RFS Advance Access published November 17, 2014

Investor Sentiment Aligned: A Powerful Predictor of Stock Returns

Dashan Huang
Singapore Management University

Fuwei Jiang
Central University of Finance and Economics

Jun Tu
Singapore Management University

Guofu Zhou
Washington University in St. Louis and CAFR

We propose a new investor sentiment index that is aligned with the purpose of predicting the aggregate stock market. By eliminating a common noise component in sentiment proxies, the new index has much greater predictive power than existing sentiment indices have both in and out of sample, and the predictability becomes both statistically and economically significant. In addition, it outperforms well-recognized macroeconomic variables and can also predict cross-sectional stock returns sorted by industry, size, value, and momentum. The driving force of the predictive power appears to stem from investors' biased beliefs about future cash flows. (*JEL* C53, G11, G12, G17)

At least as early as Keynes (1936), researchers have analyzed whether investor sentiment can affect asset prices as a result of the well-known psychological fact that people with high (low) sentiment tend to make overly optimistic (pessimistic) judgments and choices. Empirically, a major challenge for testing

We are extremely grateful to David Hirshleifer (the editor) and two anonymous referees for extensive comments that significantly improved the paper. We are grateful to Malcolm P. Baker, Zhanhui Chen, Zhi Da, Bing Han, Larry Harris, Campbell R. Harvey, Robert Hodrick, Allan Timmermann, Xiaolu Wang, Jianfeng Yu, Yu Yuan, Xiaoyan Zhang, and seminar participants at the Central University of Finance and Economics, China Finance Review International Conference 2014, Q-Group 2014 Spring Seminar, ShanghaiTech University, Singapore Management University, Singapore Scholars Symposium 2013, Sun Yat-Sen University, University of Technology (in Sydney), University of Melbourne, Washington University in St. Louis, and Wuhan University for very helpful comments. We also thank Kewei Hou and Seth Pruitt for kindly providing us their data. We are grateful to Andrew J. Patton for making the monotonic test codes available. Dashan Huang acknowledges that the study was funded through a research grant (Grant No. C207/MSS13B003) from Singapore Management University. Jun Tu acknowledges that the study was funded through a research grant (Grant No. C207/MSS12B006) from Singapore Management University and a research grant from Sim Kee Boon Institute for Financial Economics.
Send correspondence to Guofu Zhou, Olin School of Business, Washington University in St. Louis, St. Louis, MO 63130; telephone: 314-935-6384. E-mail: zhou@wustl.edu.

doi:10.1093/rfs/hhu080

the importance of investor sentiment is that it is not directly observable. In their influential studies, Baker and Wurgler (2006, 2007) construct a novel investor sentiment index (BW index, hereafter) that aggregates the information from six proxies, and find that high investor sentiment predicts strongly low returns in the cross-section, such as stocks that are speculative and hard to arbitrage. Stambaugh, Yu, and Yuan (2012) show that investor sentiment is a significant negative predictor for the short legs of long-short investment strategies. Baker, Wurgler, and Yuan (2012) provide further international evidence for the forecasting power of investor sentiment.[1] However, whether investor sentiment can predict the aggregate stock market at the usual monthly frequency is still an open question, because existing studies, such as Baker and Wurgler (2007) and Baker, Wurgler, and Yuan (2012), do not provide strong statistical evidence, whereas Brown and Cliff (2005) find significant indications only at one-year or longer horizons.

In this paper, we exploit the information of Baker and Wurgler's (2006, 2007) six sentiment proxies in a more efficient manner, to obtain a new index for the purpose of explaining the expected return on the aggregate stock market.[2] In their pioneering study, Baker and Wurgler use the first principal component (PC) of the proxies as the measure of investor sentiment. Econometrically, the first PC is the best combination of the six proxies that maximally represents the total variations of the six proxies. Because all of the proxies may have approximation errors to the true but unobservable investor sentiment, and these errors are parts of their variations, the first PC can potentially contain a substantial amount of common approximation errors that are not relevant for forecasting returns. Our idea is to align the investor sentiment measure with the purpose of explaining the returns by extracting the most relevant common component from the proxies. In other words, economically, we separate out information in the proxies that is relevant to the expected stock returns from the error or noise. Statistically, the partial least squares (PLS) method pioneered by Wold (1966, 1975) and extended by Kelly and Pruitt (2013, 2014) does exactly this job. We call the new index extracted this way the *aligned* investor sentiment index, which does incorporate efficiently all of the relevant forecasting information from the proxies, as shown by forecast encompassing tests in our applications.

Empirically, we find that the aligned sentiment index can predict the aggregate stock market remarkably well. Its monthly in- and out-of-sample R^2s in the ordinary least squares (OLS) predictive regressions are 1.70% and 1.23%, respectively, much larger than 0.30% and 0.15%, the counterparts of the BW index. Since a monthly out-of-sample R^2 of 0.5% signals substantial economic value (Campbell and Thompson 2008), our aligned investor sentiment index is

1 There are a number of other applications and related studies. The latest number of Google citations of Baker and Wurgler (2006) is over 1,470.

2 The same method may be applied to explain the expected return on any other asset.

not only statistically significant, but also economically significant in providing sizable utility gains or certainty equivalent returns for a mean-variance investor.

Our finding of strong market predictability of investor sentiment is a complement, in a unique way, to early studies by Baker and Wurgler (2006, 2007) and many others who find that investor sentiment plays an important role in explaining the cross-section of stock returns. Because forecasting and understanding how the market risk premium varies over time is one of the central issues in financial research that has implications in both corporate finance and asset pricing (see, e.g., Spiegel 2008; Cochrane 2011), our study suggests that investor sentiment is related to many central problems in finance beyond its effect on certain segments of the market. De Long et al. (1990) provide theoretical explanations why sentiment can cause asset price to deviate from its fundamental value in the presence of limits of arbitrage, even when informed traders recognize the opportunity. Nevertheless, almost all such theories deal with one risky asset in the analysis, that is, they effectively study the role of investor sentiment on the aggregate market. Hence, the empirical results of our paper provide strong empirical evidence supporting those theoretical models on investor sentiment.

It is of interest to compare how well the aligned investor sentiment index performs relative to alternative economic predictors. Of the well-known macroeconomic predictors, we consider all of the 14 variables used by Goyal and Welch (2008), such as the short-term interest rate (Fama and Schwert 1977; Breen, Glosten, and Jagannathan 1989; Ang and Bekaert 2007), dividend yield (DY) (Fama and French 1988; Ang and Bekaert 2007), earnings-price ratio (Campbell and Shiller 1988), term spreads (Campbell 1987; Fama and French 1988), book-to-market (BM) ratio (Kothari and Shanken 1997; Pontiff and Schall 1998), stock volatility (French, Schwert, and Stambaugh 1987; Guo 2006), inflation (Fama and Schwert 1977), and corporate issuing activity (Baker and Wurgler 2000). In addition, we consider the consumption-wealth ratio (Lettau and Ludvigson 2001), consumption surplus ratio (Campbell and Cochrane 1999), output gap (Cooper and Priestley 2009), and a new powerful predictor developed by Kelly and Pruitt (2013) based on 100 BM ratios. The in-sample R^2s of these 18 individual predictors vary from 0.01% to 2.07% (only 4 of them exceed 1%). Apart from the Kelly and Pruitt (KP) predictor, all others have R^2s below 1.70% of the aligned investor sentiment. When each of these economic predictors is used as a control, the aligned investor sentiment is still significant. Out-of-sample, all the economic predictors have negative R^2s except the KP predictor. In contrast, the aligned investor sentiment, with an R^2 of 1.23%, is both statistically and economically significant and performs the best.

Cross-sectionally, we compare how the aligned investor sentiment index performs relative to the BW index. When stocks are sorted by industry, the BW index has an impressive in-sample R^2 of 1.10% in explaining the time-varying returns on the more speculative and hard-to-value technology firms, but the aligned investor sentiment index raises it to 1.92%. When stocks are sorted by size, value, and momentum, the aligned investor sentiment index always increases the predictive power, and more than doubles the R^2s on average. Hence, the aligned investor sentiment index is useful cross-sectionally as well.

We also explore the economic driving force of the predictive power of the aligned investor sentiment. We ask whether the predictability comes from time variations in cash flows or discount rates. We find that the aligned investor sentiment index negatively and significantly forecasts future aggregate dividend growth (a standard cash flow proxy), but does not forecast future dividend-price ratio (a proxy of discount rate), supporting that the cash-flow channel is the source for predictability. This result is robust for alternative aggregate cash-flow proxies, such as aggregate earning growth and real GDP growth. In addition, the ability of investor sentiment to forecast the cross-section of stock returns is strongly correlated with its ability to forecast the cross-section of future cash flows as well. Hence, our findings suggest that low aggregate stock market return following high investor sentiment seems to represent investors' overly optimistic beliefs about future cash flows that cannot be justified by subsequent economic fundamentals (Baker and Wurgler 2007). Moreover, we also examine the relation of the aligned investor sentiment to some alternative behavior predictors. Even though the aggregate accruals predictor of Hirshleifer, Hou, and Teoh (2009) is the best among other behavioral predictors and has good performances from 1 to 12 months, we find that aligned investor sentiment outperforms it at the monthly frequency. However, the aligned investor sentiment index and aggregate accruals are complementary and their performance difference diminishes as the horizon increases.

1. Econometric Methodology

In this section, we provide first the econometric method for constructing our aligned sentiment index, following Wold (1966, 1975) and, especially, Kelly and Pruitt (2013, 2014). Then, we analytically compare it with the BW index to understand their difference.

1.1 Aligned index S^{PLS}

We assume that the one-period ahead expected excess stock return explained by investor sentiment follows the standard linear relation,

$$E_t(R_{t+1}) = \alpha + \beta S_t, \tag{1}$$

where S_t is the true but unobservable investor sentiment that matters for forecasting asset returns. The realized stock return is then equal to its conditional expectation plus an unpredictable shock,

$$\begin{aligned} R_{t+1} &= E_t(R_{t+1}) + \varepsilon_{t+1} \\ &= \alpha + \beta S_t + \varepsilon_{t+1}, \end{aligned} \tag{2}$$

where ε_{t+1} is unforecastable and unrelated to S_t.

Let $x_t = (x_{1,t}, ..., x_{N,t})'$ denote an $N \times 1$ vector of individual investor sentiment proxies at period t $(t = 1, ..., T)$. In Baker and Wurgler (2006, 2007), x_t is the close-end fund discount rate, share turnover, number of initial public offerings (IPOs), first-day returns of IPOs, dividend premium, and the equity share in new issues. We assume that $x_{i,t}$ $(i = 1, ..., N)$ has a factor structure,

$$x_{i,t} = \eta_{i,0} + \eta_{i,1} S_t + \eta_{i,2} E_t + e_{i,t}, \qquad i = 1, ..., N, \tag{3}$$

where S_t is the investor sentiment that matters for forecasting asset returns, $\eta_{i,1}$ is the factor loading that summarizes the sensitivity of sentiment proxy $x_{i,t}$ to movements in S_t, E_t is the common approximation error component of all the proxies that is irrelevant to returns, and $e_{i,t}$ is the idiosyncratic noise associated with measure i only. The key idea here is to impose the above-mentioned factor structure on the proxies to efficiently estimate S_t, the collective contribution to the true yet unobservable investor sentiment, and at the same time, to eliminate E_t, their common approximation error, and $e_{i,t}$ from the estimation process.

In Baker and Wurgler (2006, 2007), investor sentiment is estimated as the first principle component (PC) of the cross-section of $x_{i,t}$s. By its econometric design, the PC is a linear combination of $x_{i,t}$s that explains the largest fraction of the total variations in $x_{i,t}$s, and hence is unable to separate S_t from E_t. In fact, the larger the variance of E_t is, the more important the role it will play in the PC approach (see the next subsection for some analytical insights). Then, it is possible that the PC may fail to generate significant forecasts for future stock returns, even when stock returns are indeed strongly predictable by the true investor sentiment S_t. This failure indicates the need for an improved econometric method that *aligns* investor sentiment estimation toward forecasting future stock returns.

To overcome this econometric difficulty, following Wold (1966, 1975) and, especially, Kelly and Pruitt (2013, 2014), we apply the partial least squares (PLS) approach to extract S_t effectively and filter out the irrelevant component E_t, whereas the PC method cannot be guaranteed to do so. The key idea is that PLS extracts the investor sentiment, S_t, from the cross-section according to its covariance with future stock returns and chooses a linear combination of sentiment proxies that is optimal for forecasting. In doing so, PLS can be implemented by the following two steps of OLS regressions. In the first-step, we run N time-series regressions. That is, for each individual investor sentiment proxy x_i, we run a time-series regression of $x_{i,t-1}$ on a constant and realized stock return R_t,

$$x_{i,t-1} = \pi_{i,0} + \pi_i R_t + u_{i,t-1}, \qquad t = 1, ..., T. \tag{4}$$

The loading π_i captures the sensitivity of each sentiment proxy $x_{i,t-1}$ to investor sentiment S_{t-1} instrumented by future stock return R_t. Because the expected component of R_t is driven by S_{t-1}, sentiment proxies are related to the expected stock returns and are uncorrelated with the unpredictable return shocks, as shown in Equations (2) and (3). Therefore, the coefficient π_i in the first-stage time-series regression (4) approximately describes how each sentiment proxy depends on the true investor sentiment.

In the second-step, we run T cross-sectional regressions. More specifically, for each time period t, we run a cross-sectional regression of $x_{i,t}$ on the corresponding loading $\hat{\pi}_i$ estimated in the time-series regression (4),

$$x_{i,t} = c_t + S_t^{\mathrm{PLS}} \hat{\pi}_i + v_{i,t}, \qquad i = 1, \cdots, N, \tag{5}$$

where S_t^{PLS}, the regression slope in Equation (5), is the estimated investor sentiment (the aligned sentiment index hereafter). That is, in Equation (5), the first-stage loadings become the independent variables, and the aligned investor sentiment S_t^{PLS} is the regression slope to be estimated.

Intuitively, PLS exploits the factor nature of the joint system, Equations (2) and (3), to infer the relevant aligned sentiment factor S_t^{PLS}. If the true factor loading π_i was known, we could consistently estimate S_t^{PLS} by simply running cross-sectional regressions of $x_{i,t}$ on π_i period-by-period. Because π_i is unknown, however, the first-stage regression slopes provide a preliminary estimation of how $x_{i,t}$ depends on S_t^{PLS}. In other words, PLS uses time $t+1$ stock returns to discipline the dimension reduction to extract S_t relevant for forecasting and discards common and idiosyncratic components such as E_t and $e_{i,t}$ that are irrelevant for forecasting.

Mathematically, when we use full-sample information in the first-step time-series regressions, the $T \times 1$ vector of aligned investor sentiment index, $S^{\mathrm{PLS}} = (S_1^{\mathrm{PLS}}, \ldots, S_T^{\mathrm{PLS}})'$, can be expressed as a one-step linear combination of $x_{i,t}$s,

$$S^{\mathrm{PLS}} = X J_N X' J_T R (R' J_T X J_N X' J_T R)^{-1} R' J_T R, \tag{6}$$

where X denotes the $T \times N$ matrix of individual investor sentiment measures, $X = (x_1', \ldots, x_T')'$, and R denotes the $T \times 1$ vector of excess stock returns as $R = (R_2, \ldots, R_{T+1})'$. The matrices J_T and J_N, $J_T = I_T - \frac{1}{T}\iota_T \iota_T'$ and $J_N = I_N - \frac{1}{N}\iota_N \iota_N'$, enter the formula because each regression is run with a constant. I_T is a T-dimensional identity matrix and ι_T is a T-vector of ones. The weight on each individual measure $x_{i,t}$ in S_t^{PLS} is based on its covariance with the excess stock return to capture the intertemporal relationship between the aligned investor sentiment and the expected excess stock return.

1.2 Comparison of S^{PLS} with S^{BW}

To obtain analytical insights on the difference between S^{PLS} and S^{BW}, we consider a simple case of Equation (3), in which there are only two individual sentiment proxies, x_1 and x_2, that have the following factor structure:

$$x_1 = S + E + e_1, \tag{7}$$

$$x_2 = \eta_1 S + \eta_2 E + e_2, \tag{8}$$

where S is the true but unobservable investor sentiment, E is the common noise, and e_i $(i = 1, 2)$ are the idiosyncratic noises. η_1 and η_2 are the sensitivity parameters of x_2 to the investor sentiment and common noise. Without loss of generality, we assume further that these variables are independent of each other and have means zero and variances σ_S^2, σ_E^2 and σ_e^2, where the idiosyncratic noises e_1 and e_2 have the same variance. Then the covariance matrix of x_1 and x_2 is

$$\Sigma = \begin{pmatrix} \sigma_S^2+\sigma_E^2+\sigma_e^2 & \eta_1\sigma_S^2+\eta_2\sigma_E^2 \\ \eta_1\sigma_S^2+\eta_2\sigma_E^2 & \eta_1^2\sigma_S^2+\eta_2^2\sigma_E^2+\sigma_e^2 \end{pmatrix}. \tag{9}$$

With some algebra, we can solve the weights of the BW index on those proxies, which are the eigenvector corresponding to the larger eigenvalue of Σ, as

$$w^{\mathrm{BW}} \propto \begin{pmatrix} \frac{(1-\eta_1^2)\sigma_S^2+(1-\eta_2^2)\sigma_E^2}{2} + \sqrt{[\frac{(1-\eta_1^2)\sigma_S^2+(1-\eta_2^2)\sigma_E^2}{2}]^2+(\eta_1\sigma_S^2+\eta_2\sigma_E^2)^2} \\ \eta_1\sigma_S^2+\eta_2\sigma_E^2 \end{pmatrix}, \tag{10}$$

where $\propto$ is the proportion operator, indicating that the weights can be scaled by any positive real number. As long as $\eta_2 \neq 0$ in Equation (10), the BW index will have the common noise component in the weights. The greater the value of σ_E^2 is, the greater its influence on w^{BW} will be. Hence, the noise component can drastically alter the index. Indeed, if σ_E^2 approaches infinity, the weights converge to $(1, \eta_2)$. Hence, when σ_E^2 is large enough, the population BW index will be driven largely by the noise, and so will its sample estimate, the widely used BW index.

On the contrary, based on the theoretical results of Kelly and Pruitt (2014), the new index S^{PLS} will eliminate the noise asymptotically and converge to S. Hence, S^{PLS} should outperform S^{BW} in the presence of a common noise component.

