bull-bear spread from the sentiment survey data are shown here. Individual sentiment comes from the AAII; institutional sentiment is from II.

liquidity). The direction of the relation between sentiment and these variables is examined more closely.

3.2.1. Market performance

In the first group are variables which are based on recent market performance. One of the most common technical indicators is the ratio of the number of advancing issues to declining issues (ADV/DEC). We calculate this measure using data for NYSE issues. The

ARMS index is a modification of ADV/DEC, which incorporates volumes. This measure is the ratio of the number of advances to declines standardized by their respective volumes

$$\mathrm{ARMS}_t = \frac{\mathrm{Adv}_t/\mathrm{Adv\ Vol}_t}{\mathrm{Dec}_t/\mathrm{Dec\ Vol}_t}$$

The number of new highs to new lows (HI/LO) is also designed to capture the relative strength of the market. The three variables in this group are in both the weekly and monthly samples. Because the numerator and denominator of these variables tend to move in opposite directions, these data are truncated at zero and have a few extreme positive outcomes.[13]

3.2.2. Type of trading activity

A second category of indicators includes variables that relate to particular types of trading activity. At the monthly level, we have the percent change in margin borrowing (ΔMARGIN) as reported by the Federal Reserve. This measure is frequently cited as a bullish indicator since it represents investors using borrowed money to invest. Also available on a monthly basis, the percent change in short interest (ΔSHORTIR) is viewed as a bearish indicator. In the weekly sample, we have the ratio of short sales to total sales (SHORTSLS). For both weekly and monthly samples, we have the ratio of specialists' short sales to total short sales (SPECIAL). The argument is made that the specialists are well-informed and relatively savvy investors, so when their short-selling becomes relatively large, the market is likely to decline. Finally, the ratio of odd-lot sales to purchases (ODDLOT) is a bearish measure.[14]

3.2.3. Derivatives variables

The third category of variables relates to derivatives trading activity. Since the derivatives markets have become sizeable after the start of the monthly sample, these data are only available for the weekly analysis. The ratio of CBOE equity put to call (PUT/CALL) trading volume is widely viewed as a bearish indicator. The Commodities Futures Trading Commission (CFTC) reports the change in the net position in SPX futures by trader type. We use these data on non-commercial traders ($\mathrm{FUT}^{\mathrm{p}}$) as a proxy for institutional sentiment and activity by small traders ($\mathrm{FUT}^{\mathrm{a}}$) as a proxy for individual sentiment. As defined by the CFTC, non-commercial traders are essentially market professionals excluding financial and trading companies. Monthly forecasts of commodity market returns collected by Market Vane (MKTVANE) are an alleged bullish predictor of futures market behavior that ''is derived by tracking the buy and sell recommendations of

[13] In the subsequent regression analysis, we have reproduced the results after taking the log of all three ratios in order to make the data more normally distributed and the results are qualitatively similar. As inputs to the Kalman filter and principal component analysis, we use log transformations.

[14] Weekly data for ODDLOT is only available after February 24, 1989; consequently, we set weekly values equal to the monthly values for this period. This results in 66 interpolated values in the weekly sample of 596 observations.

leading market advisers" (see http://home.earthlink.net/~marketvane). Last, we construct a measure of expected volatility relative to current volatility. This is

$$\mathrm{VOL}_t = \ln(\mathrm{VIX}_t/\mathrm{SIG}_t)$$

where VIX is the S&P 100 Index option volatility and SIG is the realized volatility calculated from Open-High-Low-Close data on the S&P 100 Index.[15] Thus, a positive VOL measures higher anticipated volatility and can be interpreted as bearish. Except for $\mathrm{FUT}^{\mathrm{a}}$, we view these variables as primarily a measure of institutional sentiment because of institutions' larger presence in the derivatives markets.

3.2.4. Other sentiment proxies

A few other variables do not fall neatly within one of the above three categories. The closed-end fund discount (CEFD) is measured by taking the weekly or monthly equal-weighted average of all domestic equity fund discounts. A small literature has developed around the controversy over the closed-end fund discount as a measure of individual sentiment (see, for example, Lee et al. (1991)).

We also have monthly data on the net purchases of mutual funds (FUND FLOW). Neal and Wheatley (1998) find that this is useful in predicting the premium of small stocks over large stocks. In addition, we have data on the proportion of fund assets held in cash (FUNDCASH).[16] Presumably, a fund's cash holdings will be negatively related to its optimism about the market, ceteris paribus. We expect the cash holdings to be more correlated with institutional sentiment than with individual sentiment. Indeed we find that FUNDCASH has a correlation of -0.36 with $\mathrm{SENT}^{\mathrm{p}}$ and -0.25 with $\mathrm{SENT}^{\mathrm{a}}$.

Finally, we include monthly data on initial public offering first day returns (IPORET) and the number of offerings (IPON) as reported by Ritter (see http://bear.cba.ufl.edu/ritter). IPO activity is often associated with market tops and is considered a measure of sentiment because of information asymmetries between managers and investors.[17]

3.3. Market variables

In addition to the sentiment measures, we collect data on several market factors. In both samples we use a return on large stocks (R_{BIG}) and the part of small stock returns orthogonal to large stock returns (R_{SOB}). We do this orthogonalization as opposed to small minus big (SMB) to make sure the results are driven by the "small" component, not the "minus big" part. Since the SOB portfolio is not tradeable, we also use the small stock portfolio return R_{SML} in a portion of the analysis. For the weekly sample, we use the Standard and Poor's 500 Index as the portfolio of large stocks and the Russell 2000 Index as the portfolio of small stocks. For the monthly data, the returns are from the Center for

[15] We use the method of Parkinson (1992) to calculate daily volatilities, then take the weekly average to avoid biases associated with differing numbers of weekly trading days.

[16] These data were kindly provided by Greg Kadlec. FUNDCASH is available at the monthly frequency beginning in 1970.

[17] See Ibbotson and Ritter (1995) for a discussion and citations on related literature.

