Retail Derivatives and Sentiment:

A Sentiment Measure Constructed from Issuances of Retail Structured Equity Products☆

Brian J. Henderson,[a] Neil D. Pearson,[b,c,*] and Li Wang[d]

August 15, 2020

Abstract

We use retail Structured Equity Product (SEP) issuances to construct a new sentiment measure for individual stocks. The SEP sentiment measure predicts negative abnormal returns on the SEPs' reference stocks based on a variety of benchmarks including behavioral factor models and factors based on idiosyncratic volatility, short interest, and the 52-week high effect. Consistent with our interpretation that SEP issuances reflect investor sentiment, aggregate SEP issuances are highly correlated with the Baker-Wurgler sentiment index. Tobit regressions reveal that proxies for attention and sentiment predict demand for SEPs, providing additional evidence consistent with the hypothesis that SEP issuances reflect sentiment.

JEL classification: G13, G14, G23

Keywords: Structured equity products, sentiment, cross-section of returns, predictability

---

☆We thank Kent Daniel, David Hirshleifer, and Lin Sun for sharing the returns of their short- and long-horizon behavioral factors, Hao Xu for excellent research assistance, Lin Sun for a helpful conversation, and Gergana Jostova, Badrinath Kottimukkalur, and Allen Poteshman for their comments. We also thank, participants at the 2020 Cancun Derivatives Conference, seminars at Case Western Reserve and Texas A&M Universities, and a University of Oklahoma Zoom seminar for their feedback. We are responsible for any errors.

[a] George Washington University, Funger Hall #502, 2201 G Street NW, Washington, DC 20052. Tel.: +1 (202) 994-3669. E-mail: bjhndrsn@gwu.edu.

[b] Department of Finance, University of Illinois at Urbana-Champaign, 1206 South Sixth Street, Champaign, Illinois 61820. Tel.: +1 (217) 244-0490. E-mail: pearson2@illinois.edu.

[c] CDI Research Fellow, Canadian Derivatives Institute, 3000, Chemin de la Côte-Sainte-Catherine, Montréal, Québec H3T 2A7, Canada.

[d] Department of Banking and Finance, Weatherhead School of Management, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, Ohio 44106. Tel.: +1 (216) 368-0802. E-mail: lxw429@case.edu.

* Corresponding author: Department of Finance, University of Illinois at Urbana-Champaign, 1206 South Sixth Street, Champaign, Illinois 61820. Tel.: +1 (217) 244-0490. Fax: +1 (217) 244-3102.

E-mail: pearson2@illinois.edu (N.D. Pearson).

Retail Derivatives and Sentiment:
A Sentiment Measure Constructed from Issuances of Structured Equity Products

August 15, 2020

Abstract

We use retail Structured Equity Product (SEP) issuances to construct a sentiment measure for individual stocks. The SEP sentiment measure predicts negative abnormal returns on the SEPs' reference stocks based on a variety of benchmarks including behavioral factor models and factors based on idiosyncratic volatility, short interest, and the 52-week high effect. Consistent with our interpretation that SEP issuances reflect investor sentiment, aggregate SEP issuances are highly correlated with the Baker-Wurgler sentiment index. Tobit regressions reveal that proxies for attention and sentiment predict demand for SEPs, providing additional evidence consistent with the hypothesis that SEP issuances reflect sentiment..

JEL classification: G13, G14, G23

Keywords: Structured equity products, sentiment, cross-section of returns, predictability

1. Introduction

We propose a new measure of retail investor sentiment for individual stocks that is constructed from issuance volumes of retail structured equity products (SEPs). Stocks for which SEP issuance volumes are high during a month experience large negative abnormal returns during the next month. The negative abnormal returns persist robustly across a wide range of benchmarks, including standard and contemporary factor models.[1] Most individual stock SEPs are based on stocks with market capitalizations falling in the top decile of CRSP stocks, and almost all of the reference stocks have market capitalizations falling in the top quintile. Thus, SEPs issuance volumes predict negative abnormal returns on widely followed, liquid, large capitalization stocks that typically can be readily sold short.