2. Data

The aggregate stock market return is computed as the excess return as usual, which is the continuously compounded log return on the S&P 500 index (including dividends) minus the risk-free rate. The six individual investor sentiment proxies of Baker and Wurgler (2006, 2007) are

- *Close-end fund discount rate* (CEFD): value-weighted average difference between the net asset values of closed-end stock mutual fund shares and their market prices;
- *Share turnover* (TURN): log of the raw turnover ratio detrended by the past 5-year average, where raw turnover ratio is the ratio of reported share volume to average shares listed from the NYSE Fact Book;
- *Number of IPOs* (NIPO): monthly number of initial public offerings;
- *First-day returns of IPOs* (RIPO): monthly average first-day returns of initial public offerings;

- *Dividend premium* (PDND): log difference of the value-weighted average market-to-book ratios of dividend payers and nonpayers; and
- *Equity share in new issues* (EQTI): gross monthly equity issuance divided by gross monthly equity plus debt issuance.

The data on these measures are available from Jeffrey Wurgler's website, which provides the updated data.[3] The data span from July 1965 through December 2010 (546 months), and they have been widely used in a number of studies such as Baker and Wurgler (2006, 2007, 2012); Yu and Yuan (2011); Baker, Wurgler, and Yuan (2012); Stambaugh, Yu, and Yuan (2012); Yu (2013), and others. Because the data for the latest months are not available yet, our study here is confined to December 2010.

Using the PLS procedures in Section 1, we obtain the aligned investor sentiment index S^{PLS} from the six individual sentiment proxies,

$$\begin{aligned} S^{\mathrm{PLS}} = & -0.22\ \mathrm{CEFD} + 0.16\ \mathrm{TURN} - 0.04\ \mathrm{NIPO} \\ & + 0.63\ \mathrm{RIPO} + 0.07\ \mathrm{PDND} + 0.53\ \mathrm{EQTI}, \end{aligned} \tag{11}$$

where, following Baker and Wurgler (2006, 2007), each underlying individual measure is standardized, regressed on the growth of industrial production, the growth of durable consumption, the growth of nondurable consumption, the growth of service consumption, the growth of employment, and a dummy variable for NBER-dated recessions (to remove the effect of business-cycle variation), and smoothed with six-month moving average values (to iron out idiosyncratic jumps in the individual sentiment measures). The share turnover, average first-day return of IPOs, and dividend premium are lagged 12 months, relative to the other three measures in that these three variables take more time to reveal the same sentiment. Four of the six sentiment proxies (CEFD, TURN, RIPO, and EQTI) in S^{PLS} have the same signs as those in the BW index. However, it is interesting to note that, among the six proxies, RIPO and EQTI are the two most important underlying components in S^{PLS}, as they have the highest absolute coefficients. In contrast, they are just as important as the other proxies in BW index. Even though the weights for NIPO and PDND in S^{PLS} have opposite signs to those in BW index, their values are nearly zero and statistically insignificant.

Though the indices S^{PLS} and S^{BW} are constructed differently, they are highly correlated with each other with a positive correlation of 0.74. Consistent with the high correlation, Figure 1 shows that S^{PLS} appears to capture almost the same anecdotal accounts of fluctuations in sentiment with S^{BW}. Investor sentiment was low after the 1961 crash of growth stocks. It subsequently rose to a peak in the 1968 and 1969 electronics bubble. Sentiment fell again to a trough during the 1973 to 1974 stock market crash, but it picked up and

[3] The web page is http://people.stern.nyu.edu/jwurgler/.

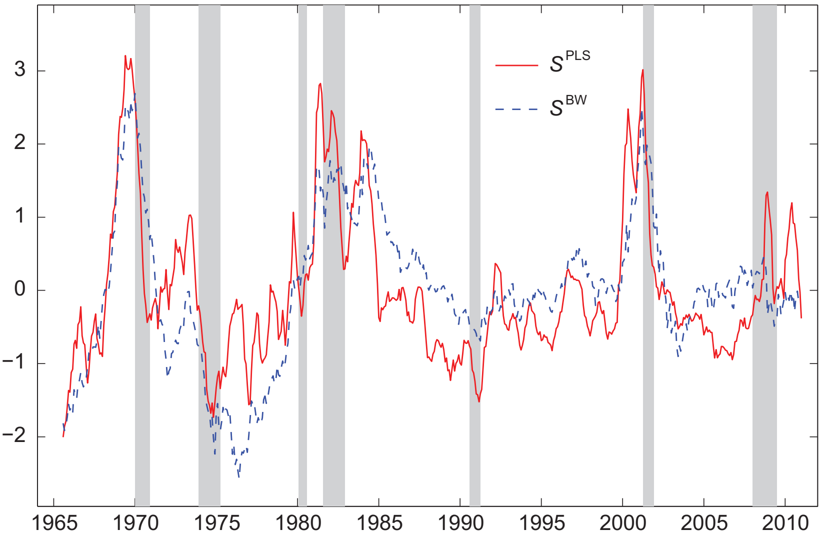

Figure 1
The investor sentiment index, July 1965 to December 2010
The solid line depicts the aligned investor sentiment index S^{PLS} extracted from the Baker and Wurgler's six individual investor sentiment proxies by applying the partial least squares method. The dashed line depicts the Baker and Wurgler (2006) investor sentiment index S^{BW} as the first principle component of the six investor sentiment measures. The six individual investor sentiment measures are available from Jeffrey Wurgler's website: the close-end fund discount rate, share turnover, number of IPOs, average first-day returns of IPOs, dividend premium, and equity share in new issues. Each underlying individual investor sentiment measure is standardized, smoothed with six-month moving average, and regressed on the growth of industrial production, the growth of durable consumption, the growth of nondurable consumption, the growth of service consumption, the growth of employment, and a dummy variable for NBER-dated recessions to remove the effect of macroeconomic conditions. The share turnover, average first-day return of IPOs, and dividend premium are lagged 12 months relative to the other three measures. The estimated investor sentiment indexes are standardized to have zero mean and unit variance. The vertical bars correspond to NBER-dated recessions.

reached a peak in the biotech bubble of the early 1980s. In the late 1980s, sentiment dropped but rose again in the early 1990s. It again reached a peak during the Internet bubble in the late 1990s. Sentiment dropped to a trough during the 2008 to 2009 subprime crisis but rose in 2010. Moreover, S^{PLS} appears to lead S^{BW} in many cases by several months, and looks slightly less persistent, suggesting that S^{PLS} may better capture the short-term variations in the expected excess market return compared to S^{BW} because the realized returns are volatile.

For interest of comparison, we also consider 18 monthly economic variables that are linked directly to economic fundamentals, which are the log dividend-price ratio (DP), log DY, log earnings-price ratio (EP), log dividend-payout ratio (DE), stock return variance (SVAR), BM ratio, net equity expansion (NTIS), Treasury bill rate (TBL), long-term bond yield (LTY), long-term bond return (LTR), term spread (TMS), default yield spread (DFY), default return spread (DFR), inflation rate (INFL), consumption-wealth ratio (CAY), log consumption surplus ratio (CSR), output gap (OG), and Kelly and Pruitt's

Table 1
Summary statistics

	Mean	Std	Skew	Kurt	Min	Max	$\rho(1)$	SR
R^m (%)	0.31	4.46	−0.67	5.41	−24.8	14.9	0.06	0.07
R^f (%)	0.46	0.25	0.72	4.33	0.00	1.36	0.98	
S^{PLS}	0.00	1.00	1.19	4.10	−2.01	3.21	0.96	
S^{BW}	0.00	1.00	0.10	3.19	−2.58	2.69	0.98	
DP	−3.56	0.42	−0.37	2.24	−4.52	−2.75	0.99	
DY	−3.56	0.42	−0.38	2.26	−4.53	−2.75	0.99	
EP	−2.82	0.47	−0.77	5.26	−4.84	−1.90	0.99	
DE	−0.74	0.32	3.08	19.0	−1.22	1.38	0.98	
SVAR (%)	0.23	0.45	9.48	116	0.01	6.55	0.49	
BM	0.52	0.28	0.57	2.25	0.12	1.21	0.99	
NTIS	0.01	0.02	−0.84	3.78	−0.06	0.05	0.98	
TBL (%)	5.49	2.95	0.72	4.33	0.03	16.3	0.98	
LTY (%)	7.29	2.40	0.89	3.34	3.03	14.8	0.99	
LTR (%)	0.65	3.06	0.40	5.55	−11.2	15.2	0.03	
TMS (%)	1.79	1.55	−0.33	2.63	−3.65	4.55	0.95	
DFY (%)	1.07	0.47	1.70	6.71	0.32	3.38	0.96	
DFR (%)	0.01	1.46	−0.29	10.0	−9.75	7.37	-0.06	
INFL (%)	0.36	0.35	−0.20	7.20	−1.92	1.79	0.61	
CAY (%)	0.08	1.82	0.18	2.24	−3.35	3.97	0.98	
CSR	−2.82	0.49	−3.41	15.4	−5.13	−2.47	0.99	
OG	0.00	0.06	−0.34	3.11	−0.18	0.13	0.99	
$\mathrm{BM}^{\mathrm{KP}}$	−0.00	0.86	−0.80	3.48	−2.68	1.53	0.98	

This table provides summary statistics for the excess market return (R^m, the log return on the S&P 500 index in excess of the risk-free rate), risk-free rate (R^f), aligned investor sentiment index (S^{PLS}) extracted by the partial least squares, Baker-Wurgler investor sentiment index (S^{BW}), the log dividend-price ratio (DP), log dividend yield (DY), log earnings-price ratio (EP), log dividend-payout ratio (DE), stock return variance (SVAR), book-to-market ratio (BM), net equity expansion (NTIS), Treasury bill rate (TBL), long-term bond yield (LTY), long-term bond return (LTR), term spread (TMS), default yield spread (DFY), default return spread (DFR), inflation rate (INFL), consumption-wealth ratio (CAY), log consumption surplus ratio (CSR), output gap (OG), and Kelly-Pruitt's BM ratio predictor ($\mathrm{BM}^{\mathrm{KP}}$). For each variable, the time-series average (Mean), standard deviation (Std), skewness (Skew), kurtosis (Kurt), minimum (Min), maximum (Max), and first-order autocorrelation ($\rho(1)$) are reported. The monthly Sharpe ratio (SR) is the mean excess market return divided by its standard deviation. The sample period is over July 1965 to December 2010.

disaggregated BM ratio factor ($\mathrm{BM}^{\mathrm{KP}}$). Details on these economic predictors are provided in the Appendix.

Table 1 reports the summary statistics of the data. The monthly excess market return has a mean of 0.31% and a standard deviation of 4.46%, implying a monthly Sharpe ratio of 0.07. Even though the excess market return has little autocorrelation, most of the other variables are quite persistent. The summary statistics are generally consistent with the literature.

3. Empirical Results

3.1 Forecasting the market

Consider the standard predictive regression model,

$$R^m_{t+1} = \alpha + \beta S^k_t + \epsilon_{t+1}, \qquad k = \mathrm{PLS}, \mathrm{BW}, \mathrm{EW}, \tag{12}$$

where R^m_{t+1} is the excess market return (i.e., the monthly log return on the S&P 500 index in excess of the risk-free rate), S^{PLS}_t is the aligned investor sentiment index, and S^{BW}_t is the BW index. For comparison, we also consider a naive investor sentiment index, S_t^{EW}, that places equal weights on the six individual sentiment proxies of Baker and Wurgler (2006, 2007). The null hypothesis of interest is that investor sentiment has no predictive ability, $\beta=0$. In this case, Equation (12) reduces to the constant expected return model, $R_{t+1}^m=\alpha+\epsilon_{t+1}$. Because finance theory suggests a negative sign of β, we test $H_0:\beta=0$ against $H_A:\beta<0$, which is closer to theory than the common alternative of $\beta\neq0$. Econometrically, Inoue and Kilian (2004) suggest the use of the one-sided alternative hypothesis, which usually increases the power of the test.

Statistically, there are three issues that may have adverse effect on the statistical inference about the aligned sentiment index. First, there is potentially a spurious regression concern when a predictor is highly persistent (Ferson, Sarkissian, and Simin 2003). Second, owing to the well-known Stambaugh (1999) small-sample bias, the coefficient estimate of the predictive regression can be biased in finite sample, which may distort the t-statistic when the predictor is highly persistent and correlated with the excess market return. Third, the first-step regression for the in-sample PLS estimation, Equation (4), introduces a look-forward bias as it uses future information. Although Kelly and Pruitt (2013, 2014) show that this bias will vanish as the sample size becomes large enough, it is still a concern with the finite sample here.

We employ three strategies to alleviate potential concerns over the above three issues. First, we base our inference on the empirical p-values using a wild bootstrap procedure that accounts for the persistence in predictors, correlations between the excess market return and predictor innovations, and general forms of return distribution. Second, we calculate the Stambaugh (1999) bias-adjusted regression coefficients following Amihud, Hurvich, and Wang (2009). Third, we construct a look-ahead bias-free PLS forecast. To calculate S_t^{PLS} at time t, we run the first-step time-series regression, Equation (4), now with information up to time t only. Then, the regression slopes are used as independent variables for the second-step regression, Equation (5), the slope of which therefore is the aligned sentiment S^{PLS} at time t. Repeating this procedure recursively, we obtain a look-ahead bias-free aligned sentiment index. In the paper, we use the first 12-year data (one fourth of the samples) as the initial training sample when computing recursively the look-ahead bias-free aligned investor sentiment.

Table 2 reports the results of the predictive regression. Panel A provides the estimation results for the BW index, S^{BW}, over the sample period of July 1965 through December 2010. Consistent with theory, S^{BW} is a negative return predictor: high sentiment is associated with the expected excess market return in the next month with a regression slope, β, of -0.24. However, S^{BW} only generates a small Newey-West t-statistic (which is computed using a lag of 12 throughout) of -1.21 and an R^2 of only 0.30%. In this sense, the forecasting power of S^{BW} for the excess market return is insignificant, confirming the earlier finding of Baker and Wurgler (2007).

Panel B of Table 2 reports the performance for the equally-weighted naive investor sentiment index, S^{EW}. Interestingly, this simple index, which requires

Table 2
Forecasting market return with investor sentiment

	β (%)	t-stat	R^2 (%)	R^2_{up} (%)	R^2_{down} (%)	R^2_{high} (%)	R^2_{low} (%)
Panel A: BW investor sentiment index							
S^{BW}	−0.24	−1.21	0.30	0.12	0.51	0.90	−0.67
Panel B: Naive investor sentiment index							
S^{EW}	−0.27*	−1.39	0.38	0.21	0.74	1.05	−0.74
Panel C: Aligned investor sentiment index							
S^{PLS}	−0.58***	−3.04	1.70	1.54	2.11	2.74	−0.00
	(OLS forecast)						
	−0.59***	−3.08	1.70	1.53	2.12	2.75	−0.02
	(Stambaugh bias-adjusted forecast)						
	−0.57**	−2.24	1.21	0.77	2.40	1.96	−0.15
	(Look-ahead bias-free forecast)						
Panel D: Individual investor sentiment proxies							
CEFD	0.16	0.89	0.14	0.00	0.42	0.45	−0.40
TURN	−0.13	−0.69	0.08	−0.05	0.39	0.38	−0.39
NIPO	0.04	0.18	0.01	0.01	0.00	−0.06	0.13
RIPO	−0.47**	−2.35	1.16	1.61	0.25	2.09	−0.34
PDND	−0.05	−0.27	0.02	0.03	0.00	−0.12	0.24
EQTI	−0.40**	−2.26	0.80	0.41	1.69	0.70	0.98
Kitchen sink			3.02	2.11	5.07	3.48	2.21

This table provides in-sample estimation results for the predictive regression

$$R^m_{t+1} = \alpha + \beta S_t + \epsilon_{t+1},$$

where R^m_{t+1} denotes the monthly excess market return and S_t is a predictor or a set of predictors. Panel A considers the Baker-Wurgler investor sentiment index (S^{BW}), and Panel B considers the naive investor sentiment index (S^{EW}), which is defined as the equally-weighted average of the six sentiment proxies. Panel C reports the results of the aligned investor sentiment index (S^{PLS}) with the ordinary least squares (OLS) approach, the Stambaugh (1999) small-sample bias-adjusted approach, and the look-ahead bias-free approach, respectively. Panel D reports the results using the six individual sentiment proxies separately. The kitchen sink represents the case using all the proxies in a multivariate predictive regression. R^2_{up} (R^2_{down}) statistics are calculated over NBER-dated business-cycle expansions (recessions), and R^2_{high} (R^2_{low}) are calculated over high (low) sentiment periods, respectively. The Newey-West t-statistic (with a lag of 12) is reported. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively, based on one-sided wild bootstrapped p-values. The sample period is over July 1965 to December 2010.

no estimation of combining weights at all, performs as well as S^{BW}. The regression slope β is equal to −0.27, slightly more negative than −0.24. The t-statistic is slightly larger in absolute value, with marginally statistical significance at the 10% level. The R^2 is slightly greater too.

Panel C of Table 2 reports the estimation results for the aligned investor sentiment. Like S^{BW} and S^{EW}, S^{PLS} is a negative return predictor for the aggregate market, and it performs the best among the three indices. With the standard OLS predictive regression, S^{PLS} has a regression slope of −0.58 that is statistically significant at the 1% level based on the wild bootstrap p-value. When correcting for the Stambaugh (1999) small-sample bias, the regression slope is virtually identical to the OLS regression, −0.59 versus −0.58. The biased adjusted slope is slightly larger in absolute value because the correlation between the forecasting error and the innovation in the predictor S^{PLS} is positive, in contrast to the case when the DY is the predictor. Because we use the first 12-year data for sample training, the estimation results for the look-ahead bias-free aligned sentiment is based on the sample period of July 1977 through December 2010. Interestingly, the results are again almost the same as the OLS regression. The regression slope is −0.56, with a Newey-West t-statistic of −2.24.