Research in Securities Prices. The large stock portfolio is the largest quintile of NYSE/AMEX/NASDAQ stocks with breakpoints based on the NYSE stocks. The small stock portfolio is CRSP deciles 6–8.

For the weekly sample, all the data are combined as of Friday of each week. Weekly returns are based on Friday prices when possible. In cases where Friday prices are not available due to holidays, etc., the most recent previous price is used instead. For both samples, data on counts (e.g., highs and advances) are averages of daily observations in the period.

4. Methodology and results

The contemporaneous relations between changes in many measures of investor sentiment and market returns are strong. For example, correlations between survey measures and returns reported in Table 2 are consistently positive and significantly different from zero especially at the monthly frequency. Univariate regressions (not tabulated here) with returns as the dependent variable consistently reveal statistically and economically significant coefficients on changes in sentiment measures.[18] However, the problem of analyzing the relation between investor sentiment and market returns is of course not this straightforward—correlation is not causality. Similarly, it is likely that there exists feedback between market returns and sentiment measures further complicating the causal relations. Anecdotal evidence supports this: When the market is on a bull run, as it was in the late 1990s, investors appear to become more bullish. For this reason, we subsequently undertake more sophisticated tests to determine the causal relation between sentiment and returns.

As a starting point, we explore this issue by regressing one period ahead returns on the sentiment proxies. The returns we consider are those on the large stock portfolio and the orthogonalized small stock portfolio. The sentiment survey variables are included in both levels and changes.[19] Since autocorrelation in returns is relatively small and dies out quickly, we only include one lag of each return in the regressions. The analysis is performed for both the monthly and weekly frequencies. The results (available on request) are largely as one would expect in an informationally efficient market. The sentiment variables offer little ability to predict short-horizon returns. The few variables that are significant do not seem to show a reliable pattern across the four regressions. The variable with the most evidence of some ability to predict future returns is the level of institutional sentiment. The statistically negative coefficient suggests that high levels of optimism push current prices up, lowering subsequent returns. However, the results from this analysis are subject to an important caveat: collinearity is a problem. Consequently, isolating exactly which variables are providing the explanatory power is difficult. In a way, the collinearity

[18] For example, at the monthly frequency with R_{BIG} as the dependent variable, the coefficient on changes in $\mathrm{SENT}^{\mathrm{p}}$ has an associated t-statistic of 9.34 (p-value < 0.0001) and an R^2 of 0.176.

[19] As a robustness check, we also examined unexpected sentiment defined as residuals from both univariate ARMA models and the residuals from the VAR equation subsequently discussed in Section 4. Results using these measures are very similar.

is heartening since it suggests strong relations between our sentiment measures. Ideally, we would like to efficiently include all of the information about the true unobservable sentiment in an analysis of the relation between sentiment and market performance. Thus, we turn to identifying composite measures of sentiment.

4.1. Identifying unobserved sentiment measures

By examining correlations and regression results, it is clear that many of the popular sentiment proxies we consider are related to the survey data and to each other. For example, when regressing survey sentiment on the indirect sentiment proxies, many of the proxies are significantly related to the survey data. These results are summarized in Table 3. In many cases, these relations are of a contrarian nature. In the monthly sample, high sentiment is associated with increases in short interest and relatively high specialist short-selling. Odd-lot selling is positively related to sentiment, in conflict with the conventional interpretation. The closed-end fund discount (measured here as a premium to net asset value) is significantly negatively related to survey sentiment. The negative coefficient indicates that when the discount widens investors are optimistic, counter to the argument in Lee et al. (1991). The results are consistent with the claim that IPOs tend to occur during bullish periods. The mutual fund data suggest that during times of high sentiment investors are putting money into mutual funds and the funds are holding relatively little cash. The weekly results also show significant relations between sentiment measures but are somewhat less strong, perhaps because timing issues and collinearity problems are more severe.

Conceptually, it is appealing to extract the common component(s) of the sentiment measures and hope that it represents a cleaner measure of investor sentiment. In order to exploit as much information as possible, we combine the various sentiment measures and use two well-established methods to extract common features of the data: the Kalman filter and principal component analysis. Appendix A provides details on the methodology and some robustness checks.

In brief, we extract our measures of investor sentiment by identifying a single state variable using the Kalman filter and the first two principal components of the selected series. We will refer to these as the filtered sentiment, principal component one, and principal component two. We do this once for the monthly data and repeat the procedure in the weekly sample. We then split the inputs to the estimation procedures for the weekly analysis into institutional and individual groups based on the previous analysis and our subjective priors. Fig. 2 shows the survey data for the institutional series along with the filtered estimate and the first principal component. The three series have the same general pattern, although differences in each arise at high frequencies. We also perform a number of diagnostic tests to assess the quality of these estimates.

We find that the extracted signals are nearly perfectly explained by their inputs (as expected), yet no single input dominates the others. The Kalman filter estimate tends to be highly correlated with the first principal component and less correlated with the second principal component. Correlations between market return series and the filtered estimate and first principal component are consistently positive in both the monthly and

Table 3
Survey sentiment and technical variables

		Monthly	Weekly	
		$SENT^p$	$SENT^p$	$SENT^a$
Trading volume	ADV/DEC	48.5813***	− 0.8145*	0.9775
	ARMS	− 13.9606*	0.8817	6.3919
	HI/LO	− 0.0567	0.0802*	0.3499***
Type of trade	ΔMARGIN	− 0.0279		
	ΔSHORTIR	19.0383***		
	SPECIAL	24.5495*	12.1160**	5.8992
	ODDLOT	6.1902***	− 2.2866***	− 4.9089***
	SHORTSLS		1.1186	− 10.1899
Derivatives	PUT/CALL		− 0.4326	0.6284
	VOL		− 0.5776	3.9741**
	FUT^p		0.0000	− 0.0000
	FUT^a		0.0000	0.0001
	MKTVANE		0.0177	0.0864
Other	CEFD	− 0.2481***	− 0.1096	− 0.2508
	IPORET	− 0.0620		
	IPON	0.0572**		
	FUNDFLOW	3.1082***		
	FUNDCASH	− 0.9419***		
Control variables	R_{BIG}	− 0.8269*	0.4876***	1.4093***
	$R_{BIG,t-1}$	0.2638	0.7444***	0.9913***
	$R_{BIG,t-2}$		0.9747***	− 0.1523
	$R_{BIG,t-3}$		0.4899***	0.4056
	R_{SOB}	− 0.2286	0.1158	− 0.0208
	$R_{SOB,t-1}$	0.2239	− 0.0273	0.5448
	$R_{SOB,t-2}$		0.0742	− 0.0614
	$R_{SOB,t-3}$		0.1289	0.3544
	CONST	− 51.2297***	− 2.5817	− 10.3119
$\bar{R}^2$		0.7105	0.9261	0.5475