Our interpretation of this predictability is that SEPs issuance volumes proxy for retail investor sentiment, so that high issuance volumes are correlated with high retail investor sentiment and overvaluation, which is subsequently corrected. This interpretation is based on two features of the SEPs market. First, the pricing of SEPs based on individual stocks is extremely disadvantageous to retail investors. Henderson and Pearson (2011) find that the offering prices of one popular brand of SEPs, Morgan Stanley SPARQS, were on average about 8% greater than reasonable estimates of the SPARQS' fair values. The SPARQS had an original time-to-maturity of about one year, so that if the underlying stock was fairly priced the 8% markup implies that a SPARQS investor was accepting an abnormal return or "alpha" of about –8% per year. Vokatá (2020) studies a more recent and much larger sample comprising almost the universe of U.S. yield enhancement products and finds average mispricing of similar

[1] These benchmarks include the Fama-French three and five-factor models (Fama and French, 1993, 2015), the Carhart (1997) four-factor model, production-based factor models (Hou, Xue, and Zhang, 2015; Hou et al., 2020), the mispricing and behaviorally motivated factor models proposed by Stambaugh and Yuan (2017) and Daniel, Hirshleifer, and Sun (2020), and all of the models with additional factors based on idiosyncratic volatility, short interest, and the 52-week high effect.

magnitude. Given the overpricing, it is unlikely that many sophisticated investors purchase SEPs.[2]

Second, the SEPs marketing process reveals information about the demands of the unsophisticated retail investors who purchase SEPs. As described in Egan (2019), issuers offer products based on a range of underlying stocks, often on a monthly cycle, and then investors make their purchase decisions from among the offered SEPs. Given this model of SEPs distribution, issuers presumably design their offering menus to include the stocks they expect retail investors to demand, and then of course the investors' purchase decisions reveal their preferences from among the underlying stocks offered. Through this process, SEP issuances reveal the stocks that interest unsophisticated investors.

SEPs issuance volumes are typically small relative to the market capitalizations and trading volumes of their underlying stocks. We hypothesize that the SEPs sentiment measure is correlated with overvaluation and predicts stock returns because the demands of the investors who purchase SEPs are correlated with the demands of those who purchase the underlying stocks rather than the SEPs.

In addition to the main abnormal return results, we also provide a range of other evidence consistent with the hypothesis that SEP issues reflect sentiment. This includes showing that an aggregate market-wide measure of SEP issuance volumes is highly correlated with the well-known Baker-Wurgler (2006) index and that proxies for investor attention and sentiment predict SEP issuances.

[2]Consistent with a lack of sophistication, Egan (2019) shows that some investors buy strictly dominated SEPs. Célérier and Vallée (2017) provide evidence of large markups on the SEPs in their sample and argue that issuers structure SEPs to cater to yield-chasing investors who focus on the SEPs' headline rates. As pointed out by Vokatá (2020), regulatory investigations provide evidence that some SEP investors do not understand the products' terms (U.S. Securities and Exchange Commission, 2011) and that some broker-dealers aggressively market the products to investors with little investing experience.

We construct the sentiment measure using data on the SEPs based on a single common stock and issued during the period running from January 2004 through December 2019. We construct the measure for each month and underlying reference stock as the ratio of SEPs issuance proceeds to the underlying stock's market capitalization, normalized by the sum of the ratios. Due to the normalization the sentiment measures sum to one and thus are interpretable as portfolio weights. We then test whether this measure predicts future stock returns.

Specifically, we use the portfolio weights to construct portfolios of the underlying reference stocks, and then examine the returns of the portfolios. We find that the SEPs-based sentiment measure predicts large negative abnormal returns on the underlying stocks during the month following the portfolio formation month. Using the Fama and French (1993) three, Carhart (1997) four, and Fama and French (2015) five-factor models as benchmarks, the abnormal returns based on these models are −1.23% (t-statistic −4.52), −1.21% (t-statistic −4.45), and −1.10% (t-statistic −3.90), respectively. These results are consistent with the hypothesis that retail demand revealed by SEPs issuances is an indicator of sentiment that causes some stocks to be overvalued.