After carefully examining the potential biases for the slope of S^{PLS}, we still have a value around −0.58%. Economically, the OLS coefficient suggests that a one-standard-deviation increase in S^{PLS} is associated with a −0.58% decrease in expected excess market return for the next month. On the one hand, recall that the average monthly excess market return during our sample period is only 0.31%, thus the slope of −0.58% implies that the expected excess market return based on S^{PLS} varies by about two times larger than its average level, signaling strong economic significance (Cochrane 2011). On the other hand, if we annualize the 0.58% decrease in one month by the multiplication of 12, the annualized level of 6.96% is somewhat large. In this case, one may interpret this as the model-implied expected change that may not be identical to the reasonable expected change of the investors in the market. Empirically, this level is comparable with conventional macroeconomic predictors. For example, a one-standard-deviation increase in the DP ratio, the CAY, and the net payout ratio tends to increase the risk premium by 3.60%, 7.39%, and 10.2% per annum, respectively (see, e.g., Lettau and Ludvigson 2001; Boudoukh et al. 2007).

The R^2 of S^{PLS} with OLS forecast is 1.70%, substantially greater than 0.30% and 0.38% of S^{BW} and S^{EW}. With the correction for the Stambaugh bias, it barely changes the value. However, the look-ahead bias-free index has a smaller R^2 of 1.21%. This is expected as the look-ahead information is eliminated. Economically, if this level of predictability can be sustained out-of-sample, it will be of substantial economic significance (Kandel and Stambaugh 1996). Indeed, Campbell and Thompson (2008) show that, given that the large unpredictable component inherent in the monthly market returns, a monthly out-of-sample R^2 statistic of 0.5% can generate significant economic value. This point will be analyzed further in Section 3.3.

For comparison, Panel D of Table 2 reports the predictive abilities of the six individual sentiment proxies on the market. The slopes of CEFD, TURN, RIPO, and EQTI are consistent with the theoretical predictions, but the signs of NIPO and PDND are not. However, the predictability of the latter two is very weak, with R^2s of 0.01% and 0.02%, confirming why they have negligible weights (−0.04 and 0.07) in constructing S^{PLS}. RIPO and EQTI display higher power in forecasting the excess market returns, consistent with their relatively higher weights in forming the S^{PLS} index. Overall, S^{PLS} beats all the individual proxies, providing direct support to Baker and Wurgler (2006, 2007) that an aggregate sentiment index is more desirable than any individual proxies.

An interesting question is how well the prediction performs if we use all six sentiment proxies in one single multiple predictive regression. This is known as a kitchen sink model in the predictability literature. The last row of Panel D reports the in-sample R^2, 3.02%. This is the highest value of all the predictive R^2s in Table 2. However, Goyal and Welch (2008), among others, find that the kitchen sink model usually suffers from a serious over-fitting issue and its out-of-sample performance is very poor. We will show later that this is also true in our case here, even though the number of regressors is as few as six.

From an economic point of view, while the overall R^2 is interesting, it is also important to analyze the predictability during business cycles to understand better about the fundamental driving forces. Following Rapach, Strauss, and Zhou (2010) and Henkel, Martin, and Nardari (2011), we compute the R^2 statistics separately for economic expansions (R^2_{up}) and recessions (R^2_{down}),

$$R^2_{\text{c}} = 1 - \frac{\sum_{t=1}^{T} I_t^{\text{c}} (\hat{\epsilon}_{i,t})^2}{\sum_{t=1}^{T} I_t^{\text{c}} (R_t^m - \bar{R}^m)^2} \qquad c = \text{up, down}, \tag{13}$$

where I_t^{up} (I_t^{down}) is an indicator that takes a value of one when month t is in an NBER expansion (recession) period and zero otherwise; $\hat{\epsilon}_{i,t}$ is the fitted residual based on the in-sample estimates of the predictive regression model in Equation (12); $\bar{R}^m$ is the full-sample mean of R_t^m; and T is the number of observations for the full sample. Note that, unlike the full-sample R^2 statistic, the R^2_{up} and R^2_{down} statistics can be both positive or negative.

Columns 5 and 6 of Table 2 report the R^2_{up} and R^2_{down} statistics. Panels A and B show that the return predictability is higher over recessions for S^{BW} and S^{EW}. Panel C and the last row of Panel D show that S^{PLS} and the kitchen sink model present strong in-sample forecasting ability during both expansions and recessions, although the predictability is relatively stronger during recessions vis-á-vis expansions. Regarding individual sentiment proxies, the predictability for CEFD, TURN, and EQTI is stronger during recessions, whereas NIPO, RIPO, and PDND display stronger abilities during expansions. Comparing Panels C and D, the better performance of S^{PLS} over both expansions and recessions is due to the fact that S^{PLS} places the largest two weights on RIPO and EQTI that have stronger predicting power in expansions and recessions, respectively. It is perhaps this reason why the aligned sentiment index is useful in forecasting the aggregate market during both expansions and recessions, though the power is generally stronger over recessions.

In the last two columns of Table 2, we divide the whole sample into high- and low-sentiment periods to investigate the possible sources of the improved predictive power of S^{PLS}. Following Stambaugh, Yu, and Yuan (2012), we classify a month as high (low) sentiment if the sentiment level (S^{PLS}) in the previous month is above (below) its median value for the sample period, and compute the R^2_{high} and R^2_{low} statistics for the high- and low-sentiment periods, respectively, in a manner similar to Equation (13). Interestingly, consistent with Shen and Yu (2013), who find that the predictive power of the BW index S^{BW} is significant during high-sentiment periods and insignificant

Table 3
Forecast encompassing tests

	CEFD	TURN	NIPO	RIPO	PDND	EQTI	S^{BW}	S^{PLS}
CEFD		0.35	0.50	0.01	0.44	0.02	0.12	0.01
TURN	0.45		0.50	0.01	0.45	0.02	0.12	0.01
NIPO	0.39	0.32		0.01	0.43	0.02	0.12	0.01
RIPO	0.51	0.52	0.50		0.47	0.06	0.48	0.07
PDND	0.40	0.34	0.49	0.01		0.02	0.12	0.01
EQTI	0.47	0.50	0.50	0.08	0.49		0.38	0.06
S^{BW}	0.55	0.53	0.51	0.03	0.43	0.03		0.02
S^{PLS}	0.54	0.52	0.50	0.40	0.46	0.19	0.64	

This table reports p-values for the Harvey, Leybourne, and Newbold (1998) statistic. The statistic corresponds to a one-sided (upper-tail) test of the null hypothesis that the predictive regression forecast for the monthly excess market return based on one of the predictors given in the first column encompasses the forecast based on one of the predictors given in the first row, against the alternative hypothesis that the forecast given in the first column does not encompass the forecast given in the first row. The predictors are the Baker-Wurgler investor sentiment index S^{BW}, the aligned investor sentiment index S^{PLS}, and the six individual investor sentiment proxies. The sample period is over July 1965 to December 2010.

during low-sentiment periods, we find that the predictive power of S^{PLS} also concentrates over high-sentiment periods. For example, over high-sentiment periods, S^{PLS} has an R^2_{high} of 2.74% (versus 0.90% of S^{BW}). In contrast, over low-sentiment periods, S^{PLS} has an R^2_{low} of zero (versus −0.67% of S^{BW}). In short, consistent with Shen and Yu (2013), we find that investor sentiment's predictive power mainly comes from high-sentiment periods, even with our new investor sentiment index, during which mispricing is more likely due to short-sale constraints.

Summarizing Table 2, the aligned investor sentiment S^{PLS} exhibits statistically and economically significant in-sample predictability for the monthly excess market return, whereas the BW index S^{BW} does not. In addition, S^{PLS} predicts the market in both expansions and recessions, as well as in high-sentiment periods. The results are consistent with our early econometric objective of enhancing the forecasting power by eliminating the common noise component of the proxies, which is made possible with the PLS developed further by Kelly and Pruitt (2013, 2014).

To further assess the relative information content in S^{PLS}, S^{BW}, and the six proxies, we conduct a forecast encompassing test. Harvey, Leybourne, and Newbold (1998) develop a statistic for testing the null hypothesis whether a given forecast contains all of the relevant information found in a competing forecast (i.e., the given forecast encompasses the competitor) against the alternative that the competing forecast contains relevant information beyond that in the given forecast.

Table 3 reports p-values of the test. We summarize the results with three observations. First, none of the individual investor sentiment measures of Baker and Wurgler (2006, 2007) encompasses all of the remaining individual measures, indicating potential gains from combining individual measures into a common index to make use of additional information. Second, S^{BW} fails to encompass two of the six individual measures, implying that S^{BW} does not make

Table 4
Comparison with economic return predictors

	Panel A: Univariate predictive regressions $R_{t+1}^{m}=\alpha+\psi Z_t^k+\epsilon_{t+1}$			Panel B: Bivariate predictive regressions $R_{t+1}^{m}=\alpha+\beta S_t^{\rm PLS}+\psi Z_t^k+\epsilon_{t+1}$				
	ψ (%)	t-stat	R^2 (%)	β (%)	t-stat	ψ (%)	t-stat	R^2 (%)
DP	0.47	0.99	0.20	−0.59***	−3.02	0.49	1.02	1.91
DY	0.54	1.13	0.26	−0.58**	−3.01	0.53	1.14	1.96
EP	0.21	0.43	0.05	−0.58**	−3.03	0.19	0.38	1.74
DE	0.36	0.50	0.07	−0.59**	−3.06	0.44	0.61	1.80
SVAR	−1.09**	2.29	1.23	−0.55**	−2.82	−0.99**	2.00	2.70
BM	0.15	0.20	0.01	−0.59**	−2.95	0.38	0.49	1.76
NTIS	−3.70	−0.33	0.03	−0.59**	−2.90	−1.16	−0.10	1.71
TBL	−0.07	−0.94	0.19	−0.57**	−2.62	−0.01	−0.15	1.71
LTY	0.00	0.05	0.00	−0.62**	−2.90	0.06	0.66	1.80
LTR	0.15**	2.22	1.07	−0.57**	−2.97	0.15**	2.21	2.72
TMS	0.23**	1.83	0.61	−0.54**	−2.73	0.18*	1.39	2.06
DFY	0.46	0.90	0.23	−0.68***	−3.36	0.81**	1.59	2.38
DFR	0.18	0.89	0.36	−0.58**	−3.01	0.18	0.88	2.05
INFL	0.18	0.27	0.02	−0.58**	−3.02	0.23	0.34	1.73
CAY	0.24***	2.73	0.97	−0.53**	−2.73	0.20**	2.21	2.36
CSR	−0.59	−1.24	0.40	−0.63***	−3.29	−0.75	−1.62	1.99
OG	−0.09***	−2.79	1.55	−0.54**	−2.78	−0.09**	2.54	3.01
BM^{KP}	0.64***	2.97	2.07	−0.47**	−2.41	0.59**	2.73	3.41
ECON^{PC}	0.06	0.63	0.09	−0.60**	−3.01	0.06	0.69	1.82
ECON^{PLS}	0.91***	4.77	4.12	−0.46**	−2.37	0.84***	4.38	5.16
$(S\text{+ECON})^{\text{PC}}$	0.05	0.60	0.08					
$(S\text{+ECON})^{\text{PLS}}$	−1.02***	−5.28	5.12					

This table reports the in-sample estimation results for the predictive regression of monthly excess market return on one of the 18 economic predictors Z^k, and on both the lagged aligned sentiment index ($S^{\rm PLS}$) and Z^k, respectively. The first column of the first 18 rows are the individual economic variables. ECON^{PC} is the first principal component (PC) factor extracted from the 18 economic variables. ECON^{PLS} is the extracted PLS predictor based on the 18 economic variables. $(S\text{+ECON})^{\text{PC}}$ and $(S\text{+ECON})^{\text{PLS}}$ are two predictors extracted by applying the PC approach and the PLS approach to the union of 18 economic variables and six individual sentiment proxies, respectively. The Newey-West t-statistic as well as R^2 are reported. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively, based on one-sided wild bootstrapped p-values. The sample period is over July 1965 to December 2010.

full use of all of the relevant information in the individual measures. Third, as expected, $S^{\rm PLS}$ encompasses all of the individual investor sentiment measures as well as $S^{\rm BW}$ at the conventional significant level. Therefore, the forecast encompassing test suggests that $S^{\rm PLS}$ is an efficient index that incorporates all of the relevant forecasting information, which helps in understanding why it has superior forecasting performance as reported in Table 2.

3.2 Comparison with economic predictors

In this subsection, we compare the forecasting power of aligned investor sentiment index $S^{\rm PLS}$ with economic predictors and examine whether its forecasting power is driven by omitted economic variables related to business-cycle fundamentals or changes in investor risk aversion.

First, we consider the predictive regression on a single economic variable,

$$R_{t+1}^{m}=\alpha+\psi Z_t^k+\epsilon_{t+1}, \qquad k=1,...,18, \tag{14}$$

where Z_t^k is one of the 18 economic predictors described in Section 2 and the Appendix.

The first 18 rows of Panel A of Table 4 report the estimation results for Equation (14). Out of the 18 economic predictors, only stock return variance (SVAR), long-term government bond return (LTR), term spread (TMS), consumption-wealth ratio (CAY), output gap (OG), and Kelly and Pruitt's disaggregated BM ratio factor (BM^{KP}) exhibit significant predictive abilities for the market at the 5% or better significance levels. Among these six significant economic variables, three of them have R^2s larger than 1% (LTR, OG, BM^{KP}), and one has R^2 larger than 2% (BM^{KP}). Hence, S^{PLS} outperforms 17 out of 18 individual economic predictors, except for BM^{KP}, in forecasting the monthly excess market returns in sample.

We then investigate whether the forecasting power of S^{PLS} remains significant after controlling for economic predictors. To analyze the incremental forecasting power of S^{PLS}, we conduct the following bivariate predictive regressions based on S_t^{PLS} and Z_t^k,

$$R_{t+1}^m = \alpha + \beta S_t^{PLS} + \psi Z_t^k + \epsilon_{t+1}, \qquad k = 1, ..., 18. \tag{15}$$

We are interested in the regression slope β of S_t^{PLS}, and test $H_0 : \beta = 0$ against $H_A : \beta < 0$ based on the wild bootstrapped p-values.

Panel B of Table 4 shows that the estimates of the slope β in Equation (15) are negative and large, in line with the results in the predictive regression, Equation (12), reported in Table 2. More importantly, β remains statistically significant when augmented by the economic predictors. All of the R^2s in Equation (15) are substantially larger than those in Equation (14) based on the economic predictors alone. These results demonstrate that S^{PLS} contains sizable complementary forecasting information beyond what is contained in the economic predictors.[4]

The next question of interest is how well PLS and PC perform when they are applied to all the economic variables or combining economic variables with the Baker and Wurgler (2006) proxies. The last four rows of Table 4 report the results. Based on all 18 economic variables, the PC predictor, $ECON^{PC}$, has an R^2 of only 0.09%, much smaller than 4.12% of the PLS predictor, $ECON^{PLS}$. When combining all the economic variables with the sentiment proxies, $(S+ECON)^{PLS}$ yields an in-sample R^2 of 5.12% and is significant at the 1% level, whereas $(S+ECON)^{PC}$ has again a small R^2 of only 0.08%, and is insignificant. In comparison with earlier results, the PLS not only outperforms the PC in all cases, but also has substantially higher R^2s when the predictors are combined. However, later in Section 3.3, we find that the strong in-sample predictability of $ECON^{PLS}$ and $(S+ECON)^{PLS}$ is not sustainable out of sample.

3.3 Out-of-sample forecasts

Even though the in-sample analysis provides parameter estimates that are more efficient and thus more precise return forecasts by utilizing all available

4 This result does not apply to S^{BW} and is not reported for brevity, but it is available upon request.

data, Goyal and Welch (2008), among others, argue that out-of-sample tests seem more relevant for assessing genuine return predictability in real time and avoiding the in-sample over-fitting issue. In addition, out-of-sample tests are much less affected by the small-sample size distortions such as the Stambaugh bias and the look-ahead bias concern of the PLS approach (Kelly and Pruitt 2013). Hence, it is of interest to investigate the out-of-sample predictive performance of investor sentiment.

The key requirement for out-of-sample forecasts at time t is that we can only use information available up to t to forecast stock returns at $t+1$. Following Goyal and Welch (2008), Kelly and Pruitt (2013), and many others, we run the out-of-sample analysis by estimating the predictive regression model recursively, based on different measures of investor sentiment,

$$\hat{R}^m_{t+1} = \hat{\alpha}_t + \hat{\beta}_t S^k_{1:t;t}, \tag{16}$$

where $\hat{\alpha}_t$ and $\hat{\beta}_t$ are the OLS estimates from regressing $\{R^m_{s+1}\}_{s=1}^{t-1}$ on a constant and a sentiment measure $\{S^k_{1:t;s}\}_{s=1}^{t-1}$. Like our in-sample analogues in Table 2, we consider the recursively estimated BW investor sentiment index $S^{\mathrm{BW}}_{1:t;t}$, the equally-weighted naive investor sentiment index $S^{\mathrm{EW}}_{1:t;t}$, and the recursively estimated aligned investor sentiment index $S^{\mathrm{PLS}}_{1:t;t}$.

For interest of comparison, we consider also the combination forecast that is widely used in econometric forecasting applications and that often beats sophisticated optimally estimated forecasting weights (Timmermann 2006). In finance, Rapach, Strauss, and Zhou (2010) show that a simple equally-weighted average of univariate regression forecasts can consistently predict the market risk premium. It is hence of interest to see how well it performs in the context of using the six individual sentiment proxies.

Let p be a fixed number chosen for the initial sample training, so that the future expected return can be estimated at time $t = p+1, p+2, \ldots, T$. Hence, there are q $(=T-p)$ out-of-sample evaluation periods. That is, we have q out-of-sample forecasts: $\{\hat{R}^m_{t+1}\}_{t=p}^{T-1}$. More specifically, we use the data over July 1965 through December 1984 as the initial estimation period, so that the forecast evaluation period spans over January 1985 through December 2010. The length of the initial in-sample estimation period balances having enough observations for precisely estimating the initial parameters with the desire for a relatively long out-of-sample period for forecast evaluation.[5]

We evaluate the out-of-sample forecasting performance based on the widely used Campbell and Thompson (2008) R^2_{OS} statistic, the Diebold and Mariano (1995) t-statistic modified by McCracken (2007), and the Clark and West (2007) MSFE-adjusted statistic. The R^2_{OS} statistic measures the proportional reduction

[5] Hansen and Timmermann (2012) and Inoue and Rossi (2012) show that out-of-sample tests of predictive ability have better size properties when the forecast evaluation period is a relatively large proportion of the available sample, as in our case.

in mean squared forecast error (MSFE) for the predictive regression forecast relative to the historical average benchmark,

$$R^2_{\mathrm{OS}} = 1 - \frac{\sum_{t=p}^{T-1}(R^m_{t+1} - \hat{R}^m_{t+1})^2}{\sum_{t=p}^{T-1}(R^m_{t+1} - \bar{R}^m_{t+1})^2}, \tag{17}$$

where $\bar{R}^m_{t+1}$ denotes the historical average benchmark corresponding to the constant expected return model ($R^m_{t+1} = \alpha + \epsilon_{t+1}$),

$$\bar{R}^m_{t+1} = \frac{1}{t}\sum_{s=1}^{t} R^m_s. \tag{18}$$

Goyal and Welch (2008) show that the historical average is a very stringent out-of-sample benchmark, and individual economic variables typically fail to outperform the historical average. The R^2_{OS} statistic lies in the range $(-\infty, 1]$. If $R^2_{OS} > 0$, it means that the forecast $\hat{R}^m_{t+1}$ outperforms the historical average $\bar{R}^m_{t+1}$ in term of MSFE.