Regressions of survey sentiment data on the technical variables. The monthly regression uses 336 observations from January 1970 to December 1997. The weekly regressions use 596 observations from July 1987 through December 1998. Newey and West (1987) standard errors use four and eight lags for the monthly and weekly regressions, respectively.

* Indicate significance at the 10% level.

** Indicate significance at the 5% level.

*** Indicate significance at the 1% level.

weekly samples.[20] The correlation between the filtered estimates from the institutional and individual samplesis only 0.59 despite the fact that they share several common inputs.

[20] The correlation with R_{BIG} is low to moderate for the filtered estimates (0.05–0.41) and moderate for the first principal components (0.33–0.65). The correlation with R_{SOB} is consistently positive for both the filtered estimates and the first principal components (0.18–0.35). The correlations between the return series and the second principal components are positive for the monthly sample but generally negative for the weekly sample. The correlation with R_{BIG} is 0.09 for the monthly estimate and ranges from −0.26 to −0.47 for the weekly estimates. The correlation with R_{SOB} is 0.28 for the monthly sample and ranges from 0.02 to −0.14 for the weekly sample.

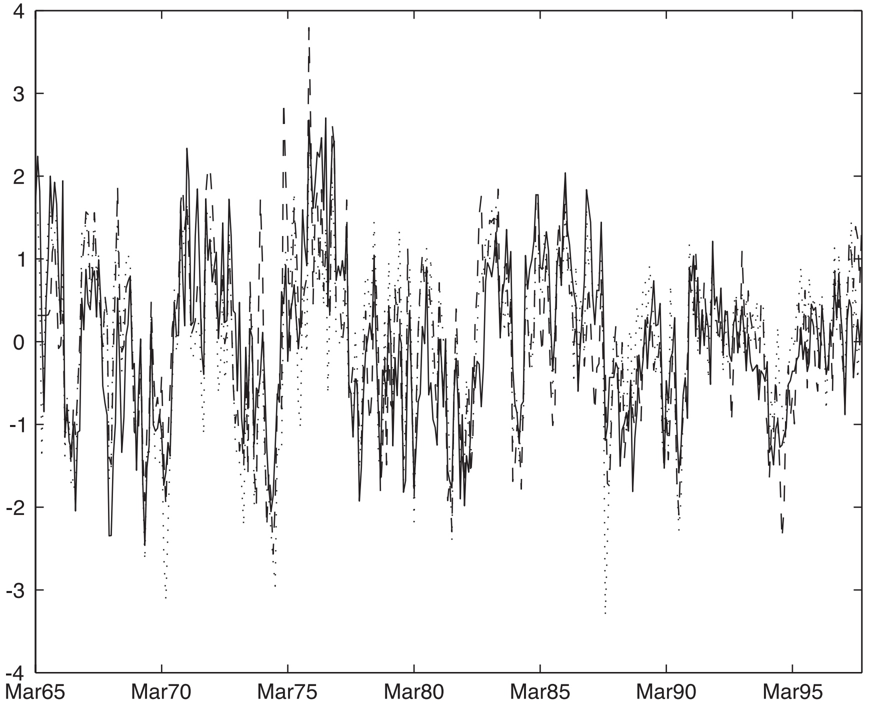

Fig. 2. Plot of the monthly institutional sentiment survey (solid line) along with the sentiment estimates from the Kalman filter (dashed line) and the first principal component (dotted line).

As a robustness check on our method, we can compare the monthly averages of our weekly estimates with the estimates from the monthly analysis. For the filtered estimates and first principal components, these correlations are large and positive (0.64 and 0.78, respectively). This suggests that the methods are capturing similar primary features in the monthly and weekly samples. However, the correlation between second principal components is − 0.39 implying a more cautious interpretation of this series is in order. In sum, it appears we have been able to successfully extract measures of unobserved sentiment (or some other common signal) from the various indicators.

4.2. Sentiment and market returns as a system

Some of our earlier analysis suggests that market returns and sentiment may act as a system. For this reason, we estimate a set of VAR models with the sentiment series and market returns. The goal is to see how sentiment and market returns interact and identify the (statistical) causality between sentiment and the market.

The general model is

$$Y_t = \mu + \sum_{i=1}^{p} \Phi_i Y_{t-i} + \varepsilon_t \tag{1}$$

where Y_t=[Kal_t^{a} Kal_t^{p} $R_{\text{BIG},t}$ $R_{\text{SOB},t}$] for the weekly VAR and Y_t=[$\Delta\text{Kal}_t^{\text{p}}$ $R_{\text{BIG},t}$ $R_{\text{SOB},t}$] for the monthly VAR. Model selection metrics such as AIC and BIC suggest $p=2$ in the monthly data and $p=4$ in the weekly data. Likelihood ratio tests for the model order indicate several additional lags are needed. In the interest of parsimony, we stick to the smaller model orders. The results with larger parameterizations and with models including volume and volatility are very similar with respect to the sentiment results.

We estimate the models using both levels and changes in sentiment measures since it is not easily determined which specification should reveal the primary effects of sentiment. From an econometric viewpoint, differencing the sentiment series will mitigate some of the problems with outdated newsletters entering into the current observation. From a theoretical standpoint, both levels and changes in sentiment may affect market returns. For example, suppose investor sentiment decreases from very bullish to bullish. Because sentiment is still bullish, one might anticipate a positive return, but on the other hand, since sentiment has decreased one might also expect a reduction in the return.