We also consider alternative benchmarks consisting of the four and five-factor production-based or q-factor models described by Hou, Xue, and Zhang (2015) and Hou et al. (2020), the Stambaugh and Yuan (2017) mispricing four-factor model, and the Daniel, Hirshleifer, and Sun (2019) behaviorally-motivated three-factor model. Using these models, the alphas during the first month after portfolio formation are −1.09% (t-statistic −4.06), −0.92% (t-statistic −3.31), −1.12% (t-statistic −3.66), and −0.85% (t-statistic −3.09), respectively. Thus, the predictability is slightly mitigated using behaviorally-motivated factors compared to using the other factors. This suggests that the SEP sentiment measure shares a common component with the mispricing and behavioral factors, which is to be expected. However, most of the abnormal return survives benchmarking using the mispricing and behaviorally-motivated factors. The SEP-sentiment-weighted portfolio returns have significantly negative loadings on the Hou et al. (2020) expected growth factor EG, the Stambaugh and Yuan (2017) management factor MGMT, and the Daniel, Hirshleifer, and Sun (2020) financing and post-earning-announcement-drift factors FIN and PEAD.

We also construct and add to all of the models additional factors based on variables that have been shown to predict stock returns. These consist of factors based on idiosyncratic volatility (Ang et al., 2006), short interest (Akbas et al., 2017), and the 52-week high (George and Hwang, 2004). After controlling for these factors, the alphas during the first month after portfolio formation are similar to those obtained when the additional factors are not included. An alternative approach using Fama-Macbeth (1973) cross-sectional regressions also finds that SEPs sentiment predicts negative stock returns controlling for other variables.

We validate our interpretation of SEP issues as reflecting investor sentiment by providing a range of additional evidence consistent with this hypothesis. Our SEP-based sentiment measure is for individual stocks and thus is not directly comparable to the well known Baker-Wurgler (2006) sentiment index, which is a market-wide measure. However, we use the SEP data to construct an aggregate market-wide measure and compare that to the Baker-Wurgler index. Consistent with the view that SEP issues reflect sentiment, the aggregate measure constructed from SEP issuance volumes is highly correlated with the Baker-Wurgler index.

Panel Tobit regressions reveal that determinants of SEP issuance volumes include proxies for investor attention and sentiment, including Google search volumes, the Stambaugh, Yu, and Yuan (2015) mispricing measure, the Barker-Wurgler (2006) sentiment index, past returns, and past realized stock return volatilities. This is consistent with SEP issues reflecting investor sentiment. Long portfolios formed from stocks for which the Tobit models predict negative sentiment experience positive abnormal returns during the month following portfolio formation. However, both Fama-Macbeth regressions and returns of sorted portfolios show that the ability of SEPs issuance volumes to predict returns is incremental to that of these other variables.

We also find that a significant fraction of the abnormal return subsequent to SEP issues is realized during a short window around the first earnings announcement following the SEP issue. This is consistent with biased expectations being partially corrected when information is released.

Our paper is related to the literature on retail investors and stock prices. Early literature provides evidence that retail demand can drive stock prices away from fundamental values, consistent with retail demand being a proxy for investor sentiment and with the view of retail investors as noise traders (Hvidkjaer, 2008; Barber, Odean, and Zhu, 2009). However, more recent studies conclude that retail traders profit by providing liquidity to other investors and that retail order flow predicts future stock returns (Kelley and Tetlock, 2013; Chen, et.al., 2014; Kelley and Tetlock, 2017; Barrot, Kaniel, and Sraer, 2016; Boehmer et al., 2019).[3] Using sentiment proxies based on mutual fund flows, Frazinni and Lamont (2008) construct a proxy for retail demand for different stocks from quarterly mutual fund flows and find that the measure predicts short-term positive and long-run negative future stock returns. On the other hand, Keswani and Stolin (2008) use monthly mutual fund flow data and find robust evidence for a

[3] Aboody, Even-Tov, Lehavy, and Trueman (2018) propose that the overnight return is a measure of retail investor sentiment. However, for the top size quartile stocks, which are close to our sample stocks, the abnormal return of the portfolio based on the sentiment proxy is only 42 basis points, with weak statistical significance.

smart money effect in buying (but not selling) decisions made by both individuals and institutions.