The second statistic we report is Diebold and Mariano (1995) statistic modified by McCracken (2007) (DM-test hereafter), which tests for the equality of the mean squared forecast errors (MSFE) of one forecast relative to another. Here our null hypothesis is that the historical average has a MSFE that is less than, or equal to, that of the predictive regression model. Comparing a predictive regression forecast to the historical average entails comparing nested models, as the predictive regression model reduces to the historical average under the null hypothesis. McCracken (2007) shows that the modified DM-test statistic follows a nonstandard normal distribution when testing nested models, and provides bootstrapped critical values for the nonstandard distribution.

The third statistic is the MSFE-adjusted statistic of Clark and West (2007) (CW-test hereafter). It tests the null hypothesis that the historical average MSFE is less than or equal to the predictive regression forecast MSFE against the one-sided (upper-tail) alternative hypothesis that the historical average MSFE is greater than the predictive regression forecast MSFE, corresponding to H_0: $R^2_{\mathrm{OS}} \leq 0$ against H_A : $R^2_{OS} > 0$. Clark and West (2007) show that the test has an asymptotically standard normal distribution when we compare different forecasts from the nested models. Intuitively, under the null hypothesis that the constant expected return model generates the data, the predictive regression model produces a noisier forecast than the historical average benchmark because it estimates slope parameters with zero population values. We thus expect the benchmark model's MSFE to be smaller than the predictive regression model's MSFE under the null. The MSFE-adjusted statistic accounts for the negative expected difference between the historical average MSFE and predictive regression MSFE under the null, so that it can reject the null even if the R^2_{OS} statistic is negative.

Panel A of Table 5 shows that the BW index S^{BW} generates a positive R^2_{OS} statistic (0.15%), and thus delivers a lower MSFE than the historical average.

Table 5
Out-of-sample forecasting results

	R^2_{OS} (%)	DM-test	CW-test	$R^2_{OS,\mathrm{up}}$ (%)	$R^2_{OS,\mathrm{down}}$ (%)
Panel A: Investor sentiment indexes					
S^{BW}	0.15	0.58	0.96	0.09	0.49
S^{EW}	0.38	1.18	1.76**	0.16	1.03
S^{Com}	0.42	1.30	1.76**	0.29	0.83
S^{PLS}	1.23	4.54***	1.97**	0.90	3.21
$\mathrm{ECON}^{\mathrm{PLS}}$	0.07	0.21	1.26	−5.41	1.70
$(S\text{+ECON})^{\mathrm{PLS}}$	0.29	0.91	1.48*	−4.78	2.31
Panel B: Individual investor sentiment proxies					
CEFD	0.06	0.19	0.56	−0.02	0.33
TURN	−0.02	−0.07	−0.07	−0.50	0.40
NIPO	−0.54	−1.69	−2.41	−0.72	0.06
RIPO	0.97	3.07**	1.54*	1.60	−1.12
PDND	−0.11	−0.34	−0.20	0.08	−0.74
EQTI	0.69	2.15**	1.46*	0.37	1.75
Kitchen sink	0.27	0.84	1.73**	−0.15	1.68

This table reports the out-of-sample performances of various measures of investor sentiment in predicting the monthly excess market return. Panel A provides the results using the Baker-Wurgler sentiment index S^{BW}, the equally-weighted naive sentiment index S^{EW}, the combination forecast S^{Com}, the aligned sentiment index S^{PLS}, $\mathrm{ECON}^{\mathrm{PLS}}$ extracted from the 18 economic variables, and $(S\text{+ECON})^{\mathrm{PLS}}$ extracted from the union of the 18 economic variables and the six sentiment proxies. Panel B are generated by using one of six individual sentiment proxies or by using all of them in a multivariate regression (the kitchen sink model). All of the predictors and regression slopes are estimated recursively using the data available at the forecast formation time t. R^2_{OS} is the Campbell and Thompson (2008) out-of-sample R^2. DM-test is the modified Diebold and Mariano (1995) t-statistic and CW-test is the Clark and West (2007) MSFE-adjusted statistic. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. $R^2_{OS,\mathrm{up}}$ ($R^2_{OS,\mathrm{down}}$) statistics are calculated over NBER-dated business-cycle expansions (recessions). The out-of-sample evaluation period is over January 1985 to December 2010.

However, this outperformance is not statistically significant according to the DM- and CW-test statistics. Thus, S^{BW} has weak out-of-sample predictive ability for the aggregate stock market, confirming our previous in-sample results (Table 2). The equally-weighted naive sentiment index S^{EW} slightly improves the performance to 0.38%, significant with the CW-test but insignificant with the DW-test, owing to the reduced estimation errors for index weights. Consistent with Rapach, Strauss, and Zhou (2010), the combination forecast S^{Com} can further enhance the forecasting performance to 0.42%, again significant with the CW-test but not so with the DW-test.

In contrast, S^{PLS} exhibits much stronger out-of-sample predictive ability for the aggregate market. Its R^2_{OS} is 1.23%, exceeding all of the R^2_{OS}s substantially in Table 5 with other forecasting approaches. The DM- and CW-test statistics of S^{PLS} are 4.54 and 1.97, suggesting that S^{PLS}'s MSFE is significantly smaller than that of the historical average at the 5% or better significant level. In addition, the fifth and sixth columns of Table 5 show that, even though the predictability of S^{BW}, S^{EW}, and S^{Com} is only concentrated in recessions, S^{PLS} presents strong out-of-sample forecasting ability during both expansions and recessions, although the ability is relatively stronger during recessions as well.

The last two rows of Panel A report the out-of-sample performances of $\mathrm{ECON}^{\mathrm{PLS}}$ and $(S\text{+ECON})^{\mathrm{PLS}}$. $\mathrm{ECON}^{\mathrm{PLS}}$ generates a small positive R^2_{OS} of

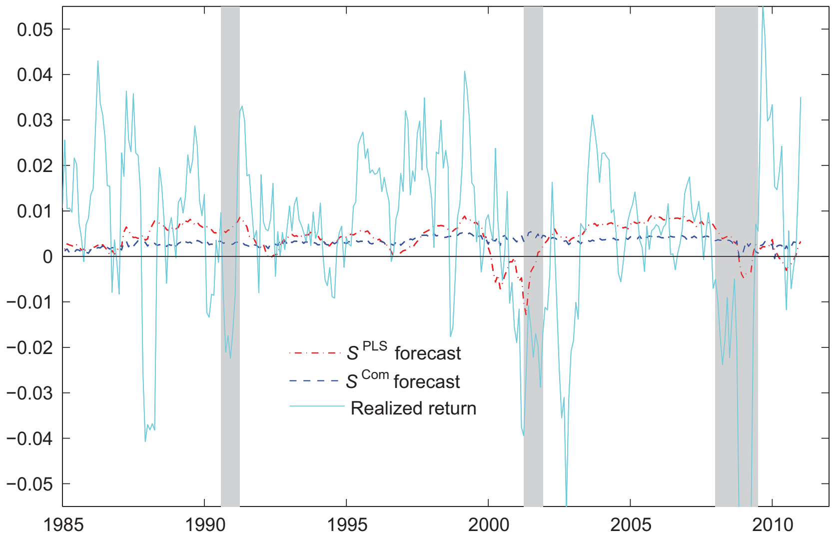

Figure 2
Excess market return forecasts of S^{PLS} and S^{Com}, January 1985 to December 2010
The dotted line depicts the out-of-sample predictive regression forecast for excess market return based on the recursively constructed aligned investor sentiment index S^{PLS}. The dashed line depicts the out-of-sample excess market return forecast based on combination forecast S^{Com} as the simple equally-weighted average of univariate predictive regression forecasts. The solid line depicts the excess market return smoothed with six month moving average. The sentiment indices and excess market return forecasts are estimated recursively based on information up to the period of forecast formation period t alone. The vertical bars correspond to NBER-dated recessions.

0.07%, which is much smaller than its in-sample value (4.12%), seen earlier in Table 2. This is not surprising since all economic variables, except for the Kelly-Pruitt's predictor, generate negative R^2_{OS}s, and they are very instable predictors, as emphasized by Goyal and Welch (2008). When the six sentiment proxies are added to the economic variables, the R^2_{OS} is improved to 0.29%, significant under the CW-test. However, the economic magnitude is still smaller than that of S^{PLS} (1.23%), when applying the PLS method to the sentiment variables alone. The results suggest that, even though more predictors tend to improve in-sample performance using the PLS, the out-of-sample performance may not necessarily be improved.

Because the combination forecast is widely known as a viable predictor and it performs the second best here, it is of interest to examine further its relation with S^{PLS}. A simple correlation analysis shows that the combination forecast and the S^{PLS} forecast have a high correlation of 79%. Intuitively, this suggests that they are likely to capture very much similar sentiment shifts of the same proxies. Hence, their economic sources of predictability are likely the same. To understand their differences in forecasting power, Figure 2 depicts the forecasted returns based on S^{PLS} and S^{Com} for the January 1985 through December 2010 out-of-sample period. It is clear that the PLS forecasted returns are much more volatile than the combination forecasts. As the actual realized excess returns (plotted in the figure as the six-month moving average for better visibility) are even more volatile than the PLS forecasted returns. This may explain intuitively why the PLS method does a better job than the combination forecast approach here in capturing the expected variation in the market return.

An interesting observation from Figure 2 is that there are long periods during which S^{PLS} provides negative predicted values of the expected excess market returns. During these periods, the sentiment is high and mispricing is possible, especially on the short legs of various long-short investment strategies (Stambaugh, Yu, and Yuan 2012), due to limits to arbitrage. On the contrary, Pettenuzzo, Timmermann, and Valkanov (2014) provide two well-motivated economic constraints, non-negative equity premiums and bounds on the conditional Sharpe ratio, and find that they improve substantially the out-of-sample predictability of a number of macroeconomic variables. An interesting open question, which is out of the scope of this paper, is whether one can improve further Pettenuzzo, Timmermann, and Valkanov's novel approach to place their constraints on and off in some optimal fashion overtime to account for the possible mispricing or the case of negative conditional equity premiums that are highly unlikely in standard asset pricing models.

To understand further the difference between the PLS and the BW index, Figure 3 plots the weights of S^{PLS} and S^{BW} on the six individual proxies over the out-of-sample period. Figure 3 shows that the PLS weights vary over time gradually and vary more than the PC weights, while they do vary around the full-sample values. This fact is not surprising, because PLS is a target-driven approach and it extracts information according to the covariance with the forecast target. On the contrary, the PC only picks the weights that track the volatility of the proxies. By design, the stability of the PLS weights depends on the forecasting target's variation, in addition to the variations of the individual sentiment proxies. Because the excess market return, the target here, is volatile (see Table 1 and Figure 2), the PLS weights should be less persistent than the BW index weights. This provides an additional intuitive reason why S^{PLS} outperforms S^{BW} in- and out-of-sample, in that S^{PLS} incorporates the changing market dynamics more timely.

Panel B of Table 5 shows the out-of-sample performance for the six individual sentiment proxies. Three out of six generate positive R^2_{OS} statistics, but only two, RIPO and EQTI, are significant according to the DM- and CW-tests. Because RIPO and EQTI are also the only two variables that generate significant in-sample predictability, we are interested in whether their weights on S^{PLS} are persistently large over time. The positive answer to this question provides supportive evidence why S^{PLS} outperforms S^{BW} in and out of sample. In Figure 3, except for the first several months, the weights of RIPO and EQTI in S^{PLS} are always larger than 0.50, and all the rest are less than 0.3 in absolute value, suggesting that RIPO and EQTI are two dominant proxies in constructing S^{PLS} because of their significant forecasting power.

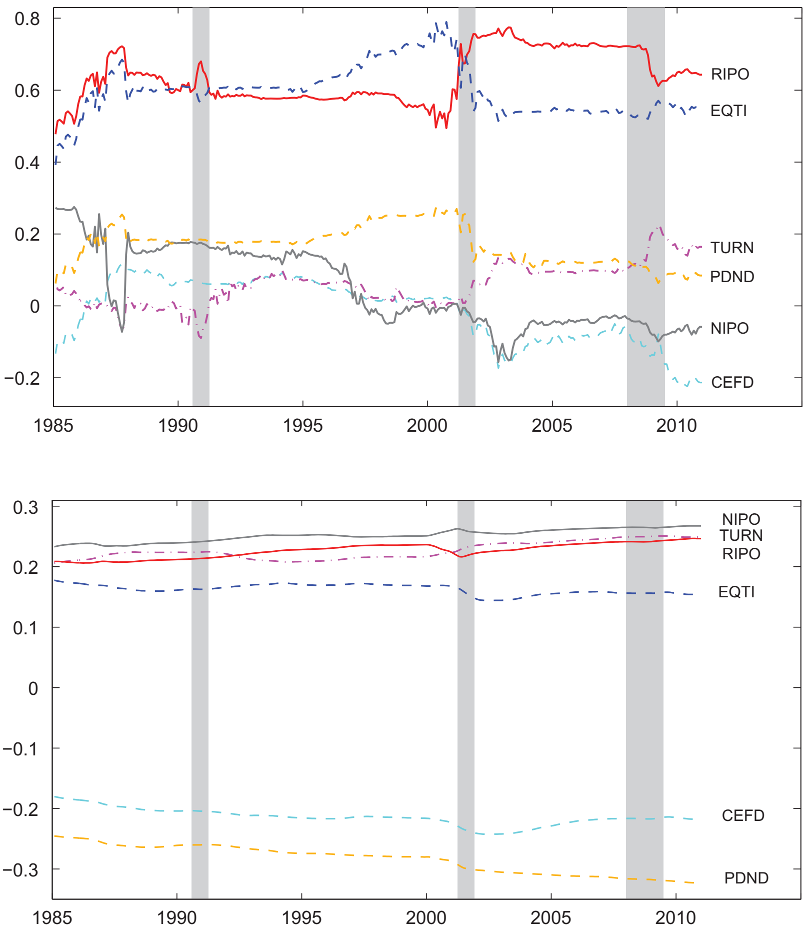

Figure 3
Weights of S^{PLS} and S^{BW} on individual investor sentiment proxies, January 1985 to December 2010
The upper panel depicts the weights of the six sentiment proxies for the recursively constructed aligned investor sentiment index S^{PLS}. The lower panel depicts the weights of the six sentiment proxies for the recursively constructed Baker and Wurgler's sentiment index S^{PC}. The index weights are estimated recursively based on information up to the period of forecast formation period t alone based on PLS and PC methods separately for S^{PLS} and S^{PC}. The six individual sentiment proxies are the close-end fund discount rate (CEFD), share turnover (TURN), number of IPOs (NIPO), first-day returns of IPOs (RIPO), dividend premium (PDND), and equity share in new issues (EQTI). The vertical bars correspond to NBER-dated recessions.

For comparison, we also estimate recursively the kitchen sink model and evaluate its out-of-sample performance. The bottom row of Panel B reports the results. The kitchen sink model generates a positive R^2_{OS} of 0.27%, with significance based on the CW- but not DW-test. Usually the kitchen sink performs badly with many predictors. In our case here, there are only six predictors, and so its performance is not as bad as it often is in other applications.

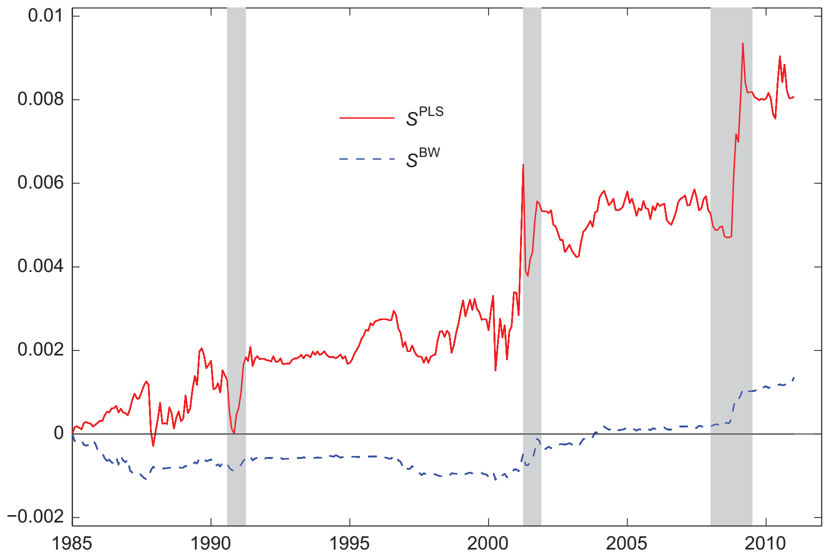

Figure 4
The difference in cumulative squared forecast error (CSFE), January 1985 to December 2010
The solid line depicts the difference between the cumulative squared forecast error (CSFE) for the historical average benchmark and the CSFE for the out-of-sample predictive regression forecast based on the recursively constructed aligned investor sentiment index S^{PLS}. The dashed line depicts the difference between the CSFE for the historical average benchmark and the CSFE for the out-of-sample predictive regression forecast based on the recursively constructed Baker and Wurgler's sentiment index S^{BW}. The sentiment indices and regression coefficients are estimated recursively based on information up to the period of forecast formation period t alone. The vertical bars correspond to NBER-dated recessions.

Nevertheless, the R^2_{OS} of 0.27% is lower than that of S^{Com}, and is substantially lower than the R^2_{OS} of S^{PLS}. The weak performance of the kitchen sink model is consistent with Goyal and Welch (2008) and Rapach, Strauss, and Zhou (2010), in that while the kitchen sink model may have good in-sample forecasting power, its out-of-sample performance tends to be worse than that of the simple combination forecast.