Table 4 reports the results from estimating the monthly sample VAR using sentiment levels. The blocks of rows indicate the contribution of each independent variable at lags one and two. The p-values from an F-test in each column indicate the joint significance of the lagged values of a given independent variable using White (1980) heteroskedastic-consistent standard errors. The first block of rows and the first column are of primary interest here.

The first block of rows shows that the level of our filtered sentiment variable is a powerful predictor of itself. Both the 1- and 2-month lags are positive and significant at the

Table 4
Monthly VAR-returns and filtered sentiment levels

Independent variable	Lag	Dependent variable		
		Kal	R_{BIG}	R_{SOB}
Kal	1	0.5089***	0.2598	− 0.5837*
	2	0.2385***	− 0.1089	− 0.0032
p-value		0.0000***	0.8017	0.0118**
R_{BIG}	1	0.0452***	− 0.0183	0.2079***
	2	− 0.0099	− 0.0683	0.0158
p-value		0.0000***	0.4815	0.0002***
R_{SOB}	1	0.0193*	0.1470**	0.1061*
	2	− 0.0071	− 0.0814	0.0486
p-value		0.1839	0.0879*	0.0657*
Constant		− 0.0364	1.1078***	− 0.2334
Block exogenity		0.0000***	0.0296**	0.0000***
$\bar{R}^2$		0.5981	0.0037	0.0592

Variables are the Kalman filter estimate of sentiment (Kal), returns on large stocks (R_{BIG}), and the return on small stocks orthogonal to large stocks (R_{SOB}). Large stock returns are the first two deciles of CRSP stocks; the small stocks used to create the SOB portfolio are CRSP deciles 6–8. White (1980) standard errors correct for heteroskedasticity. p-value indicates joint significance of all lags of an explanatory variable. Block exogenity reports the p-value of an F-test that the coefficients on all lags of all independent variables (other than own-lags) are jointly zero. The data are 406 monthly observations from March 1965 through December 1998.

* Indicate significance at the 10% level.

** Indicate significance at the 5% level.

*** Indicate significance at the 1% level.

1% level. The next column reveals that sentiment does not predict large stock returns. However, in the last column, the sentiment variable is negatively related to the small stock returns (at the 10% level with a 1-month lag and the 5% level in the joint test). Looking down the first column, the impact of both large and SOB stock returns on sentiment is evident. Large stock returns Granger-cause sentiment at the 1% level and the 1-month lag R_{SOB} return is significant at the 10% level. These results confirm that speculators are swayed by recent market performance. If bullish speculators act on their beliefs by trading, this is consistent with momentum in returns. Lagged levels of sentiment and market returns explain substantial variation in sentiment as indicated by the high $\bar{R}^2$ of 0.60. However, lagged sentiment and market returns explain very little of either large or small stock returns.

Estimating the system using the change in sentiment (Table 5) shows that a change in the level of sentiment is also a powerful predictor of itself. Conversely, there is no evidence that changes in sentiment influence either of the return variables. The $\bar{R}^2$ for each equation reveals that lagged sentiment changes and market returns explain roughly 10% of the variability in $\mathrm{SENT}^{\mathrm{p}}$ but very little of the variability in the returns (though more for R_{SOB} than for R_{BIG}).

The results from estimating the weekly sample VAR with sentiment levels are presented in Table 6. Here, we have included both our estimates for individual ($\mathrm{Kal}^{\mathrm{a}}$) and institutional ($\mathrm{Kal}^{\mathrm{p}}$) filtered sentiment in the system. The results for the sentiment series resemble those reported in Table 4 for the monthly sample. Individual sentiment is strongly positively related to its past levels and positively related to recent large stock

Table 5
Monthly VAR-returns and filtered sentiment changes

Independent variable	Lag	Dependent variable		
		ΔKal	R_{BIG}	$\mathrm{R}_{\mathrm{SOB}}$
ΔKal	1	− 0.3404***	0.0813	− 0.1266
	2	− 0.1399**	− 0.2279	0.0737
p-value		0.0000***	0.8162	0.8706
R_{BIG}	1	0.0315***	− 0.0004	0.1651***
	2	− 0.0212**	− 0.0346	− 0.0462
p-value		0.0005***	0.8622	0.0016***
R_{SOB}	1	0.0192*	0.1526**	0.0998*
	2	− 0.0078	− 0.0633	0.0254
p-value		0.2090	0.0938*	0.1519
Constant		− 0.0113	1.0588***	− 0.1245
Block exogenity		0.0004***	0.0303**	0.0002***
$\bar{R}^2$		0.0847	0.0036	0.0388

Variables are the changes in the Kalman filter estimate of sentiment (ΔKal), returns on large stocks (R_{BIG}), and the return on small stocks orthogonal to large stocks (R_{SOB}). Large stock returns are the first two deciles of CRSP stocks; the small stocks used to create the SOB portfolio are CRSP deciles 6–8. White (1980) standard errors correct for heteroskedasticity. *p*-value indicates joint significance of all lags of an explanatory variable. Block exogenity reports the *p*-value of an *F*-test that the coefficients on all lags of all independent variables (other than own-lags) are jointly zero. The data are 406 monthly observations from March 1965 through December 1998.

* Indicate significance at the 10% level.