Our SEPs sentiment measure differs from these other measures because it is based on the purchases of the subset of retail investors who buy SEPs and thus are almost certainly unsophisticated. In contrast to the measures based on mutual fund flows, the selection of stocks is not made by mutual fund managers. Our measure is also available on a monthly rather than a quarterly basis, allowing us to study the relation between retail investor demand and subsequent returns over a monthly horizon.

The remainder of the paper is organized as follows. Section 2 describes the data we use, focusing on the SEPs issuance data. Section 3 details our construction of the sentiment measure from the SEPs issuance volumes. Section 4 examines the relation between this sentiment measure and future stock returns. Section 5 provides additional evidence consistent with the hypothesis that SEP issues reflect sentiment, and Section 6 explores whether long portfolios formed based on predicted SEP issuances provide positive abnormal returns. Section 7 briefly concludes.

2. Data

We construct the SEPs sentiment measure from an extensive dataset of U.S. publicly issued SEPs. SEPs are equity-linked notes issued by financial institutions, typically investment banks or investment banking subsidiaries of commercial banks. They are liabilities of the issuer but have payoffs linked to the stock price of an unrelated company, a stock index, or a basket of stocks or stock indexes. We restrict our analysis to SEPs based on individual equities because our purpose is to construct a sentiment measure for individual stocks.

We collect the SEPs sample data from filings available through the SEC's EDGAR database. Issuers of SEPs use shelf registration statements and pricing supplements to the registration statements or free writing prospectuses that describe each offering.[4] We collect the terms of the SEPs from the filings, including the reference asset and issue proceeds (size). We match the SEPs reference stocks to the CRSP database using the name and/or ticker symbol of the underlying stock extracted from the filings.

The SEPs sentiment measure requires a sufficiently well populated set of cross-sectional observations. Despite the fact that public U.S. SEPs found in EDGAR date back to 1994, the cross-section was not well populated until 2004. Therefore, in this study we use SEPs issued between January of 2004 and December of 2019 to construct the SEPs sentiment measure. The beginning of 2006 is another plausible starting date since the SEPs market continued to grow in size from 2004 to 2006. We verify that our main results are robust to the choice of the sample start date. Figure 1 presents the number of SEPs issuances (left scale) and the aggregate principal amounts of SEPs issues (right scale) per month from January 2000 to December 2019. Prior to 2004, SEPs monthly issuance activity was low: the number of issues per month never passed 20 and the total principal amount was always below $200 million. After 2004, monthly SEPs issuances and proceeds continued to increase, peaking at $1.5 billion of proceeds in January of 2008 and 429 issues in March 2017. During the 2008 financial crisis, SEPs issues contracted, as did equity market capitalizations, but recovered in late 2009.

We further categorize the sample SEPs based on the industry of the reference stock using the Fama-Fench 12-industry classification.[5] Figure 2 presents the time series of the total

---

[4] The SEC's EDGAR database is accessible at http://www.sec.gov/edgar.shtml. The pricing supplements are filed as Form 424B2 or Form 424B3, while the free writing prospectuses are Form FWP.

[5] The 12-industry classification scheme is available from Ken French's website: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

proceeds of issues based on stocks in the three most frequent reference industries (Business Equipment, Energy, and Finance). Stocks of companies in the Business Equipment industry are the most frequent reference stocks. The only times when Business Equipment was not the dominant industry for SEPs reference equities were during the recovery following the 2008 crisis and April 2018, when SEPs frequently referenced stocks in the Finance industry.