Following Goyal and Welch (2008) and Rapach, Strauss, and Zhou (2010), Figure 4 presents the time-series plots of the differences between the cumulative squared forecast error (CSFE) for the historical average benchmark forecast and the CSFE for the predictive regression forecasts based on investor sentiment indices S^{PLS} and S^{BW} over January 1985 through December 2010, where the proxy weights for each index are given in Figure 3. This time-series plot is an informative graphical device on the consistency of out-of-sample forecasting performance over time. When the difference in CSFE increases, the model forecast outperforms the historical average, whereas the opposite holds when the curve decreases. It thus illustrates whether an investor sentiment-based forecast has a lower MSFE than the historical average for any particular out-of-sample period.

The solid line in Figure 4 shows that our aligned investor sentiment index S^{PLS} consistently outperforms the historical average. The curve has slopes that are predominantly positive, indicating that the good out-of-sample performance of S^{PLS} steps from the whole sample period rather than some special episodes. The figure also graphically illustrates the performances over the NBER-dated business cycles, complementing Table 5. For comparison, the dashed line plots the difference in CSFE for the BW index. The dashed line shows that S^{BW} fails to consistently outperform the historical average. The curve is positively sloped in the 2000s, but it is negatively sloped over the extended periods from the mid 1980s to 1990s. Overall, Figure 4 shows that S^{PLS} is a powerful and reliable predictor for the excess market returns, and it consistently outperforms S^{BW} across different sample periods.

In summary, this subsection shows that the aligned investor sentiment S^{PLS} displays strong out-of-sample forecasting power for the aggregate stock market. In addition, S^{PLS} substantially outperforms the BW index S^{BW}, the naive index S^{EW}, the simple combination forecast S^{Com}, the kitchen sink model, and all the individual sentiment proxies in an out-of-sample setting, consistent with our previous in-sample results (Tables 2–4).

3.4 Predictability with longer horizons

Although the focus of our paper is on the predictability of investor sentiment over the monthly horizon, in this subsection we investigate its forecasting power over longer horizons. Because investor sentiment is persistent, intuitively it may have a long-run effect on the stock market as well. In addition, because of the limits of arbitrage, mispricings from investor sentiment may not be eliminated completely by arbitrageurs over a short horizon. In the literature, there is some in-sample evidence on the long-run predictability of investor sentiment. For example, using some survey data, Brown and Cliff (2004, 2005) show that the predictive power of investor sentiment is significant in the long run (over 1 year) but insignificant in the short run (less than 1 year). Baker, Wurgler, and Yuan (2012) show that global sentiment in year $t-1$ predicts significantly the following 12-month country-level market returns over 1980–2005, on the basis of a pooling regression.

Table 6 reports the in- and out-of-sample forecasting results of S^{PLS} on the excess market return over long horizons. For comparison, we also show the results with the BW index. Three observations follow the table. First, S^{PLS} can significantly predict the long-run excess market returns up to 12 months. The forecasting power increases as the horizon increases and then declines, in sample and out of sample. More specifically, the in-sample forecasting power peaks at 9 months and the out-of-sample forecasting power peaks at 12 months. Second, the predictive power of S^{BW} for long-horizon excess market return is small and insignificant in general. Third, over the one-year (12 months) horizon, our results are generally consistent with Brown and Cliff (2005), although we

Table 6
Investor sentiment and long-horizon predictability

	Aligned investor sentiment index, S^{PLS}				BW investor sentiment index, S^{BW}			
Horizon	β (%)	t-stat	R^2 (%)	R^2_{OS} (%)	β (%)	t-stat	R^2 (%)	R^2_{OS} (%)
1 month	−0.58***	−3.04	1.70	1.23**	−0.24	−1.21	0.30	0.15
3 month	−1.57***	−3.64	3.90	2.75**	−0.62	−1.21	0.61	0.43
6 month	−2.84***	−3.54	5.99	3.63**	−1.23	−1.18	1.13	0.46
9 month	−3.58**	−2.86	6.24	3.72*	−1.59	−1.05	1.24	0.07
12 month	−4.09**	−2.40	6.11	4.55*	−1.74	−0.89	1.11	-0.14
24 month	−4.33	−1.41	3.76	2.77	−0.24	−0.08	0.01	-0.17

This table reports the in- and out-of-sample long-horizon forecasting results for the excess market return with lagged investor sentiment,

$$R^m_{t\to t+h}=\alpha+\beta S^k_t+\epsilon_{t\to t+h}, \qquad k=\mathrm{PLS},\mathrm{BW},$$

where $R^m_{t\to t+h}$ is the h-month ahead excess market return from t to $t+h$, S^{BW}_t is the Baker-Wurgler investor sentiment index in month t, and S^{PLS}_t is the aligned investor sentiment index. We report the regression slopes, Newey-West t-statistic, in-sample R^2 and out-of-sample R^2_{OS}. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The in- and out-of-sample periods are over July 1965 to December 2010 and January 1985 to December 2010, respectively.

use different measures for investor sentiment. For example, for a one-standard-deviation positive shock to sentiment, Brown and Cliff's (2005) sentiment predicts a 3% decrease in the aggregate stock market return over the next one year, while our aligned sentiment index S^{PLS} predicts a decrease of 4%, quantitatively similar.

In sum, the aligned sentiment index S^{PLS} strengthens substantially the predictability of the BW index, not only at the monthly frequency but also over longer horizons beyond one month. It significantly predicts the market returns from one month up to one year into the future, both in and out of sample.

3.5 Asset allocation implications

Now we examine the economic value of stock market forecasts based on the aligned investor sentiment index S^{PLS}. Following Kandel and Stambaugh (1996), Campbell and Thompson (2008), and Ferreira and Santa-Clara (2011), among others, we compute the certainty equivalent return (CER) gain and Sharpe Ratio for a mean-variance investor who optimally allocates across equities and the risk-free asset using the out-of-sample predictive regression forecasts.

At the end of period t, the investor optimally allocates

$$w_t=\frac{1}{\gamma}\frac{\hat{R}^{ms}_{t+1}}{\hat{\sigma}^2_{t+1}} \tag{19}$$

of the portfolio to equities during period $t+1$, where γ is the risk aversion coefficient, $\hat{R}^{ms}_{t+1}$ is the out-of-sample forecast of the simple excess market return, and $\hat{\sigma}^2_{t+1}$ is the variance forecast. The investor then allocates $1-w_t$ of the portfolio to risk-free bills, and the $t+1$ realized portfolio return is

$$R^p_{t+1}=w_t R^{ms}_{t+1}+R^f_{t+1}, \tag{20}$$

where R_{t+1}^{f} is the gross risk-free return. Following Campbell and Thompson (2008), we assume that the investor uses a five-year moving window of past monthly returns to estimate the variance of the excess market return and constrains w_t to lie between 0 and 1.5 to exclude short sales and to allow for at most 50% leverage.

The CER of the portfolio is

$$\mathrm{CER}_p = \hat{\mu}_p - 0.5\gamma\hat{\sigma}_p^2, \tag{21}$$

where $\hat{\mu}_n$ and $\hat{\sigma}_n^2$ are the sample mean and variance, respectively, for the investor's portfolio over the q forecasting evaluation periods. The CER gain is the difference between the CER for the investor who uses a predictive regression forecast of market return generated by Equation (16) and the CER for an investor who uses the historical average forecast, Equation (18). We multiply this difference by 12 so that it can be interpreted as the annual portfolio management fee that an investor would be willing to pay to have access to the predictive regression forecast instead of the historical average forecast. To examine the effect of risk aversion, we consider portfolio rules based on risk-aversion coefficients of 1, 3, and 5, respectively. In addition, we also consider the case of 50bps transaction costs, which is generally considered as a relatively high number.

For assessing the statistical significance, following DeMiguel, Garlappi, and Uppal (2009), we test whether the CER gain is indistinguishable form zero by applying the standard asymptotic theory as in their paper. In addition, we also calculate the monthly Sharpe ratio of the portfolio, which is the mean portfolio return in excess of the risk-free rate divided by the standard deviation of the excess portfolio return. Following again DeMiguel, Garlappi, and Uppal (2009), we use the approach of Jobson and Korkie (1981) corrected by Memmel (2003) to test whether the Sharpe ratio of the portfolio strategy based on predictive regression is statistically indifferent from that of the portfolio strategy based on historical average.

Table 7 shows that the BW index S^{BW} generates small economic gains for a mean-variance investor, consistent with the small R_{OS}^2 statistics in Table 5. Specifically, S^{BW} has a negative CER gain of −0.78% when the risk aversion is 1, and small positive CER gains of 0.75% and 0.53%, when the risk aversions are 3 and 5, respectively. The net-of-transactions-costs CER gains for S^{BW} is even lower, ranging from −0.83% to 0.70%. The Sharpe ratios of S^{BW} range from 0.09 to 0.11 under alternative risk aversions. S^{EW} performs slightly better than S^{BW} with CER gains varying from −0.60% to 1.3%, and the net-of-transactions-costs CER gains varying from −0.67% to 1.23%. Consistent with Rapach, Strauss, and Zhou (2010), the combination forecast S^{Com} performs reasonably well. All the CER gains of S^{Com} are positive, ranging from 0.80% to 1.56%; and the Sharpe ratios lie in the range of 0.10 to 0.13.

Of all the sentiment indices, S^{PLS} stands out again in term of the economic value. The CER gains for S^{PLS} across the risk aversions are consistently positive

Table 7
Asset allocation results

	No transaction cost		50pbs transaction cost	
Predictor	CER gain (%)	Sharpe ratio	CER gain (%)	Sharpe ratio
Panel A: Risk aversion $\gamma=1$				
S^{BW}	−0.78	0.11	−0.83	0.10
S^{EW}	−0.60	0.11	−0.67	0.11
S^{Com}	0.80*	0.13	0.69	0.13
S^{PLS}	4.39**	0.18**	4.17**	0.16**
ECON^{PLS}	0.61	0.12	−0.41	0.10
$(S\text{+ECON})^{PLS}$	1.89	0.14	1.14	0.13
Panel B: Risk aversion $\gamma=3$				
S^{BW}	0.75*	0.10**	0.70*	0.09*
S^{EW}	1.30**	0.11**	1.23**	0.10**
S^{Com}	1.56***	0.12**	1.46***	0.12***
S^{PLS}	4.14***	0.18***	3.82**	0.17**
ECON^{PLS}	1.55	0.14	0.44	0.11
$(S\text{+ECON})^{PLS}$	3.22*	0.17*	2.11	0.15
Panel C: Risk aversion $\gamma=5$				
S^{BW}	0.53*	0.09*	0.52*	0.08*
S^{EW}	0.91**	0.09***	0.87**	0.08**
S^{Com}	0.82**	0.10**	0.74**	0.10**
S^{PLS}	2.47**	0.18**	2.08**	0.17**
ECON^{PLS}	−0.32	0.13	−1.16	0.10
$(S\text{+ECON})^{PLS}$	1.43	0.17	1.42	0.15

This table reports the portfolio performance measures for a mean-variance investor with a risk-aversion coefficient (γ) of 1, 3, and 5, respectively, who allocates monthly between equities and risk-free bills using the out-of-sample forecasts of the excess market returns based on lagged investor sentiment. S^{BW} is the Baker-Wurgler sentiment index, S^{EW} is the equally-weighted naive sentiment index, S^{Com} is the combination forecast, and S^{PLS} is the aligned sentiment index. ECON^{PLS} and $(S\text{+ECON})^{PLS}$ are the two PLS predictors extracted from the 18 economic variables and the union of 18 economic variables and six individual sentiment proxies, respectively. CER gain is the annualized certainty equivalent return gain for the investor, and the monthly Sharpe ratio is the mean portfolio return in excess of the risk-free rate divided by its standard deviation. The portfolio weights are estimated recursively, using the data available at the forecast formation time t. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The out-of-sample evaluation period is over January 1985 to December 2010.

and economically large, ranging from 2.34% to 4.39%. More specifically, an investor with a risk aversion of 1, 3, or 5 would be willing to pay an annual portfolio management fee up to 4.39%, 4.14%, and 2.34%, respectively, to have access to the predictive regression forecast based on S^{PLS} instead of using the historical average forecast. The net-of-transactions-costs CER gains of the S^{PLS} portfolios range from 2.08% to 4.17%, well above those of S^{BW}, S^{EW}, and S^{Com}, and they are of economic significance. The Sharpe ratios of portfolios formed based on S^{PLS} range from 0.15 to 0.19, which more than double the market Sharpe ratio, 0.07, with a buy-and-hold strategy (Table 1). In addition, all the CER gains and Shape ratio gains of S^{PLS} in all of the risk aversion cases are statistically significant.

In the last two lines of each panel of Table 7, we report the portfolio gains of the PLS predictor extracted from the 18 economic variables and their union with the sentiment variables. In accordance with the R^2_{OS}s in Table 5, ECON^{PLS} and $(S\text{+ECON})^{PLS}$ generate only limited economic values for the mean-variance investor in terms of both the CER gain and the Sharpe ratio. Hence, from the asset allocation perspective, S^{PLS} performs the best among all the alternatives.

Overall, Table 7 demonstrates that the aligned investor sentiment S^{PLS} can generate sizable economic value for a mean-variance investor, while S^{BW} cannot. The results are robust to common risk aversion specifications and a common level of transaction cost.

3.6 Forecasting characteristics portfolios

Investor sentiment has different impacts on different stocks. In particular, stocks that are speculative, difficult to value, hard to arbitrage, and in the short leg are likely to be more sensitive to investor sentiment (Baker and Wurgler 2006, 2007; Stambaugh, Yu, and Yuan 2012; Antoniou, Doukas, and Subrahmanyam 2013). In this subsection, we investigate how well the aligned investor sentiment S^{PLS} can forecast portfolios sorted on industry, size, BM, and momentum. This study not only helps to strengthen our previous findings for aggregate stock market predictability, but also helps to enhance our understanding for the economic sources of return predictability.[6]

Consider now the predictive regression,

$$R^j_{t+1} = \alpha_j + \beta_j S^{\mathrm{PLS}}_t + \epsilon^j_{t+1}, \tag{22}$$

where R^j_{t+1} is the monthly excess returns for the 10 industry, 10 size, 10 BM, and 10 momentum portfolios, respectively, with the null hypothesis $H_0 : \beta_j = 0$ against the alternative hypothesis $H_A : \beta_j < 0$ based on wild bootstrapped p-values.

Panel A of Table 8 reports the estimation results for in-sample univariate predictive regressions for 10 industry portfolios with investor sentiment over the period of July 1965 through December 2010.[7] Affirming our findings for the market portfolio in Table 2, S^{PLS} substantially enhances the return forecasting performance relative to S^{BW} across all industries, with the R^2s about two to ten times higher than the corresponding R^2s of S^{BW}.

In addition, almost all of the regression slope estimates for S^{PLS} and S^{BW} are negative; thus, the negative predictability of investor sentiment for subsequent stock returns are pervasive across industry portfolios. The regression slope estimates and R^2 statistics vary significantly across industries, illustrating large cross-sectional difference in the exposures to investor sentiment. Specifically, technology, energy, and telecom industries are the most predictable by investor sentiment, whereas utility, health, and nondurable present the lowest predictability.

6 See, for example, Ferson and Harvey (1991); Baker and Wurgler (2006, 2007); Hong, Torous, and Valkanov (2007); Cohen and Frazzini (2008); and Menzly and Ozbas (2010).

7 Monthly value-weighted returns for portfolios sorted on industry, size, BM ratio, and momentum are available from Kenneth French's data library.

Table 8
Forecasting characteristics portfolios with investor sentiment

	S^{PLS} (%)	t-stat	R^2 (%)	S^{BW} (%)	t-stat	R^2 (%)
Panel A: Industry portfolios						
Nondurable	−0.38	−1.91	0.74	−0.02	−0.08	0.00
Durable	−0.46	−1.82	0.52	−0.13	−0.54	0.04
Manufacture	−0.66**	−3.15	1.70	−0.27	−1.17	0.27
Energy	−0.67**	−2.59	1.47	−0.44**	−1.84	0.64
Technology	−0.95**	−2.90	1.92	−0.72**	−2.22	1.10
Telecom	−0.56**	−2.76	1.35	−0.27*	−1.40	0.33
Shop	−0.43	−1.87	0.64	0.05	0.19	0.01
Health	−0.35	−1.49	0.48	−0.01	−0.03	0.00
Utility	−0.28	−1.52	0.46	−0.11	−0.60	0.07
Other	−0.69**	−2.77	1.55	−0.32	−1.28	0.33
Panel B: Size portfolios						
Small	−1.06***	−3.47	2.54	−0.82***	−2.80	1.52
2	−0.90**	−3.01	1.88	−0.66***	−2.32	1.00
3	−0.89***	−3.29	2.00	−0.57**	−2.07	0.82
4	−0.89***	−3.52	2.16	−0.59***	−2.24	0.95
5	−0.85***	−3.44	2.12	−0.54**	−2.10	0.84
6	−0.82***	−3.50	2.22	−0.50**	−2.04	0.85
7	−0.76***	−3.27	1.97	−0.44**	−1.84	0.68
8	−0.63**	−2.79	1.46	−0.36*	−1.52	0.47
9	−0.64**	−3.09	1.75	−0.29*	−1.38	0.37
Large	−0.56**	−2.89	1.65	−0.22	−1.11	0.26
Panel C: Book-to-market portfolios						
Growth	−0.75**	−2.93	1.97	−0.37*	−1.46	0.49
2	−0.58**	−2.82	1.42	−0.21	−0.98	0.19
3	−0.64***	−3.27	1.78	−0.26	−1.27	0.30
4	−0.57**	−2.74	1.34	−0.28	−1.29	0.32
5	−0.53**	−2.91	1.32	−0.26	−1.33	0.31
6	−0.57**	−2.94	1.51	−0.34**	−1.66	0.53
7	−0.59***	−3.05	1.67	−0.33*	−1.57	0.52
8	−0.54**	−2.74	1.32	−0.31*	−1.51	0.44
9	−0.52**	−2.68	1.13	−0.29	−1.36	0.35
Value	−0.62**	−2.78	1.08	−0.39*	−1.54	0.43
Panel D: Momentum portfolios						
Loser	−1.14***	−3.07	1.92	−0.84**	−2.34	1.06
2	−0.66*	−2.15	1.05	−0.32	−1.09	0.26
3	−0.58*	−2.43	1.12	−0.20	−0.83	0.13
4	−0.53*	−2.41	1.13	−0.20	−0.91	0.17
5	−0.48*	−2.42	1.08	−0.18	−0.89	0.15
6	−0.68***	−3.37	2.10	−0.33*	−1.56	0.50
7	−0.54**	−2.76	1.40	−0.23	−1.16	0.26
8	−0.67***	−3.69	2.11	−0.30*	−1.53	0.43
9	−0.72***	−3.57	2.07	−0.43**	−2.04	0.72
Winner	−1.00***	−3.56	2.43	−0.67***	−2.52	1.10

This table reports in-sample estimation results for predictive regression

$$R^j_{t+1} = \alpha_j + \beta_j S^k_t + \epsilon^j_{t+1}, \qquad k = \text{PLS}, \text{BW},$$

where R^j_{t+1} is the monthly excess returns (in percentage) for the 10 industry, 10 size, 10 book-to-market, and 10 momentum portfolios, respectively. S^{PLS}_t is the aligned investor sentiment index at period t, and S^{BW}_t is the Baker-Wurgler investor sentiment index at period t. We report the slopes, Newey-West t-statistics, as well as the R^2s. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively, based on one-sided wild bootstrapped p-values. Portfolio returns are value-weighted and available from Kenneth French's data library. The sample period is over July 1965 to December 2010.