** Indicate significance at the 5% level.

*** Indicate significance at the 1% level.

Table 6
Weekly VAR-returns and filtered sentiment levels

Independent variable	Lag	Dependent variable			
		Kal^a	Kal^p	R_{BIG}	R_{SOB}
Kal^a	1	0.3477***	0.0535	0.1843	− 0.0765
	2	0.0700	− 0.0199	0.0556	0.0960
	3	− 0.0106	0.0060	− 0.1687	0.0832
	4	0.1492***	0.0285	− 0.0440	0.0003
p-value		0.0000***	0.5791	0.3727	0.4302
Kal^p	1	0.1282**	0.5195***	− 0.1780	− 0.0659
	2	0.0368	− 0.0080	− 0.2687	− 0.0729
	3	− 0.1245**	0.0831	0.2970*	− 0.0966
	4	0.0645	0.1431***	0.1510	0.1048
p-value		0.0044***	0.0000***	0.0261**	0.2495
R_{BIG}	1	0.1353***	0.1023***	− 0.0613	0.2779***
	2	0.0504**	0.0819***	0.0510	0.0615
	3	0.0495**	0.0495***	0.0170	0.0220
	4	0.0209	0.0016	− 0.0097	0.0240
p-value		0.0000***	0.0000***	0.4620	0.0000***
R_{SOB}	1	0.0118	− 0.0222	0.0702	0.0699
	2	− 0.0273	− 0.0348	0.0526	− 0.0201
	3	0.0187	− 0.0300	0.0163	0.0561
	4	− 0.0020	− 0.0392*	− 0.0487	0.0262
p-value		0.7813	0.0412**	0.6953	0.2911
Constant		− 0.0627*	− 0.0552*	0.2363***	− 0.0881
Block exogenity		0.0000***	0.0000***	0.0070***	0.0000***
$\bar{R}^2$		0.4986	0.5982	0.0039	0.1697

Variables are the Kalman filter estimates of individual and institutional sentiment (Kal^a and Kal^p), returns on large stocks (R_{BIG}), and the return on small stocks orthogonal to large stocks (R_{SOB}). Large stock returns are the first two deciles of CRSP stocks; the small stocks used to create the SOB portfolio are CRSP deciles 6–8. White (1980) standard errors correct for heteroskedasticity. *p*-value indicates joint significance of all lags of an explanatory variable. Block exogenity reports the *p*-value of an *F*-test that the coefficients on all lags of all independent variables (other than own-lags) are jointly zero. The data are 596 weekly observations from July 1987 through December 1998.

* Indicate significance at the 10% level.

** Indicate significance at the 5% level.

*** Indicate significance at the 1% level.

returns. Similarly, strong relations are found for institutional sentiment. Institutional sentiment is a significant predictor of individual sentiment but there is no evidence of the individual sentiment measure influencing institutional sentiment (despite the significant pairwise correlation).

Only limited evidence suggests that sentiment might predict subsequent market returns. The Granger-causality tests fail to reject the null hypothesis of no predictability in returns for all but one variable: Institutional sentiment appears positively related to subsequent large stock returns at about the 3% level. Note that this is counter to conventional wisdom that it is more likely that individual investors would be most easily swayed by sentiment and their effect would be most felt in small stocks. This significance is driven by the third lag of Kal^p. The reliability of this relation is called into question by the absence of an effect from the first two lags.

Estimates from the weekly data using changes in the sentiment measures (Table 7) are consistent with the results in levels. Both (ΔKal^a) and (ΔKal^p) are very significant predictors of their own lagged changes and there is again evidence that institutional sentiment affects individual sentiment but not the converse. Large stock returns are strong positive predictors of changes in sentiment where as there is a some what weaker negative relation between small stock returns and changes in each type of investor sentiment. Overall, these variables are able to explain about 25% of the variation in each of the sentiment measures.

The effects of investor sentiment on subsequent returns are again quite limited. As before, the only significant relation is between institutional sentiment and large stock returns. Specifically, changes in institutional sentiment are negatively related to future large stock returns at about the 1% level. Though the coefficients are negative for each lag,

Table 7
Weekly VAR-returns and filtered sentiment changes

Independent variable	Lag	Dependent variable			
		ΔKal^a	ΔKal^p	R_{BIG}	R_{SOB}
ΔKal^a	1	− 0.5230***	0.0684	0.2059	− 0.0923
	2	− 0.3574***	0.0558	0.2876**	− 0.0136
	3	− 0.2856***	0.0523	0.1292	0.0522
	4	− 0.0788*	0.0695	0.1273	0.0391
p-value		0.0000***	0.2801	0.2174	0.5679
ΔKal^p	1	0.1325**	− 0.4177***	− 0.2149	− 0.0351
	2	0.1453**	− 0.3962***	− 0.5095***	− 0.0880
	3	− 0.0142	− 0.2766***	− 0.2316	− 0.1613
	4	− 0.0061	− 0.1410***	− 0.1672	− 0.0356
p-value		0.0066***	0.0000***	0.0111**	0.4920
R_{BIG}	1	0.1237***	0.0981***	− 0.0650	0.2787***
	2	0.0257	0.0693***	0.0434	0.0613
	3	0.0188	0.0366**	0.0179	0.0216
	4	− 0.0114	− 0.0112	− 0.0034	0.0220
p-value		0.0000***	0.0000***	0.4795	0.0000***
R_{SOB}	1	− 0.0169	− 0.0312	0.0685	0.0699
	2	− 0.0575**	− 0.0372	0.0559	− 0.0161
	3	− 0.0066	− 0.0300	0.0228	0.0597
	4	− 0.0353	− 0.0407*	− 0.0387	0.0284
p-value		0.0836*	0.0169**	0.7082	0.2427
Constant		− 0.0381	− 0.0417	0.2372**	− 0.0858
Block exogenity		0.0000***	0.0000***	0.0027***	0.0000***
$\bar{R}^2$		0.2511	0.2361	0.0077	0.1672

Variables are the changes in the Kalman filter estimates of individual and institutional sentiment (ΔKal^a and ΔKal^p), returns on large stocks (R_{BIG}), and the return on small stocks orthogonal to large stocks (R_{SOB}). Large stock returns are the first two deciles of CRSP stocks; the small stocks used to create the SOB portfolio are CRSP deciles 6–8. White (1980) standard errors correct for heteroskedasticity. *p*-value indicates joint significance of all lags of an explanatory variable. Block exogenity reports the *p*-value of an *F*-test that the coefficients on all lags of all independent variables (other than own-lags) are jointly zero. The data are 596 weekly observations from July 1987 through December 1998.

* Indicate significance at the 10% level.