Our analysis utilizes additional data. Stock prices and returns are from the Center for Research on Securities Prices (CRSP). SIC industry classification codes come from Compustat, and from CRSP when the Compustat codes are not available. We obtain factor portfolio returns data from Ken French's website.[6] The returns on the Daniel, Hirshleifer, and Sun (2020) behavioral factors were provided by those researchers, and the mispricing proxies and mispricing factor returns are from Robert Stambaugh's website.[7] We obtain Google trends search volume directly from Google's website.[8] The Baker-Wurgler sentiment indexes SENT and SENT⊥ come from Jeffrey Wurgler's website.[9] Option trading volumes and implied volatilities are from Option Metrics. Earnings announcement dates come from IBES.

3. Construction of the SEP sentiment measure

The SEP sentiment measure utilizes issuances, and thus investor purchases, of newly issued SEPs. The SEPs in our sample offer long exposure to the underlying reference stocks, consistent with SEPs buyers having optimistic views about the underlying stocks.

We construct the sentiment measure using data on the SEPs based on a single common stock and issued during the period running from January 2004 through December 2019. We do

[6] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.
[7] http://finance.wharton.upenn.edu/~stambaug/
[8] See Da, Engelberg, and Gao (2011) for a description of the Google trends search volume data. We download the data following their procedure and use AACC as a reference to convert all search volumes to the same basis. We describe our approach for constructing a cross-sectional measure of Google search volumes in the Appendix.
[9] http://people.stern.nyu.edu/jwurgler/. We use the index SENT.

not include SEPs based on stock indexes, baskets of common stocks, or multiple common stocks because our focus is on a cross-sectional measure for individual stocks. To compute the sentiment measure for stock n during month t, we first compute the total proceeds (dollar amount) $Q_{t,n}$ of all SEP issues linked to stock n in month t. When a stock is not referenced by any SEP issued during month t then $Q_{t,n} = 0$.

The main variable in our analysis scales the SEP proceeds by the market capitalization of the underlying stock. Let $V_{t-1,n}$ be the market capitalization of stock n at the end of month $t-1$, and consider the ratio of SEP proceeds to the market capitalization of the stock at the end of month $t-1$, that is $Q_{t,n}/V_{t-1,n}$. We define the sentiment measure $G_{t,n}$ for stock n in month t to be the ratio $Q_{t,n}/V_{t-1,n}$ normalized by the sum of that ratio for all stocks during month t:

$$G_{t,n} = \frac{Q_{t,n}/V_{t-1,n}}{\sum_{n=1}^{N} Q_{t,n}/V_{t-1,n}}. \tag{1}$$

Due to the normalization the sentiment measures $G_{t,n}$ sum to one and thus can be used as portfolio weights.

For example, in June 2015 there were 26 SEP issues linked to Ford Motor common stock (ticker symbol "F") with total proceeds of \$88.49 million. During the same month, there were 253 single-stock SEPs on 113 unique stocks with total proceeds of \$644.76 million. Using Ford Motor's end of May market capitalization of \$59.23 billion, the proceeds of \$88.49 million represented $Q_{t,n}/V_{t-1,n} = 0.15\%$ of its end of May market capitalization. The sentiment measure for Ford Motor is then this ratio, normalized by the sum of the same ratios for all stocks so that the sentiment measures sum to one. Despite the fact that the proceeds of SEPs based on Ford Motor accounted for 13.72% (\$88.49 / \$644.76) of total SEP proceeds, Ford Motor's sentiment measure for June 2015 was only $G_{t,n} = 3.60\%$ due to its large equity market capitalization.