The remaining panels of Table 8 show that S^{PLS} improves sharply the forecasting performance relative to S^{BW} for the cross-sectional stock returns of size, BM, and momentum portfolios as well. S^{PLS} significantly forecasts all of the 10 characteristic portfolios sorted on size, BM, and past return, respectively, whereas S^{BW} only significantly forecasts 9, 5, and 5 corresponding characteristic portfolios. In addition, all the R^2s of S^{PLS} are much larger than the corresponding R^2s of S^{BW}. For example, the R^2 of S^{PLS} for the largest cap portfolio is 1.65%, whereas the corresponding R^2 of S^{BW} is 0.26%.

Consistent with the literature, there is a fairly large dispersion of regression slope estimates in the cross-section. Stocks that are small, distressed (high BM ratio), with high growth opportunity (low BM ratio), or past losers are more predictable by investor sentiment. Interestingly, among the four groups of portfolios, the slopes on the size portfolios are monotonically increasing in absolute value from large to small firms. Based on the monotonicity test of Patton and Timmermann (2010), we find that the increasing pattern is a true feature of the data that is statistically significant at the 5% level.

4. Economic Explanations

In this section, we explore first the source of predictability at both the market and portfolio levels. Then, we explore the relation of investor sentiment with aggregate volatility, accruals, cash flows, and consumer sentiment.

4.1 Cash flow and discount rate predictability

Valuation models suggest that stock prices are determined by both future expected cash flows and discount rates. From this perspective, the ability of investor sentiment to forecast aggregate stock market returns may come from either the cash flow channel or the discount rate channel or both (Baker and Wurgler 2006, 2007). Hence, it is of interest to investigate this issue.

Fama and French (1989) and Cochrane (2008, 2011), among others, argue that aggregate stock market predictability comes from the time variation in discount rates. Under the discount rate channel, high S^{PLS} predicts low future return because it predicts low discount rates. However, S^{PLS} may represent investors' biased belief about future cash flows not justified by economic fundamentals (Baker and Wurgler 2006, 2007). Since S^{PLS} is a negative predictor for future stock market return, the cash-flow channel implies that the low stock market return following high S^{PLS} reflects the downward correction of overpricing induced by overly optimistic cash-flow forecasts under high investor sentiment, when the true fundamental is revealed in the next period.[8]

[8] The overly optimistic cash-flow forecasts relative to the rational expectation under high sentiment can be driven by various reasons, including overreaction to good cash-flow news owing to over-extrapolation and representativeness bias (Kahneman and Tversky 1974), underreaction to bad cash flow news owing to conservatism bias (Barberis, Shleifer, and Vishny 1998), or cognitive dissonance (Antoniou, Doukas,

To test whether the predictability of S^{PLS} is from either or both of the channels, proxies of the channels are needed. We use the aggregate dividend-price ratio as our discount rate proxy, since the time variation in aggregate dividend-price ratio is primarily driven by discount rates (Cochrane 2008, 2011). We use aggregate dividend growth as our primary cash-flow proxy, which is widely examined and used in similar studies in the literature (e.g., Campbell and Shiller 1988; Fama and French 2000; Menzly, Santos, and Veronesi 2004; Lettau and Ludvigson 2005; Cochrane 2008, 2011; Binsbergen and Koijen 2010; Kelly and Pruitt 2013; Huang et al. 2014a). Considering that Fama and French (2001) document a steep-downward trend in the fraction of U.S. firms paying dividends, and that the dividends are subject to smoothing, we also examine two alternative aggregate cash-flow proxies, including the aggregate earning growth and real GDP growth rate,[9] in addition to dividend growth.

The Campbell and Shiller (1988) log linearization of stock return generates an approximate identity, as argued in Cochrane (2008, 2011) and Campbell, Polk, and Vuolteenaho (2010),

$$R_{t+1} = k + \mathrm{DG}_{t+1} - \rho \mathrm{D/P}_{t+1} + \mathrm{D/P}_t, \tag{23}$$

where R_{t+1} is the aggregate stock market return from t to $t+1$, DG_{t+1} is the log aggregate dividend-growth rate, $\mathrm{D/P}_{t+1}$ is the log aggregate dividend-price ratio, and ρ is a positive log-linearization constant. Equation (23) implies that if S_t^{PLS} predicts next period market return R_{t+1} beyond the information contained in $\mathrm{D/P}_t$, it must predict either DG_{t+1} or $\mathrm{D/P}_{t+1}$ (or both). Since DG_{t+1} and $\mathrm{D/P}_{t+1}$ represent separately cash flows and discount rates in our setting, the forecasting power of S_t^{PLS} for DG_{t+1} and $\mathrm{D/P}_{t+1}$ would point to the cash-flow predictability channel and discount-rate predictability channel, respectively.

Therefore, our study focuses on the following bivariate predictive regressions,

$$Y_{t+1} = \alpha + \beta S_t^{\mathrm{PLS}} + \psi \mathrm{D/P}_t + \upsilon_{t+1}, \qquad Y = \mathrm{D/P}, \mathrm{DG}, \mathrm{EG}, \mathrm{GDPG}, \tag{24}$$

where $\mathrm{D/P}_{t+1}$ is the log dividend-price ratio on the S&P 500 index at the end of year $t+1$, DG_{t+1} is the annual log dividen-growth rate on the S&P 500 index from year t to $t+1$, EG_{t+1} is the annual log earning growth rate on the S&P 500 index from year t to $t+1$, GDPG_{t+1} is the annual log real GDP growth rate from year t to $t+1$, S_t^{PLS} is the aligned investor sentiment index at the end of year t, and υ_{t+1} is the noise term. Following the literature, we use annual data in the above regressions to avoid spurious predictability arising from within-year seasonality. We construct $\mathrm{D/P}_{t+1}$ and DG_{t+1} according

and Subrahmanyam 2013), gradual information diffusion (Hong and Stein 1999), and Bayesian learning (Timmermann 1993, 1996; Lewellen and Shanken 2002), among others.

[9] In addition, we have examined the aggregate cash flow in Hirshleifer, Hou, and Teoh (2009) and the industrial production growth, as alternative cash flow measures, and have found similar results.

Table 9
Forecasting cash flows and discount rates with investor sentiment

Y_{t+1}	β	t-stat	ψ	t-stat	R^2 (%)
Panel A: Aligned investor sentiment, S^{PLS}					
D/P	−0.00	−0.09	0.95***	19.3	89.8
DG (%)	−3.46*	−2.35	3.55	0.73	10.3
EG (%)	−9.03**	−2.53	1.88	0.15	6.88
GDPG (%)	−0.57*	−1.99	0.33	0.39	7.53
Panel B: BW investor sentiment, S^{BW}					
D/P	−0.01	−0.55	0.95***	19.6	89.9
DG (%)	−2.02	−1.29	4.71	0.97	5.51
EG (%)	−7.12*	−2.43	4.10	0.32	4.53
GDPG (%)	−0.47	−1.42	0.46	0.56	5.72

This table reports in-sample estimation results for the bivariate predictive regressions

$$Y_{t+1} = \alpha + \beta S_t^k + \psi \mathrm{D/P}_t + \upsilon_{t+1}, \qquad Y = \mathrm{D/P}, \mathrm{DG}, \mathrm{EG}, \mathrm{GDPG}, \qquad k = \mathrm{PLS}, \mathrm{BW},$$

where $\mathrm{D/P}_{t+1}$ is the log dividend-price ratio on the S&P 500 index at the end of year $t+1$, DG_{t+1} is the annual log dividend-growth rate on the S&P 500 index from year t to $t+1$ (in percentage), EG_{t+1} is the annual log earning growth rate on the S&P 500 index from year t to $t+1$ (in percentage), GDPG_{t+1} is the annual log real GDP growth rate from year t to $t+1$ (in percentage), S_t^{PLS} is the aligned investor sentiment index at the end of year t, and S_t^{BW} is the Baker-Wurgler investor sentiment index at the end of year t. DG_{t+1} and $\mathrm{D/P}_{t+1}$ are constructed following Cochrane (2008, 2011). We report the regression slopes, Newey-West t-statistics, as well as R^2s. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively, based on one-sided wild bootstrapped p-values. The sample period is over 1965–2011.

to Cochrane (2008, 2011), based on total market returns and market returns without dividends. The sample period is from 1965 to 2011.

Panel A of Table 9 reports the results. S^{PLS} displays distinct patterns for cash-flow and discount-rate predictability. The slope of S^{PLS} for $\mathrm{D/P}_{t+1}$ is virtually equal to zero and statistically insignificant.[10] However, the slope estimate of S^{PLS} for DG_{t+1} in predictive regression, Equation (24), is −3.46, with statistical significance at the 10% level based on the one-sided wild bootstrapped p-value. From Equation (23), the significant negative predictability of S^{PLS} for DG_{t+1} and no predictability for $\mathrm{D/P}_{t+1}$ jointly indicate that S^{PLS} should present significantly negative predictive power for excess market return, which is in accord with the evidence of negative market return predictability of S^{PLS} in Tables 2 and 6. Moreover, Panel A shows that S^{PLS} also displays significant ability in predicting alternative cash-flow proxies, such as EG_{t+1} and GDPG_{t+1}, suggesting that our evidence on cash-flow channel is unlikely driven by the changes in payout policies or dividend smoothing.

Panel A also shows that the lagged dividend-price ratio $\mathrm{D/P}_t$ has strong forecasting power for future dividend-price ratio $\mathrm{D/P}_{t+1}$, with a slow mean reverting coefficient of 0.95, whereas its forecasting power for dividend-growth DG_{t+1} is statistically insignificant. This result is consistent with Cochrane (2008, 2011) that the dividend-price ratio captures the time variation in discount rates.

10 We obtain similar results when controlling the lagged dividend growth DG_t.

For comparison, Panel B of Table 9 reports the corresponding results of using BW index S^{BW} in place of S^{PLS}. The slopes of S^{BW} on $\mathrm{D/P}_{t+1}$, DG_{t+1}, and GDPG_{t+1} are not statistically significant, whereas S^{BW} has marginally significant predictive power for EG_{t+1}. This is generally consistent with the early evidence of insignificant market return predictability of S^{BW}.

In summary, the strong negative predictability of S^{PLS} for DG_{t+1}, EG_{t+1}, GDPG_{t+1}, and weak predictability for $\mathrm{D/P}_{t+1}$ in Table 9 indicate that the negative-return predictability of S^{PLS} for aggregate stock market return is coming from the cash-flow channel, different from the popular time-varying discount-rate interpretation of market-return predictability in the literature.[11] Specifically, Table 9 shows that high sentiment predicts low future aggregate cash flows. Our findings hence suggest that high sentiment causes the overvaluation of aggregate stock market because of investors' overly optimistic belief about future aggregate cash flows. When low cash flows are revealed to investors gradually, the overvaluation will diminish and stock price will fall, leading to low future aggregate stock return on average, consistent with the discussion in Baker and Wurgler (2007).

4.2 The cross-section of cash-flow channel

In order to further elucidate the economic source of the predictability of investor sentiment, we extend our analysis to cross-section at the portfolio level. Baker and Wurgler (2006, 2007) find that stock returns that are speculative and hard to arbitrage are more predictable by investor sentiment. Thus, if the predictability of investor sentiment comes from the cash-flow channel, it should have stronger forecasting power for the cash flows of speculative and hard-to-arbitrage stocks as well. This analysis complements the cash-flow channel explanation of investor sentiment's return predictability discussed in Section 4.1.

We examine the predictive ability of S^{PLS} for the cross-section of cash flows using the predictive regression,

$$DG_{t+1}^{j} = \alpha_j + \phi_j S_t^{\mathrm{PLS}} + \vartheta_{t+1}^{j}, \tag{25}$$

where DG_{t+1}^{j} is annual log dividend-growth rate for one of the characteristic portfolios examined in Table 7. We are interested in the predictive regression slope ϕ_j on S^{PLS} in Equation (25), which measures the ability of investor sentiment to forecast cash flows in the cross-section.

We then test whether the ability of investor sentiment to forecast stock returns is positively associated with its ability to forecast cash flows in the cross-section.

[11] Campbell and Ammer (1993), Chen and Zhao (2009), and Campbell, Polk, and Vuolteenaho (2010) argue that because the nominal cash flows of government bonds are fixed, any government bond return predictability should be driven by time-varying discount rates alone. Thus, government bonds provide a clean discount rate proxy without any modeling assumption and variable choice. In untabulated results, we find that S^{PLS} does not have any forecasting power for monthly excess returns of government bonds with maturities from less than 1 to 10 years.

We use a cross-sectional regression to statistically test the cash-flow channel, in the spirit of Hong, Torous, and Valkanov (2007). If the hypothesis holds, firms that are more predictable by investor sentiment should have higher cash-flow predictability as well. We run the cross-section regression

$$\beta_j = a + g\phi_j + e_j, \tag{26}$$

where ϕ_j is from Equation (25) that measures the ability of investor sentiment to forecast the cross-sectional cash flows, and β_j is from Equation (22) measuring the ability of investor sentiment to forecast the cross-section of stock returns (annualized by multiplying 12). If the cash-flow channel hypothesis holds, we expect a positive relationship between β_j and ϕ_j; that is, $g > 0$.

Panel A of Table 10 shows that firms with higher return exposures to investor sentiment have significantly higher cash-flow exposures to investor sentiment. For example, for the 10 size portfolios, the OLS estimate of g for S^{PLS} in Equation (26) is 0.83, with a Newey-West t-statistic of 12.3, indicating a significantly positive relationship between β_j and ϕ_j. Thus, small firms that are more predictable by S^{PLS} with larger negative β_j have significantly higher cash-flow predictability by S^{PLS}, with larger ϕ_j as well. In Table 10, we also document qualitatively similar patterns for S^{BW}, indicating that the cash-flow channel helps to understand the strong cross-sectional predictability of S^{BW} as well.

To delve deeper into the forecasting channel, Panel B of Table 10 shows the regression results of Equation (25). S^{PLS} is a significant negative predictor of cash flows, DG_{t+1}^{j}, for all the 10 size portfolios, consistent with our aggregate market evidence in Table 9. Most importantly, we find an interesting cross-sectional pattern: the cash flows of more speculative and hard-to-arbitrage portfolios are much more predictable by investor sentiment. For example, the R^2 increases from 7.2% for the largest size portfolio, to 17.8% for the smallest size portfolio, which is usually regarded as more speculative and harder to arbitrage; and the regression coefficient ϕ_j decreases sharply from -3.5% for the largest size portfolio, to -10.2% for the smallest size portfolio. This pattern implies that a one-standard-deviation increase in S^{PLS} is associated with a -3.5% decrease in expected dividend growth for large firms and a -10.2% decrease for small firms in the next year, suggesting that the cash flows of small size portfolios are about three times more predictable than those of large size portfolios. Statistically, based on the monotonicity test of Patton and Timmermann (2010), we find that this monotonic relationship in predicting the cash flows is genuinely there, at the usual 5% significance level.

4.3 Market-volatility risk

In this subsection, we examine whether the market-volatility risk can explain the stock return predictability of investor sentiment. Merton (1980) and French, Schwert, and Stambaugh (1987) show that lower stock market volatility implies lower market risk, leading to a lower risk premium or discount rate for next

Table 10
Cross-sectional relation between stock-return predictability and dividend-growth predictability with investor sentiment

	Aligned investor sentiment index, S^{PLS}			BW investor sentiment index, S^{BW}		
Panel A: Cross-sectional regression, $\beta_j = a + g\phi_j + e_j$						
	g	t-stat	R^2 (%)	g	t-stat	R^2 (%)
	0.83	12.3	85.3	0.89	10.6	85.4
Panel B: Forecasting dividend growth of size portfolios						
	ϕ_j	t-stat	R^2 (%)	ϕ_j	t-stat	R^2 (%)
Small	−10.20	−2.90	17.8	−9.48	−2.78	15.3
2	−7.42	−2.00	8.7	−6.33	−1.81	6.3
3	−6.27	−2.33	7.9	−5.70	−2.20	6.5
4	−8.04	−3.19	12.1	−5.19	−2.05	5.5
5	−6.57	−2.88	11.3	−3.80	−1.66	3.7
6	−5.32	−2.89	12.2	−3.83	−1.78	6.3
7	−6.34	−2.90	11.2	−4.08	−1.88	4.9
8	−5.16	−2.87	7.3	−3.63	−1.84	3.6
9	−3.61	−2.25	6.2	−2.04	−1.16	2.0
Large	−3.50	−2.30	7.2	−2.01	−1.34	2.4

Panel A reports the estimation results for the cross-sectional linear regression

$$\beta_j = a + g\phi_j + e_j,$$

where β_j is the following predictive regression slope coefficient of size portfolio j's annualized excess return on investor sentiment (in Panel B of Table 8),

$$R^j_{t+1} = \alpha_j + \beta_j S^k_t + \epsilon^j_{t+1}, \qquad k = \mathrm{PLS},\ \mathrm{BW}, \qquad j = 1, \ldots, 10;$$

and ϕ_j is the following predictive regression slope coefficient of size portfolio j's dividend growth rate DG^j_{t+1} on investor sentiment,

$$DG^j_{t+1} = \alpha_j + \phi_j S^k_t + \vartheta^j_{t+1}, \qquad k = \mathrm{PLS},\ \mathrm{BW}, \qquad j = 1, \ldots, 10.$$

S^{PLS} is the aligned investor sentiment index, S^{BW} is the Baker-Wurgler investor sentiment index, DG^j_{t+1} is the annual dividend-growth rate for size portfolio j constructed following Cochrane (2008, 2011). We report the regression slope coefficients, Newey-West t-statistics, as well as R^2s.

period. It is thus possible that the predictability of S^{PLS} is due to the fact that S^{PLS} represents time variation in expected stock market volatility.