** Indicate significance at the 5% level.

*** Indicate significance at the 1% level.

the result is driven by the coefficient on the second lag. Even with this significant relation, sentiment has almost no explanatory power for returns as indicated by the $\bar{R}^2$ of 0.008. None of the coefficients (or joint tests) reveal a significant relation between changes in sentiment and small stock returns.

Replacing the filtered sentiment with the survey and principal component estimates (results not reported) yields very similar results. The sentiment surveys do not show any statistically significant predictive power with the exception of the change in SENT^{p} being significantly negatively related to R_{SOB} at the 5% level in the monthly sample.

In summary, this section reveals that market returns are a strong predictor of subsequent levels and changes of both individual and institutional sentiment. In contrast, there is only weak evidence that even our composite sentiment measures are useful in predicting future stock returns over short horizons.

4.3. Abnormal performance

It is possible that even the limited predictability of our composite sentiment measures could be used as a profitable trading strategy. As a consequence, we perform a final test using sentiment as an indicator for trading. Intuitively, the sentiment measures can be used to determine the optimal level of investment in the market. This strategy is much like any other managed portfolio such as a mutual fund, and the fund's performance can be tested. Of course, it is necessary to determine whether a fund does better or worse than expected. Parametric asset pricing models are often used in this exercise (e.g., Jensen's alpha). However, we steer away from parametric models because of the joint hypothesis problem.

To test performance non-parametrically, we use the techniques of Chen and Knez (1996). The basic idea is to extract the stochastic discount factor from a set of basis assets, use this to calculate the "fair" price of the asset, then see if this price is significantly different from the market price (normalized to one). Let z_t be the sentiment signal, standardized to have the same standard deviation as R_{BIG} and a mean of 1. At each period, z_t dollars are invested in R_{BIG} so the payoff will be $R_{z,t} \equiv z_t R_{\text{BIG},t}$. The normalization of the trading signal means z_t ranges between 0.86 and 1.15, with 80% of the values between 0.93 and 1.07 so extreme positions are not considered. The basis assets are the 25 Fama and French (1993) portfolios, which are formed on the basis of size and book/market. Appendix B provides additional details on this analysis.

The results from the GMM estimation (reported in the text only) using R_{BIG} do not suggest any abnormal performance to the sentiment-based trading strategy. The χ^2 statistic is 0.86 (p-value of 0.35), and the t-statistic for the mispricing moment condition is also insignificant. Repeating the analysis with the return on small stocks the conclusions are similar. The p-value for the test of over-identifying restrictions is 0.25 and the t-statistics for the abnormal performance moment condition remains insignificant. We also perform the analysis using the Kalman filter and principal components estimates in place of the survey data. Overall, the results are very similar. The only signal that yields a significant χ^2 statistic (at the 10% level) is using the first principal component as a trading strategy for small stocks. Given that we performed eight of these experiments (two types of stocks for four signals), one test significant at the 10% level cannot be regarded as strong evidence.

In summary, the non-parametric performance measures provide no evidence of abnormal returns to a sentiment-based trading strategy at short (monthly) horizons. This is consistent with Graham and Harvey (1996), who find no evidence that the asset allocation recommendations of a collection of investment newsletters is useful in forecasting future market returns.

5. Conclusions

In summary, we have demonstrated that surveys measuring investor sentiment are related to other popular measures of investor sentiment and recent stock market returns. By using signal extraction techniques, we have been able to isolate common features of these indicators for a long monthly time period (34 years) and a shorter weekly time period (11 years). In the weekly sample, we have generated two separate measures that we believe represent institutional and individual investor sentiment.

For all of the aggregate sentiment measures, we find strong evidence of co-movement with the market but little evidence of short-run predictability in returns (using a variety of methods). Using the limited predictability of sentiment as a trading strategy does not appear to be profitable. Finally, our research does not suggest that sentiment is limited to individual investors. To the contrary, it appears that the strongest relations exist between our measures on institutional sentiment and large stocks. This has implications for existing research, which typically assumes "noise" traders are individuals who affect small stocks.

Acknowledgements

Funding from the Cato Center for Applied Business Research is gratefully acknowledged. We thank Kent Daniel, Greg Kadlec, Steve Slezak, Kent Womack, seminar participants at 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 Greg Kadlec for providing data. This paper is based on a working paper previously titled "Sentiment and the Stock Market".

Appendix A. Unobserved components

A.1. Methodology

We posit a single unobservable state variable, investor sentiment (S), which follows an AR(1) process

$$S_{t+1} = \mu + \gamma S_t + w_t \tag{A.1}$$

The N observable proxies for investor sentiment are collected into a vector $y_t=[S_t^1 \ldots S_t^N]'$.

These observables are modeled as a linear function of the state variable

$$y_t = \alpha + \Theta S_t + v_t \tag{A.2}$$

The scalar error from the state equation (Eq. (A.1)) has variance σ^2 and the error covariance matrix from the measurement equation (Eq. (A.2)) is $\Sigma_v = \sigma^2 I$. If these errors are normally distributed and $E[w_t v_t]=0$, the Kalman filter provides optimal linear forecasts of the state variable.

Since the parameters γ and Θ are unknown, we estimate them by maximum likelihood. With $\hat{\gamma}$, $\hat{\Theta}$, and $\hat{\sigma}^2$ in hand the Kalman filter can produce estimates of the state variable through time. This is done by iterating the equations

$$\hat{S}_{t+1|t} = \hat{\gamma}\hat{S}_{t|t-1} + K(y_t - \hat{\Theta}S_t) \tag{A.3}$$

and

$$\hat{S}_{t|t} = \hat{S}_{t|t-1} + \hat{\gamma}^{-1}K(y_t - \hat{\Theta}\hat{S}_{t|t-1}) \tag{A.4}$$

where $K = \hat{\gamma}P_{t|t-1}\hat{\Theta}'(\hat{\Theta}P_{t|t-1}\hat{\Theta}' + \hat{\Sigma}_v)^{-1}$ is known as the gain matrix and $P_{t|t-1}$ denotes the mean squared error of the one step ahead forecast of the state variable. The notation $\hat{S}_{\tau|\tau'}$ indicates the forecast of S_τ made at time τ'. Refer to Hamilton (1994) for a more detailed discussion of the Kalman filter.