In some analyses we use the market-capitalization scaled sentiment measure computed using data from three, six, and 12-month windows:

$$G_{t,n}^{3} = \frac{\sum_{j=0}^{2}(Q_{t-j,n}/V_{t-1-j,n})}{\sum_{n=1}^{N}\sum_{j=0}^{2}(Q_{t-j,n}/V_{t-1-j,n})}, \tag{2}$$

$$G_{t,n}^{6} = \frac{\sum_{j=0}^{5}(Q_{t-j,n}/V_{t-1-j,n})}{\sum_{n=1}^{N}\sum_{j=0}^{5}(Q_{t-j,n}/V_{t-1-j,n})}, \tag{3}$$

and

$$G_{t,n}^{12} = \frac{\sum_{j=0}^{11}(Qt_{-j,n}/V_{t-1-j,n})}{\sum_{n=1}^{N}\sum_{j=0}^{11}(Q_{t-j,n}/V_{t-1-j,n})}. \tag{4}$$

To verify that our results are not sensitive to the particular sentiment measure design, we also construct an alternative measure $F_{t,n} = Q_{t,n}/Q_t$ that does not involve scaling by market capitalization, where $Q_t = \sum_{t=1}^{N} Q_{t,n}$ is the sum across stocks of the proceeds of SEPs issued during month *t*. We also compute longer-term versions of this measure based on issues during the previous 3, 6, and 12 months as $F_{t,n}^{3} = \sum_{j=0}^{2} Q_{t-j,n}/\sum_{\mathrm{j}=0}^{2} Q_{t-j}$, $F_{t,n}^{6} = \sum_{j=0}^{5} Q_{t-j,n}/\sum_{j=0}^{5} Q_{t-j}$, and $F_{t,n}^{12} = \sum_{j=0}^{11} Q_{t-j,n}/\sum_{j=0}^{11} Q_{t-j}$ and use them in some analyses.

Figure 3 presents the times series of the main SEP sentiment measure $G_{t,n}$ for several stocks to illustrate the variation of this measure over time. Panel A shows the sentiment measures for the five stocks having the largest average sentiment values among the stocks having at least 60 months of data during the sample period. Those five stocks are: Chesapeake Energy, Delta Air Lines, Under Armour, United Rentals, and US Steel. Panel B shows the time series of $G_{t,n}$ for the five stocks (Caterpillar, Celgene, General Motors, Micron Tech, and Schlumberger) that have sentiment values closest to the median value among stocks having at least 60 months of data during the sample period. Both figures show that there is considerable time-series variation in the measure $G_{t,n}$, indicating that the measure does not simply capture a stock fixed effect.

4. SEP sentiment and return predictability

We hypothesize that the retail demand for SEPs issuances serves as an indicator of investor sentiment that causes some stocks to be overvalued. This hypothesis implies that the stocks underlying SEPs will underperform risk-adjusted benchmarks following SEPs issuance, consistent with "noise trader" theory (Shleifer and Summers, 1990; DeLong et al., 1990, 1991).

4.1 Abnormal returns relative to traditional return benchmarks

We use calendar time portfolio return regressions to examine the ability of the SEP sentiment measure to predict future stock returns. The portfolio constituents consist of SEPs reference stocks during a given month. If SEPs reveal sentiment associated with overpricing, then the portfolio of SEPs reference stocks will underperform risk-adjusted benchmarks. Thus, we test the null hypothesis that the average subsequent abnormal returns to the portfolios, that is the regression intercepts, are equal to zero.

We construct calendar-time portfolios and their returns using the following procedure. During each month, the cross-sectional individual stock SEPs sentiment values sum to one and thus can be interpreted as portfolio weights. For each calendar month $t+1$, we use the sentiment measure from the previous month, $G_{t,n}$, and compute the portfolio returns as:

$$R_{t+1} = \sum_{n=1}^{N} G_{t,n} r_{t+1,n} \qquad (6)$$

This procedure results in monthly calendar time portfolio returns in which stocks are weighted by their SEPs sentiment score.[10] We then compute excess returns $R_{t+1} - r_{f,t}$, where $r_{f,t}$ is the one-month risk-free rate from the end of month t, and test the predictive ability of the SEPs sentiment

[10] In untabulated results, we repeat the analyses using the three, six, and 12-month sentiment measures as the weights to construct the portfolio returns. We also construct the portfolio returns as $R_{t+1} = \sum_{n=1}^{N} F_{t,n} r_{t+1,n}$ using the SEPs sentiment measure F that does not involve scaling by market capitalization. In both cases the estimates of the alphas are similar in magnitude and statistical significance to those we obtain when we construct the portfolio returns using the weights $G_{t,n}$.

measure by estimating time-series regressions using standard factor models: the Fama-French (1993) three, Carhart (1997) four, and Fama-French (2015) five-factor models. If the SEPs sentiment measure predicts returns, then we expect the average abnormal return, or regression intercept, to be negative. The null hypothesis is that the regression intercepts or abnormal returns of the calendar-time portfolio returns are zero.