We estimate the following predictive regression

$$\mathrm{LVOL}_{t+1} = \alpha + \beta S^{\mathrm{PLS}}_t + \psi \mathrm{LVOL}_t + \nu_{t+1}, \tag{27}$$

where $\mathrm{LVOL}_{t+1} = \log(\sqrt{\mathrm{SVAR}_{t+1}})$ is the monthly aggregate stock market volatility at period $t+1$. The monthly aggregate stock market variance SVAR_{t+1} is the sum of squared daily returns on the S&P 500 index at monthly frequency,

$$\mathrm{SVAR}_{t+1} = \sum_{i=1}^{N_{t+1}} R^2_{i,t+1}, \tag{28}$$

where N_{t+1} is the number of trading days during period $t+1$, and $R_{i,t+1}$ is the daily excess return for the S&P 500 index on the ith trading day of period $t+1$ (e.g., French, Schwert, and Stambaugh 1987; Schwert 1989; Paye 2012).[12]

We are interested in the slope β on S^{PLS} in Equation (27). Given that S^{PLS} is negatively associated with future aggregate stock market return in Tables 2 and 4, the volatility risk-based argument implies that high S^{PLS} should predict low aggregate stock market volatility and thus low market risk, which in turn decreases the equity risk premium (discount rate). However, in an unreported table, we find find that S^{PLS} indeed displays positive forecasting power for the market volatility, with a $\beta = 0.028$ and a Newey-West t-statistic of 2.10, inconsistent with the volatility hypothesis.

In summary, even though we cannot fully rule out the risk-based explanation, it seems unlikely that market risk is driving the predictive power of S^{PLS} for stock market returns. To the extent that high investor sentiment proxies for more noise trading, our findings appear to provide further support for the behavioral explanation of De Long et al. (1990) that high noise trading leads to excessive volatility.[13]

4.4 Relation with alternative behavioral variables

Many studies provide evidence that behavioral biases can generate misvaluation and return predictability. For example, Hirshleifer and Teoh (2003); and Hirshleifer, Lim, and Teoh (2009), among others, show that investor attention is a limited cognitive resource, so prices do not fully and immediately reflect relevant public information. Hong and Stein (1999); Hong, Torous, and Valkanov (2007); Cohen and Frazzini (2008); Menzly and Ozbas (2010); Huang et al. (2014b); and others show that fundamental information diffuses gradually in the stock market because of market frictions and bounded rationality. Thus, it is interesting to compare the aligned investor sentiment S^{PLS} with alternative return predictors that are related to behavioral bias.

We examine three such alternative behavior predictors. First, we compare S^{PLS} with the consumer-sentiment index published by the Thomson Reuters/University of Michigan. In contrast to S^{PLS}, which is based on market sentiment proxies, the Michigan Consumer Sentiment Index is based on a large number of survey responses to queries about households' current and expected financial conditions. In practice, the Michigan Consumer Sentiment Index is reported regularly in the media, along with commentary on its significance for the economy and financial markets. The index has been used to predict household spending activity (e.g., Ludvigson 2004) as well as small-stock

[12] Because stock market volatility is highly persistent, we, following Andersen et al. (2001) and Paye (2012), include lagged volatility LVOL_t as a control variable in Equation (27) to examine the incremental forecasting power of investor sentiment for aggregate stock market volatility.

[13] Antweiler and Frank (2004) also find that higher sentiment, proxied by the number of messages posted and the bullishness messages posted on the Yahoo Finance and Raging Bull stock message boards, predicts higher future stock market volatility for a set of individual stocks.

premium as an investor sentiment proxy (e.g., Lemmon and Portniaquina 2006). However, we find here that the Michigan Consumer Sentiment Index fails to forecast significantly the future monthly excess aggregate market returns (the R^2 is 0.01%). Therefore, S^{PLS} strongly outperforms the Michigan Consumer Sentiment Index in forecasting the market.

Second, we analyze the Conference Board Consumer Confidence Index®, another popular survey-based proxy of investor sentiment. However, we find that its predictability is as weak as the University of Michigan Consumer Sentiment Index.[14]

Third, we compare S^{PLS} with the aggregate accruals proposed by Hirshleifer, Hou, and Teoh (2009). Accruals have been widely interpreted as proxies for market misvaluation, or managers' efforts to manipulate earnings and stock prices to induce such misvaluation. Sloan (1996) shows that accruals negatively predict future stock returns, because of investors' fixation on reported earnings and their failure to understand the lower persistence of accruals relative to cash flows. In other words, investors are overly optimistic (pessimistic) about the prospects of firms with high (low) accruals. Hirshleifer, Hou, and Teoh (2009) extend the cross-sectional evidence to the aggregate stock market level, and find that the aggregate stock market can be predicted *positively* by aggregate accruals. With annual data, the in-sample R^2 of the accruals is as large as 20% over 1965–2005.

With their data and procedures, we can obtain the monthly accruals and examine the predictability of the aggregate accruals over horizons. Panel A of Table 11 reports the results. It is interesting that the accruals predictor has an out-of-sample R^2 of 0.47% at the monthly frequency, even better than S^{EW}, though the sample period here is shorter. Over 1- to 12-month forecasting horizons, the accruals predictor has in general increasing and sizable predictive power on the market. Note that the forecasts are computed and evaluated as usual, based on recursive estimations and overlapping monthly data for horizons up to the 6 months, but the 12-month result here is computed by calendar years as in Hirshleifer, Hou, and Teoh (2009), for easier comparison. The slight difference in the R^2 of our paper with theirs at the 12-month frequency is because we use excess market return here, and they use the raw return (the difference in the results is minimum). Overall, our results extend Hirshleifer, Hou, and Teoh (2009) annual result to shorter horizons, showing that their aggregate accruals predictor also has good performances from 1 to 6 months.

Because Hirshleifer, Hou, and Teoh (2009) also propose a novel aggregate cash-flows predictor, it is of interest to see how well it performs too. Their predictor measures the innovations in the difference between earnings and the accruals rather than, say, dividend growth. They find that the aggregate stock market can be predicted *negatively* by the cash-flows predictor. Similar

14 We have also examined the economic policy uncertainty index developed by Baker, Bloom, and Davis (2013) and do not find any predictability either.

Table 11
Comparison of S^{PLS} with aggregate accruals and cash flows

Predictor	β (%)	t-stat	R^2 (%)	R^2_{OS} (%)
Panel A: Aggregate accruals				
1 month	0.47**	2.13	1.15	0.47
3 month	1.32**	2.63	3.02	1.30
6 month	2.45***	2.70	5.09	1.43
12 month	6.82***	3.33	20.1	8.28**
Panel B: Aggregate cash flows				
1 month	−0.41**	−2.14	0.90	0.52*
3 month	−1.51***	−3.04	3.97	2.71*
6 month	−2.70***	−2.82	6.42	5.12**
12 month	−5.23**	−2.42	11.3	11.9***
Panel C: Aligned investor sentiment, S^{PLS}				
1 month	−0.54**	−2.70	1.52	1.09**
3 month	−1.69***	−3.87	4.99	4.09***
6 month	−3.33***	−4.02	9.46	8.93***
12 month	−5.21***	−2.92	11.2	13.1**

The table reports the in- and out-of-sample forecasting results with aggregate accruals (Panel A), aggregate cash flows (Panel B), and aligned investor sentiment index S^{PLS} (Panel C), based on the predictive regression

$$R^m_{t\to t+h} = \alpha + \beta Z_t + \varepsilon_{t\to t+h},$$

where $R^m_{t\to t+h}$ is the h-month ahead excess market return on the S&P 500 index from t to $t+h$, and Z_t is the value-weighted average of firm-level scaled accruals, cash flows, or the aligned investor sentiment index at time t. Annual aggregate accruals and cash flows are constructed as in Hirshleifer, Hou, and Teoh (2009). We report the in-sample regression slopes (β); Newey-West t-statistics; in-sample R^2s; and the Campbell and Thompson (2008) out-of-sample R^2_{OS}s. The significance of out-of-sample forecasting is based on the Clark and West (2007) MSFE-adjusted statistic for testing the null hypothesis that the historical average forecast MSFE is less than or equal to the competing predictive regression forecast MSFE against the one-sided alternative hypothesis. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. We use overlapping monthly sampled data. The in- and out-of-sample periods are over July 1965 to April 2006 and January 1985 to April 2006, respectively.

to accruals predictor, we find that the cash-flows predictor also has good performances from 1 to 6 months. Panel B of Table 11 provides the results. Although the performances of the accruals and cash-flows predictors are quite similar, it seems that the latter has somewhat better out-of-sample predictability.

For comparison, Panel C of Table 11 reports the forecasting results of the S^{PLS} over the same sample period and horizons. At the monthly horizon, S^{PLS} outperforms both the accruals and cash-flows predictors. However, it should be noted that the accruals and cash-flows predictors are solid predictors, because they perform better than do the majority of known economic predictors examined earlier. As the horizon increases from 3 to 12 months, the differences between the predictability of the three predictors diminishes. The correlations of S^{PLS} with accruals and cash-flows are −0.13 and 0.17, respectively. These low levels of correlations suggest that the S^{PLS} and the other two predictors are capturing different driving forces in the aggregate market return, and they are complementary in general.

In summary, out of the three alternative behavior predictors, the accruals predictor of Hirshleifer, Hou, and Teoh (2009) stands out the best, as a solid predictor with good performances from 1 to 12 months. However, S^{PLS}

outperforms it at the monthly frequency. Nevertheless, they are complementary, and their performance differences diminish as the horizon increases.

5. Conclusion

In this paper, we propose a new investor sentiment index aligned for predicting the aggregate stock market return, based on the widely used Baker, and Wurgler's (2006) six proxies and by using the PLS method recently introduced to the finance literature by Kelly and Pruitt (2013). With this new measure, we find that investor sentiment has much greater predictive power for the aggregate stock market than previously thought. In addition, it performs much better than most of the commonly used macroeconomic variables do, and its predictability is both statistically and economically significant. Moreover, the new measure also improves substantially the forecasting power for the cross-section of stock returns formed on industry, size, value, and momentum. Economically, we find that the return predictability of investor sentiment seems to come from investors' biased belief about future cash flows rather than discount rates.

Overall, our empirical results suggest that investor sentiment is important not only cross-sectionally, as established in the literature, but also at the aggregate market level. The success of the aligned investor sentiment is due to the use of the PLS approach that exploits more efficiently the information in the proxies than existing procedures do. Hence, the aligned investor sentiment can achieve substantial improvements in forecasting stock returns either at the aggregate level or at the portfolio level. Because investor sentiment has been widely used to examine a variety of financial issues, the aligned investor sentiment, as a significant improvement of existing measures, may yield a number of future applications.

Appendix

A.1 Detailed Description of Economic Variables

This section describes the 18 economic variables in Tables 1 and 4, which are popular stock return predictors documented in the literature that are directly linked to economic fundamentals and risk aversion.

- Dividend-price ratio (DP): log of a 12-month moving sum of dividends paid on the S&P 500 index minus the log of stock prices (S&P 500 index).
- Dividend yield (DY): difference between the log of dividends and log of lagged prices.
- Earnings-price ratio (EP): difference between the log of earnings on the S&P 500 index and log of prices, where earnings are measured using a one-year moving sum.
- Dividend-payout ratio (DE): difference between the log of dividends and log of earnings on the S&P 500 index.
- Stock return variance (SVAR): sum of squared daily returns on the S&P 500 index.
- Book-to-market ratio (BM): ratio of book value to market value for the Dow Jones Industrial Average.
- Net equity expansion (NTIS): ratio of 12-month moving sums of net issues by NYSE-listed stocks to total end-of-year market capitalization of NYSE stocks.
- Treasury bill rate (TBL): interest rate on a 3-month Treasury bill (secondary market).

- Long-term yield (LTY): long-term government bond yield.
- Long-term return (LTR): return on long-term government bonds.
- Term spread (TMS): difference between the long-term yield and Treasury bill rate.
- Default yield spread (DFY): difference between BAA- and AAA-rated corporate bond yields.
- Default return spread (DFR): difference between long-term corporate bond and long-term government bond returns.
- Inflation (INFL): calculated from the consumer price index for all urban consumers (CPI); following Goyal and Welch (2008), inflation are lagged for two months, relative to stock market return, to account for the delay in consumer price index (CPI) releases.
- Consumption-wealth ratio (CAY): residual of regressing consumption on asset wealth and labor income from Lettau and Ludvigson (2001). The data is from Professor Martin Lettau's webpage.[15]
- Log consumption surplus ratio (CSR): estimated with consumption data from U.S. Bureau of Economic Analysis (BEA) as in Campbell and Cochrane (1999).
- Output gap (OG): deviation of the logarithm of total industrial production from a trend that includes both a linear component and a quadratic component (Cooper and Priestley 2009).
- Kelly and Pruitt book-to-market predictor (BM^{KP}): extracted from 100 BM ratios of size- and value-sorted portfolios with the partial least squares (PLS) approach (Kelly and Pruitt 2013).

A.2 Bootstrap Procedures for Computing Empirical *p*-Values

This section describes the wild bootstrap procedures underlying the empirical p-values. The resampling scheme for the wild bootstrap is based on Cavaliere, Rahbek, and Taylor (2010), which is a multiequation extension of the time-series wild bootstrap.

First, we begin by describing the procedure that generates the wild bootstrapped p-values for the test statistics for the predictive regressions of the excess aggregate market return reported in Tables 2, 4, and 6. The wild bootstrap procedure simulates data under the null of no return predictability. Let

$$\hat{\epsilon}_{t+1} = R^m_{t+1} - (\hat{\alpha} + \sum_{i=1}^{N} \hat{\beta}_i x_{i,t} + \sum_{i=1}^{M} \hat{\psi}_i Z_{i,t}), \tag{A1}$$

where $\hat{\alpha}$, $\hat{\beta}_i$ $(i=1,...,N)$, and $\hat{\psi}_i$ $(i=1,...,M)$ are OLS parameter estimates for the general multiple predictive regression model that includes a constant, N standardized individual investor sentiment proxies of Baker and Wurgler (2006), and M economic variables as regressors.

Following convention, we assume that the predictors in Equation (A1) follow an AR(1) process:[16]

$$x_{i,t+1} = \rho_{i,x,0} + \rho_{i,x,1} x_{i,t} + \varphi_{i,x,t+1}, \quad i=1,...,N, \tag{A2}$$

$$Z_{i,t+1} = \rho_{i,Z,0} + \rho_{i,Z,1} Z_{i,t} + \varphi_{i,Z,t+1}, \quad i=1,...,M. \tag{A3}$$

Define

$$\hat{\varphi}^c_{i,x,t+1} = x_{i,t+1} - \hat{\rho}^c_{i,x,0} - \hat{\rho}^c_{i,x,1} x_{i,t}, \quad i=1,...,N, \tag{A4}$$

$$\hat{\varphi}^c_{i,Z,t+1} = Z_{i,t+1} - \hat{\rho}^c_{i,Z,0} - \hat{\rho}^c_{i,Z,1} Z_{i,t}, \quad i=1,...,M, \tag{A5}$$

---

15 We have also examined the alternative CAY, "cayp" in Goyal and Welch (2008), and found similar results.

16 The popular specification of Equations (A2) and (A3) is only an approximation for nonlinear predictors (such as valuation ratios that have price in the denominator). See Goetzmann and Jorion (1993) for alternative bootstrap approaches that account for the nonlinearity.

where

$$(\hat{\rho}^c_{i,x,0},\ \hat{\rho}^c_{i,x,1}),\ \ i=1,...,N, \tag{A6}$$

and

$$(\hat{\rho}^c_{i,Z,0},\ \hat{\rho}^c_{i,Z,1}),\ \ i=1,...,M, \tag{A7}$$

denote vectors of reduced-bias estimates of the AR(1) parameters in Equations (A2) and (A3), respectively. The reduced-bias estimates of the AR parameters are computed by iterating on the Nicholls and Pope (1988) expression for the analytical bias of the OLS estimates (e.g., Amihud, Hurvich, and Wang 2009).

Based on these AR parameter estimates and fitted residuals, we build up a pseudo sample of observations for the excess aggregate market return, N individual investor sentiment proxies, and M macroeconomic variables under the null hypothesis of no return predictability:

$$\tilde{R}^m_{t+1}=\bar{R}^m+\hat{\epsilon}_{t+1}w_{t+1}, \tag{A8}$$

$$\tilde{x}_{i,t+1}=\hat{\rho}^c_{i,x,0}+\hat{\rho}^c_{i,x,1}\tilde{x}_{i,t}+\hat{\varphi}^c_{i,x,t+1}w_{t+1},\ \ i=1,...,N, \tag{A9}$$

$$\tilde{Z}_{i,t+1}=\hat{\rho}^c_{i,Z,0}+\hat{\rho}^c_{i,Z,1}\tilde{Z}_{i,t}+\hat{\varphi}^c_{i,Z,t+1}w_{t+1},\ \ i=1,...,M, \tag{A10}$$

where $\bar{R}^m$ is the sample mean of R^m_{t+1}, w_{t+1} is a draw from the standard normal distribution, $\tilde{x}_{i,0}=x_{i,0}$ $(i=1,...,N)$, and $\tilde{Z}_{i,0}=Z_{i,0}$ $(i=1,...,M)$.

Our wild bootstrap approach is linear and nonparametric on the joint distribution between residuals of the predictors and that of lagged stock return, in the spirit of Stambaugh (1999). Observe that, we multiply the fitted residuals $\hat{\epsilon}_{t+1}$ in Equation (A8), each $\hat{\varphi}^c_{i,x,t+1}$ in Equation (A9), and each $\hat{\varphi}^c_{i,Z,t+1}$ in Equation (A10) by the same scalar, w_{t+1}, when generating the month-$(t+1)$ pseudo residuals. Therefore, our method not only preserves the contemporaneous cross-dependence between endogenous predictors and lagged returns, but also allows the wild bootstrap to capture the general forms of conditional heteroskedasticity. Employing reduced-bias parameter estimates in Equations (A9) and (A10) helps to further ensure that we adequately capture the persistence in the predictors.