As an alternative to the Kalman filter, we also employ principal components analysis. This method allows us to identify a specific number of time series that explain the most variation in a set of data series. By construction, the components will be orthogonal to each other. Since we may examine multiple components for a given set of data, this method may provide insight into the question of whether multiple sentiment measures are present in our data. Additional details are provided in references such as Morrison (1967).

A.2. Diagnostics

We first apply these methods to the monthly time series. As discussed above, it is possible that sentiment indicators that are not significant explanatory variables for the sentiment surveys still contain valuable information. Consequently, we include all of the indicators presented in the last two columns of Table 3 as well as the II survey data.[21] To prevent having to weight each series for the Kalman filter and to facilitate subsequent analysis, we normalize each of the series to mean zero and unit standard deviation.

We refer to the output from the Kalman filter as filtered sentiment. To test the properties of this series, we regress the filtered sentiment on the input sentiment indicators. All of the

[21] Excluding any single indicator has little effect on the resulting time series.

indicators are significant explanatory variables at the 1% level. This combined with an R^2 of 0.98 reassures us that the method is in fact capturing common features of all of the inputs. Furthermore, the filtered sentiment, while closely correlated with nearly all of the inputs, is not dominated by any one. The highest pairwise correlation is with HI/LO (0.79) and the second highest is with SENT^{p} (0.61). The only correlation with a magnitude less than 0.13 is with SPECIAL (0.01). All correlations are positive with the exception of the ARM index ($-$ 0.40).

For the principal component analysis, we examine only the first two components in detail.[22] SENT^{p}, HI/LO, MARGIN, FUNDFLOW, ARMS, and IPON have the largest factor loadings for the first component. SPECIAL and CEFD have the largest loadings for the second component. The only indicator with a low factor loading in both components is SHORTIR.

Next, we estimate market sentiment series using the weekly data. First, we construct overall sentiment measures analogous to the monthly measures. As inputs, we use both sentiment surveys and all of the weekly indicators in Table 3. The results are very similar to those reported above for the monthly data. Specifically, for filtered sentiment, all variables are significantly related to the resulting sentiment measure. In the principal component analysis, only the change in CEFD and FUT^{p} do not have a factor loading greater than 0.2 for either component.

The similarity of the weekly and monthly results provides an encouraging indicator suggesting the methods are fairly robust to different inputs, sample periods, and data frequencies. This result is particularly heartening given the subjective task of separating indicators into individual and institutional groups. For this task, we rely on the analysis reported in Table 3 and our prior beliefs discussed in Section 3.1. Specifically, we select SENT^{p}, HI/LO, ADV/DEC, ARMS, SPECIAL, VOL, FUT^{P}, and SHORTSLS as the inputs for institutional sentiment. For individual sentiment indicators, we select SENT^{a}, HI/LO, ADV/DEC, ARMS, MKTVANE, CEFD, change in CEFD, ODDLOT, and FUT^{a}.

The resulting institutional filtered sentiment series shows similar properties to the aggregate filtered sentiment series. All of the included indicators are significant explanatory variables for the filtered series. Likewise, in the principal component analysis, all of the indicators have large (>0.2) factor loadings for at least one of the first two components with the exception of put/call trading volume. The institutional filtered sentiment and the first two principal components appear to capture many of the same effects (correlations of 0.65 and 0.63, respectively).

Repeating the analysis for individual sentiment indicators yields, once again, similar results. All inputs are significantly related to the resulting filtered sentiment series. Similar results hold when considering the magnitude of the factor loadings for the first two

[22] In general, the absolute and relative magnitudes of the eigenvalues indicate that there are two important principal components. Occasionally, diagnostic tests (at both the monthly and weekly frequencies) indicate that more components may be required to properly capture relevant common features in the data. In the subsequent analysis, adding more components did not change the qualitative results so they are omitted from the analysis to facilitate exposition.

principal components. It is still the case that filtered sentiment and the first principal component capture similar information (correlation of 0.76). For individual sentiment, the filtered series is only slightly positively correlated with the second principal component (correlation of 0.13).

Most importantly, it appears that the institutional and individual sentiment series resulting from these procedures are capturing at least some different intertemporal data features. The correlation between institutional and individual filtered sentiment is only 0.59 despite the fact that they share some common inputs. Likewise, the correlation between the two first principal components is 0.61 and the two second principal components is 0.49. Market returns are not an important explanatory variables for the new institutional and individual sentiment measures. The correlations between each filtered measure and lagged large stock returns averages only 0.11. The correlations with SOB returns are slightly higher, averaging about 0.25. For the large stocks, there are larger positive correlations with the first principal component (averaging 0.49) and negative correlations with the second principal components (averaging -0.40).For R_{SOB}, the correlations are similar but closer to zero.

Appendix B. Abnormal performance estimation

Following Hansen and Jagannathan (1991), a pricing kernel m_t that prices all the basis assets X_t must satisfy

$$\iota_{[25]} = E[m_t X_t] = E[X_t X_t^{'} \alpha]. \tag{B.1}$$

The last equality follows from the fact that the pricing kernel can be expressed as a portfolio of the basis assets, whose weights are given by α.

We then want to see whether m_t also prices $R_{z,t}$

$$1 = E[m_t R_{z,t}] = E[R_{z,t} X_t^{'} \alpha]. \tag{B.2}$$

Let $\eta_{n,t} = R_{n,t} X_t' \alpha - 1$ be the pricing error at time t for each asset n. Stacking these together into η_t, the 25 elements of α are parameters to estimate, while Eqs. (B.1) and (B.2) represent 26 restrictions. The system is estimated using the generalized method of moments from Hansen (1982), which minimizes the quadratic

$$J_T = \left[\frac{1}{T}\sum_{t=1}^{T}\eta_t\right]^{'} W_T \left[\frac{1}{T}\sum_{t=1}^{T}\eta_t\right]. \tag{B.3}$$

Under the null hypothesis of no abnormal performance, $TJ_T \sim \chi^2(1)$. Of course, nothing is for free, so the downside of this approach is in the choice of basis assets. If the basis assets do not span the payoff space, then the abnormal performance may in fact just represent spanning enhancements. Given the spanning assumption, finding abnormal performance may be interpreted as an arbitrage opportunity. As a result, it is necessary to rely on frictions and measurement errors for a sensible application of this approach.