Table 1 reports the coefficient estimates and associated *t*-statistics for the various factor models. The results in the left-hand part of the table use the SEPs sentiment measure beginning in 2004, while the right-hand columns present results for the sample period starting in 2006 at which point the SEPs market had further matured. Across all six regressions, the intercepts or alphas are always negative, ranging from −1.31% (*t*-statistic −4.58) to −1.10% (*t*-statistic −3.90). The negative and statistically significant intercepts show that the portfolios of SEPs reference stocks, weighted by their relative popularity among SEPs investors, underperform the standard risk-adjusted benchmarks. The magnitudes of the intercepts are economically meaningful. Thus, the SEPs sentiment measure predicts economically and statistically significant negative abnormal returns after accounting for standard risk factors. These negative abnormal returns are consistent with subsequent corrections to overpricing of the reference stocks during the months of high SEPs issuance.

The estimated factor loadings indicate that SEPs' underlying stocks have high loadings on the market factor MKTRF, ranging from 1.3590 to 1.4063. The loadings on SMB are positive and statistically significant, though only at the 10% level in the results for the five-factor model in Panel C. In Panel C the point estimates of the coefficient on RMW are large, −0.3424 and −0.34215, but the *t*-statistics are only −1.73 and −1.62, respectively.

To check whether our results are driven by the returns in only part of the sample period, we plot the time series of the realized abnormal returns (that is, the sum of the intercept and residual) computed using the Fama-French three-factor model. Figure 4 presents two plots, the realized abnormal returns (solid line), and the 12-month moving average of the realized abnormal returns (dashed line). Both plots show seemingly random variation over time and indicate that the predictability of our SEP-based sentiment measure is not concentrated in any period. We also compute the Durbin-Watson test statistic for first-order autocorrelation of the realized abnormal returns and find that it is 1.86, close to 2. This confirms the visual impression of no important positive serial correlation in the realized abnormal returns.

4.2 Alternative benchmarks

Several recent asset pricing models propose alternative or additional factors, some of which are behaviorally motivated. We estimate the abnormal returns of the calendar time portfolios relative to these alternative benchmarks. The alternative models we use are the production-based *q*-factor model of Hou, Xue, and Zhang (2015), the *q*5-factor model of Hou et al. (2020), the Stambaugh and Yuan (2017) mispricing factor model, and the Daniel, Hirshleifer, and Sun (2020) behavioral factor model. Benchmarking the SEPs sentiment portfolio returns using these factor models allows us to ascertain whether the negative abnormal returns evident using traditional factor models are explained by the investment factors or any of the behavioral or mispricing factors.

Panels A to Panel D of Table 2 report the results of the calendar time regressions of the returns on the portfolio of SEPs reference stocks benchmarked using the four alternative models. Scanning the panels, the estimated intercepts (alphas) are all negative and statistically significant, but smaller than the estimates for the traditional factor models reported in Table 1. Specifically, the estimates of alpha range from −0.85% (t-statistic −3.09) in the Daniel, Hirshleifer, and Sun (2020) model estimated using the 2004‒2019 sample to −1.14% (t-statistic −3.56) for the Stambaugh and Yuan (2017) model estimated using the 2006‒2019 sample, respectively. The results that the abnormal returns are smaller using the Daniel, Hirshleifer, and Sun (2020) behavioral factors as compared to the traditional Fama-French factors and the Hou, Xue, and Zhang (2015) production-based factors are consistent with the SEPs sentiment measure sharing a common component with the behavioral factors. However, SEPs' ability to predict future stock returns persists after controlling for these additional factors, consistent with the hypothesis that SEPs issuances provide incremental information about sentiment over and beyond the existing behavioral factors.