Using the pseudo sample of observations for

$$\{(\tilde{R}^m_{t+1},\tilde{x}_{1,t},...,\tilde{x}_{N,t},\tilde{Z}_{1,t},...,\tilde{Z}_{M,t})\}_{t=0}^{T-1}, \tag{A11}$$

we estimate the slopes and the corresponding Newey-West t-statistics for univariate predictive regressions based on each investor sentiment index in Equation (12), each macroeconomic variable in Equation (14), the bivariate predictive regressions based on aligned investor sentiment and each macroeconomic variable in Equation (15), and long-horizon predictive regressions in Section 3.4 . Note that we compute the aligned investor sentiment index, the look-ahead bias-free aligned sentiment index, the BW investor sentiment index, and naive investor sentiment index in Equations (12) and (15) using the pseudo sample of $\{\tilde{x}_{i,t}\}_{t=0}^{T-1}$ $(i=1,...,N)$ and $\{\tilde{R}^m_{t+1}\}_{t=0}^{T-1}$, so that it accounts for the estimated regressors in the predictive regressions. We store the t-statistics for all of the predictive regressions. Repeating this process 2,000 times yields empirical distributions for each of the t-statistics. For a given t-statistic, the empirical p-value is the proportion of the bootstrapped t-statistics greater (less) than the t-statistic for the original sample.

Second, we modify the previous wild bootstrap procedure to simulate data for the predictive regressions on the characteristics portfolios including the 10 industry, 10 size, 10 BM, and 10 momentum portfolios in Table 8 under the null of no predictability. Let

$$\hat{\epsilon}^j_{t+1}=R^j_{t+1}-(\hat{\alpha}^j+\sum_{i=1}^{N}\hat{\beta}^j_i x_{i,t}),\ \ j=1,...,K, \tag{A12}$$

where $\hat{\alpha}^j$ and $\hat{\beta}^j_i$ $(i=1,...,N)$ are estimated by regressing the excess returns of characteristics portfolio j on a constant and all of the N individual investor sentiment proxies. We continue to assume that $x_{i,t}$ follows an AR(1) process and use Equations (A2), (A4), and (A9). In accord with the null, we build up a pseudo sample of observations for excess returns on the characteristics portfolios

$$\tilde{R}^j_{t+1} = \bar{R}^j + \hat{\epsilon}^j_{t+1} w_{t+1}, \quad i = 1, ..., K. \tag{A13}$$

We use this process to simulate data for each characteristics portfolio j $(j = 1, ..., K)$, and compute the aligned investor sentiment index and BW investor sentiment index using the pseudo sample. We then use the pseudo sample to compute the slopes and the corresponding t-statistics for predictive regressions based on each investor sentiment index. Repeating this process 2,000 times, the empirical p-value is the proportion of the bootstrapped t-statistics greater (less) than the t-statistic for the original sample.

Third, we change the previous wild bootstrap procedure to simulate data for the predictive regressions on the dividend-price ratio, dividend growth, earning growth, and real GDP growth in Table 9 under the null. Let

$$\hat{\upsilon}_{Y,t+1} = Y_{t+1} - (\hat{\alpha}_Y + \sum_{i=1}^{N} \hat{\beta}_{Y,i} x_{i,t} + \hat{\psi} \mathrm{D/P}_t), \quad Y = \mathrm{D/P, DG, EG, GDPG}. \tag{A14}$$

Under the null, we allow for predictive power arising from lagged dividend-price ratio, but not lagged investor sentiment measures. We continue to assume that $x_{i,t}$ follows an AR(1) process and use Equations (A2), (A4), and (A9). We simulate R^m_t using Equations (A1) and (A8). In accord with the null, we build up a pseudo sample of observations for these variables

$$\tilde{Y}_{t+1} = \hat{\alpha}_Y + \hat{\psi} \widetilde{\mathrm{D/P}}_t + \hat{\upsilon}_{Y,t+1} w_{t+1}, \qquad Y = \mathrm{D/P, DG, EG, GDPG}. \tag{A15}$$

We use this process to simulate data for these discount rate and cash-flow proxies, and compute the aligned investor sentiment index and BW investor sentiment index using the pseudo sample. We then use the pseudo sample to compute the slopes and the corresponding t-statistics for bivariate predictive regressions based on each investor sentiment index. Repeating this process 2,000 times, the empirical p-value is the proportion of the bootstrapped t-statistics greater (less) than the t-statistic for the original sample.

Fourth, we alternate the previous wild bootstrap procedure to simulate data for the predictive regressions on the log aggregate stock market volatility in Section 4.3 under the null. Let

$$\hat{\nu}_{t+1} = \mathrm{LVOL}_{t+1} - (\hat{\alpha} + \sum_{i=1}^{N} \hat{\beta}_i x_{i,t} + \hat{\psi} \mathrm{LVOL}_t). \tag{A16}$$

Under the null, we allow for market volatility predictability coming from lagged volatility, but not lagged investor sentiment measures. We continue to assume that $x_{i,t}$ follows an AR(1) process and use Equations (A2), (A4), and (A9). We simulate R^m_t using Equations (A12) and (A13). In accord with the null, we generate a pseudo sample of observations for log market volatility

$$\widetilde{\mathrm{LVOL}}_{t+1} = \hat{\alpha} + \hat{\psi} \widetilde{\mathrm{LVOL}}_t + \hat{\nu}_{t+1} w_{t+1}. \tag{A17}$$

We use this process to simulate data for log market volatility, and compute the aligned investor sentiment index and the BW investor sentiment index using the pseudo sample. We then use the pseudo sample to compute the slopes and the corresponding t-statistics for bivariate predictive regressions based on investor sentiment index. Repeating this process 2,000 times, the empirical p-value is the proportion of the bootstrapped t-statistics greater (less) than the t-statistic for the original sample.

References

Amihud, Y., C. Hurvich, and Y. Wang. 2009. Multiple-predictor regressions: Hypothesis tests. *Review of Financial Studies* 22:413–34.

Andersen, T., T. Bollerslev, F. Diebold, and H. Ebens. 2001. The distribution of realized stock return volatility. *Journal of Financial Economics* 61:43–76.

Ang, A., and G. Bekaert. 2007. Return predictability: Is it there? *Review of Financial Studies* 20:651–707.

Antoniou, C., J. Doukas, and A. Subrahmanyam. 2013. Cognitive dissonance, sentiment, and momentum. *Journal of Financial and Quantitative Analysis* 48:245–75.

Antweiler, W., and M. Frank. 2004. Is all that talk just noise? The information content of internet stock message boards. *Journal of Finance* 59:1259–94.

Baker, M., and J. Wurgler. 2000. The equity share in new issues and aggregate stock returns. *Journal of Finance* 55:2219–57.

———. 2006. Investor sentiment and the cross-section of stock returns. *Journal of Finance* 61:1645–80.

———. 2007. Investor sentiment in the stock market. *Journal of Economic Perspectives* 21:129–52.

———. 2012. Comovement and predictability relationships between bonds and the cross-section of stocks. *Review of Asset Pricing Studies* 2:57–87.

Baker, M., J. Wurgler, and Y. Yuan. 2012. Global, local, and contagious investor sentiment. *Journal of Financial Economics* 104:272–87.

Baker, S., N. Bloom, and S. Davis. 2013. Measuring economic policy uncertainty. Working Paper, Stanford University.

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

Binsbergen, J., and R. Koijen. 2010. Predictive regressions: A present-value approach. *Journal of Finance* 65:1439–71.

Breen, W., L. Glosten, and R. Jagannathan. 1989. Economic significance of predictable variations in stock index returns. *Journal of Finance* 64:1177–89.

Brown, G., and M. Cliff. 2004. Investor sentiment and the near-term stock market. *Journal of Empirical Finance* 11:1–27.

———. 2005. Investor sentiment and asset valuation. *Journal of Business* 78:405–40.

Boudoukh, J., R. Michaely, M. Richardson, and M. Roberts. 2007. On the importance of measuring payout yield: Implications for empirical asset pricing. *Journal of Finance* 62:877–915.

Campbell, J. 1987. Stock returns and the term structure. *Journal of Financial Economics* 18:373–99.

Campbell, J., and J. Ammer. 1993. What moves the stock and bond markets? A variance decomposition for long-term asset returns. *Journal of Finance* 48:3–37.

Campbell, J. and J. Cochrane. 1999. By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy* 107:205–51.

Campbell, J., C. Polk, and T. Vuolteenaho. 2010. Growth or glamour? Fundamentals and systematic risk in stock returns. *Review of Financial Studies* 23:305–44.

Campbell, J., and R. Shiller. 1988. The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies* 1:195–228.

Campbell, J., and S. Thompson. 2008. Predicting the equity premium out of sample: Can anything beat the historical average? *Review of Financial Studies* 21:1509–31.

Cavaliere, G., A. Rahbek, and A. Taylor. 2010. Cointegration rank testing under conditional heteroskedasticity. *Econometric Theory* 26:1719–60.

Chen, L., and X. Zhao. 2009. Return decomposition. *Review of Financial Studies* 22:5213–49.

Clark, T., and K. West. 2007. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics* 138:291–311.

Cochrane, J. 2008. The dog that did not bark: A defense of return predictability. *Review of Financial Studies* 21:1533–75.

———. 2011. Presidential address: Discount rates. *Journal of Finance* 66:1047–1108.

Cohen, L., and A. Frazzini. 2008. Economic links and predictable returns. *Journal of Finance* 63:1977–2011.

Cooper, I., and R. Priestley. 2009. Time-varying risk premiums and the output gap. *Review of Financial Studies* 22:2801–33.

De Long, J., A. Shleifer, L. Summers, and R. Waldmann. 1990. Noise trader risk in financial markets. *Journal of Political Economy* 98:703–38.

DeMiguel, V., L. Garlappi, and R. Uppal. 2009. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies* 22:1915–53.

Diebold, F., and R. Mariano. 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13:253–63.

Fama, E., and K. 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.

———. 2000. Forecasting profitability and earnings. *Journal of Business* 73:161–75.

———. 2001. Disappearing dividends: Changing firm characteristics or lower propensity to pay? *Journal of Financial Economics* 60:3–44.

Fama, E., and G. Schwert. 1977. Asset returns and inflation. *Journal of Financial Economics* 5:115–46.

Ferreira, M., and P. Santa-Clara. 2011. Forecasting stock market returns: The sum of the parts is more than the whole. *Journal of Financial Economics* 100:514–37.

Ferson, W., and C. Harvey. 1991. The variation of economic risk premiums. *Journal of Political Economy* 99:385–415.

Ferson, W., S. Sarkissian, and T. Simin. 2003. Spurious regressions in financial economics? *Journal of Finance* 58:1393–413.

French, K., G. Schwert, and R. Stambaugh. 1987. Expected stock returns and volatility. *Journal of Financial Economics* 19:3–29.

Goetzmann, W., and P. Jorion. 1993. Testing the predictive power of dividend yields. *Journal of Finance* 48: 663–79.

Goyal, A., and I. Welch. 2008. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies* 21:1455–508.

Guo, H. 2006. On the out-of-sample predictability of stock market returns. *Journal of Business* 79:645–70.

Hansen, P., and A. Timmermann. 2012. Choice of sample split in out-of-sample forecast evaluation. Working Paper, University of California at San Diego.

Harvey, D., S. Leybourne, and P. Newbold. 1998. Tests for forecast encompassing. *Journal of Business and Economic Statistics* 16:254–59.

Henkel, S., J. Martin, and F. Nardari. 2011. Time-varying short-horizon predictability. *Journal of Financial Economics* 99, 560–80.

Hirshleifer, D., K. Hou, and S. H. Teoh. 2009. Accruals, cash flows, and aggregate stock returns. *Journal of Financial Economics* 91:389–406.

Hirshleifer, D., S. Lim, and S. H. Teoh. 2009. Driven to distraction: Extraneous events and underreaction to earnings news. *Journal of Finance* 63:2287–323.

Hirshleifer, D., and S. H. Teoh. 2003. Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics* 36:337–86.

Hong, H., and J. Stein. 1999. A unified theory of underreaction, momentum trading, and overreaction in asset markets. *Journal of Finance* 54:2143–84.

Hong, H., W. Torous, and R. Valkanov. 2007. Do industries lead stock markets? *Journal of Financial Economcis* 83:367–96.

Huang, D., F. Jiang, J. Tu, and G. Zhou. 2014a. Cost growth and stock returns. Working Paper, Singapore Management University and Washington University in St. Louis.

———. 2014b. Forecasting stock returns during good and bad times. Working Paper, Singapore Management University and Washington University in St. Louis.

Inoue, A., and L. Kilian. 2004. In-sample or out-of-sample tests of predictability: Which one should we use? *Econometric Reviews* 23:371–402.

Inoue, A., and B. Rossi. 2012. Out-of-sample forecast tests robust to the choice of window size. *Journal of Business and Economic Statistics* 30:432–53.

Jobson, J., and R. Korkie. 1981. Performance hypothesis testing with the Sharpe and Treynor measures. *Journal of Finance* 36:889–908.

Kahneman, D., and A. Tversky. 1974. Judgment under uncertainty: Heuristics and biases. *Science* 185:1124–31.

Kandel, S., and R. Stambaugh. 1996. On the predictability of stock returns: An asset allocation perspective. *Journal of Finance* 51:385–424.

Kelly, B., and S. Pruitt. 2013. Market expectations in the cross-section of present values. *Journal of Finance* 68:1721–56.

———. 2014. The three-pass regression filter: A new approach to forecasting using many predictors. *Journal of Econometrics*, forthcoming.

Keynes, J. 1936. *The general theory of employment, interest and money*. London: Macmillan.

Kothari, S., and J. Shanken. 1997. Book-to-market, dividend yield, and expected market returns: A time-series analysis. *Journal of Financial Economics* 44:169–203.

Lemmon, M., and E. Portniaquina. 2006. Consumer confidence and asset prices: Some empirical evidence. *Review of Financial Studies* 19:1499–1529.

Lettau, M., and S. Ludvigson. 2001. Consumption, aggregate wealth, and expected stock returns. *Journal of Finance* 56:815–49.

———. 2001. Resurrecting the (C)CAPM: A cross-sectional test when risk premia are time-varying. *Journal of Political Economy* 109, 1238–87.

———. 2005. Expected returns and expected dividend growth. *Journal of Financial Economics* 76:583–626.

Lewellen, J., and J. Shanken. 2002. Learning, asset-pricing tests, and market efficiency. *Journal of Finance* 57:1113–45.

Ludvigson, S. 2004. Consumer confidence and consumer spending. *Journal of Economic Perspectives* 18:29–50.

McCracken, M. 2007. Asymptotics for out of sample tests of Granger casuality. *Journal of Econometrics* 140:719–52.

Memmel, C. 2003. Performance hypothesis testing with the Sharpe ratio. *Finance Letters* 1:21–3.

Menzly, L., and O. Ozbas. 2010. Market segmentation and the cross-predictability of returns. *Journal of Finance* 65:1555–80.

Menzly, L., T. Santos, and P. Veronesi. 2004. Understanding predictability. *Journal of Political Economy* 112:1–47.

Merton, R. 1980. On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics* 8:323–61.

Nicholls, D., and A. Pope. 1988. Bias in the estimation of multiple autoregressions. *Australian Journal of Statistics* 30:296–309.

Patton, A., and A. Timmermann. 2010. Monotonicity in asset returns: New tests with applications to the term structure, the CAPM and portfolio sorts. *Journal of Financial Economics* 98:605–25.

Paye, B. 2012. 'Déjà vol': Predictive regressions for aggregate stock market volatility using macroeconomic variables. *Journal of Financial Economics* 106:527–46.

Pettenuzzo, D., A. Timmermann, and R. Valkanov. 2014. Forecasting stock returns under economic constraints. *Journal of Financial Economics*, forthcoming.

Pontiff, J., and L. D. Schall. 1998. Book-to-market ratios as predictors of market returns. *Journal of Financial Economics* 49:141–60.

Rapach, D., J. Strauss, and G. Zhou. 2010. Out-of-sample equity premium prediction: Combination forecast and links to the real economy. *Review of Financial Studies* 23:821–62.

Schwert, G. 1989. Why does stock market volatility change over time? *Journal of Finance* 44:1115–53.

Shen, J., and J. Yu. 2013. Investor sentiment and economic forces. Working Paper, University of Minnesota.

Sloan, R. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71:289–315.

Spiegel, M., 2008. Forecasting the equity premium: Where we stand today. *Review of Financial Studies* 21:1453–1454.

Stambaugh, R. 1999. Predictive regressions. *Journal of Financial Economics* 54:375–421.

Stambaugh, R., J. Yu, and Y. Yuan. 2012. The short of it: Investor sentiment and anomalies. *Journal of Financial Economics* 104:288–302.

Timmermann, A. 1993. How learning in financial markets generates excess volatility and predictability in stock prices. *Quarterly Journal of Economics* 108:1135–45.

———. 1996. Excess volatility and predictability of stock prices in autoregressive dividend models with learning. *Review of Economic Studies* 63:523–57.

———. 2006. Forecast combinations. In *Handbook of Economic Forecasting*, Vol 1, ed. G. Elliott, C. Granger, and A. Timmermann. Amsterdam: Elsevier.

West, K. 1996. Asymptotic inference about predictive ability. *Econometrica* 64:1067–84.

Wold, H. 1966. Estimation of principal components and related models by iterative least squares. In *Multivariate Analysis*, ed. P. R. Krishnaiaah, 391–420. New York: Academic Press.

———. 1975. Path models with latent variables: The NIPALS approach. In *Quantitative Sociology: International Perspectives on Mathematical and Statistical Modeling*, ed. H. Blalock, A. Aganbegian, F. Borodkin, R. Boudon, and V. Cappecchi. New York: Academic Press.

Yu, J. 2013. A sentiment-based explanation of the forward premium puzzle. *Journal of Monetary Economics* 60:474–91.

Yu, J., and Y. Yuan. 2011. Investor sentiment and the mean-variance relation. *Journal of Financial Economics* 100:367–81.