References

Avery, C., Chevalier, J., 1999. Identifying investor sentiment from price paths: the case of football betting. Journal of Business 72, 493–521.

Bakshi, G., Chen, Z., 2001. Stock valuation in dynamic economies. Working Paper.

Barber, B.M., 1994. Noise trading and prime and score premiums. Journal of Empirical Finance 1, 251–278.

Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. Journal of Financial Economics 49, 307–343.

Black, F., 1986. Noise. Journal of Finance 41, 529–543.

Brown, G.W., 1999. Volatility, sentiment, and noise traders. Financial Analysts Journal 55, 82–90.

Brown, G.W., Cliff, M.T., 2004. Investor sentiment and asset valuation. Journal of Business (forthcoming).

Campbell, J.Y., Kyle, A.S., 1993. Smart money, noise trading, and stock price behaviour. Review of Economic Studies 60, 1–34.

Chen, Z., Knez, P.J., 1996. Portfolio performance measurement: theory and applications. Review of Financial Studies 9, 511–555.

Chen, N., Kan, R., Miller, M., 1993. Are the discounts on closed-end funds a sentiment index? Journal of Finance 48, 795–800.

Clarke, R.G., Statman, M., 1998. Bullish or bearish? Financial Analysts Journal, 63–72 (May/June).

Daniel, K., Hirshleifer, D., Subrahmanyam, A., 1998. Investor psychology and security market under- and overreactions. Journal of Finance 53, 1839–1886.

DeLong, J.B., Shleifer, A., Summers, L.H., Waldmann, R.J., 1989. The size and incidence of the losses from noise trading. Journal of Finance 44, 681–696.

DeLong, J.B., Shleifer, A., Summers, L.H., Waldmann, R.J., 1990a. Noise trader risk in financial markets. Journal of Political Economy 98, 703–738.

DeLong, J.B., Shleifer, A., Summers, L.H., Waldmann, R.J., 1990b. The survival of noise traders in financial markets. NBER Working Paper #2715.

Donaldson, R.G., Kim, H.Y., 1993. Price barriers in the Dow Jones industrial average. Journal of Financial and Quantitative Analysis 28, 313–330.

Elton, E.J., Gruber, M.J., Busse, J.A., 1998. Do investors care about sentiment? Journal of Business 71, 477–500.

Fama, E., French, K., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.

Fisher, K.L., Statman, M., 2000. Investor sentiment and stock returns. Financial Analysts, 16–23 (March/April).

Fosback, N.G., 1993. Stock Market Logic: A Sophisticated Approach to Profits on Wall Street. Dearborn Financial Publishing, Chicago, IL.

Graham, J., Harvey, C.R., 1996. Market timing and volatility implied in investment newsletters' asset allocation recommendations. Journal of Financial Economics 42, 397–421.

Grossman, S., Stiglitz, J.E., 1980. On the impossibility of informationally efficient markets. American Economic Review 70, 393–408.

Hamilton, J.D., 1994. Time Series Analysis. Princeton Univ. Press, Princeton, NJ.

Hansen, L.P., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, 1029–1054.

Hansen, L.P., Jagannathan, R., 1991. Implications of securities market data for models of dynamic economies. Journal of Political Economy 99, 225–262.

Hirshleifer, D., 2001. Investor psychology and asset pricing. Journal of Finance 56, 1533–1598.

Hong, H., Stein, J.C., 1999. A unified theory of underreaction, momentum trading, and overreaction in asset markets. Journal of Finance 54, 2143–2184.

Ibottson, R., Ritter, J., 1995. Initial public offerings. North-Holland Handbooks of Operations Research and Management Science: Finance. North-Holland, Amsterdam, pp. 993–1016.

Lee, C., Shleifer, A., Thaler, R., 1991. Investor sentiment and the closed-end fund puzzle. Journal of Finance 46, 75–109.
Lee, C.M.C., Myers, J., Swaminathan, B., 1999. What is the intrinsic value of the Dow? Journal of Finance 54, 1693–1741.
Lintner, J., 1969. The aggregation of investor's diverse judgements and preferences in purely competitive security markets. Journal of Financial and Quantitative Analysis 4, 347–400.
Morrison, D.F., 1967. Multivariate Statistical Methods, 2nd ed. McGraw Hill, New York.
Neal, R., Wheatley, S.M., 1998. Do measures of investor sentiment predict returns? Journal of Financial and Quantitative Analysis 33, 523–547.
Newey, W., West, K., 1987. A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.
Otoo, M.W., 1999. Consumer sentiment and the stock market. Working Paper, Federal Reserve Board of Governors.
Parkinson, M., 1992. The extreme value method for estimating the variance of the rate of return. Journal of Business 53, 61–65.
Seyhun, N.H., 1986. Insiders' profits, costs of trading, and market efficiency. Journal of Financial Economics 16, 189–212.
Siegel, J.J., 1992. Equity risk premia, corporate profit forecasts, and investor sentiment around the stock market crash of October 1987. Journal of Business 65, 557–570.
Simon, D.P., Wiggins III, R.A., 1999. Stock returns and sentiment indicators. Working Paper, Bentley College.
Smith, A., 1776. An Inquiry into the Nature and Causes of the Wealth of Nations. W. Strahan and T. Cadell Publishers, London, U.K.
Solt, M.E., Statman, M., 1988. How useful is the sentiment index. Financial Analysts Journal, 45–55 (September/October).
Swaminathan, B., 1996. Time-varying expected small firm returns and closed-end fund discounts. Review of Financial Studies 9, 845–887.
White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48, 817–838.