Similar to the results for the traditional factor models, the coefficients on the market factor are large and highly significant in all specifications. The results presented in Panel B show that the SEPs sentiment-weighted portfolios' returns have a large and significant negative loading on the Hou et. al. (2020) expected growth factor EG. This negative relation indicates that the SEPs stocks tend to have low expected growth.

The results for the Stambaugh and Yuan (2017) model reveal that the SEP sentiment-weighted portfolio returns have large negative and highly significant loadings on the management-related factor MGMT. The coefficient estimates and t-statistics reported in Panels A‒C show that the SEPs sentiment portfolio does not have significant loadings on the performance-related factors PERF and ROE. This is consistent with the "noise trader" hypothesis that retail demand revealed by SEP issues is not justified by observable fundamentals.

Referring to Table 2 Panel D, the SEPs sentiment-weighted portfolios' returns have large and statistically significant negative loadings on the Daniel, Hirshleifer, and Sun (2020) short-and long-horizon behavioral factors PEAD and FIN. These negative coefficients indicate that the SEPs stocks tend to be ones with high long-run financing activity, but low earnings surprises in the short run.

We also form equal-weighted portfolios of SEP stocks in each month and repeat the analyses for which we report results in Tables 1 and 2. We report the results of these additional analyses in Internet Appendix Table IA1. The alphas are negative, but much smaller than those of the portfolios formed using SEP issue volumes, between −54 basis points (*t*-statistic −3.42) for the Fama-French (1993) three-factor model and −25 basis points (*t*-statistic −1.69) for the Hou et al. (2020) *q*5-factor model. As expected, these results showing that the abnormal returns are smaller for equally-weighted returns indicate that SEP issuance volume, and not just the simple fact of a SEP issue, is useful in predicting returns.

4.3 Other factors: Idiosyncratic volatility, short interest, and 52-week high

Other variables known to predict stock returns include idiosyncratic volatility (Ang et al., 2006), short interest (Akabs et al., 2017), and the 52-week high (George and Hwang, 2004). We add factors based on each of these variables to the various regression models for which we report results in Tables 1 and 2 and examine whether the inclusion of the additional factors impacts the findings discussed above.

Idiosyncratic volatility (IVOL) has been shown to predict the cross-section of stock returns (Ang et al., 2006), with high IVOL stocks experiencing negative returns in the next period.[11] We use two different IVOL factors, both of which involve value-weighted long positions in high IVOL stocks and value-weighted short positions in low IVOL stocks. The first

[11] Stambaugh, Yu, and Yuan (2015) show that the negative relation between IVOL and stock returns holds for overpriced stocks but not for underpriced stocks.

IVOL factor comes from Kenneth French's data library.[12] Stocks are sorted into quintile portfolios by their residual variances computed from the daily Fama and French (1993) three-factor model residuals measured over the previous 60 days. The IVOL factor is then the return on a value-weighted portfolio of the 20% of stocks with the highest residual variances minus the return on a value-weighted portfolio of the 20% of stocks with the lowest residual variances. The construction of the second IVOL factor follows Stambaugh, Yu, and Yuan (2015), who compute IVOL using the three-factor model residuals from the preceding month. The IVOL factor is again the difference between the value-weighted returns of the stocks in the top and bottom IVOL quintiles.

We separately add the IVOL factors to all of the models for which results are reported in Tables 1 and 2. We also estimate models that consist of just the market factor and one of the IVOL factors. The first two sets of columns in Table 3 report the estimates of the intercepts (abnormal returns) and coefficients on the IVOL factors for the various specifications. For brevity, the coefficient estimates on the other factors are not reported in the table.[13] The estimated abnormal returns do not change much when we add the IVOL factors to the models: they range from −94 basis points to −128 basis points and remain highly statistically significant. In the models formed by adding the IVOL factor to the single-factor market model and the Fama-French (1993, 2015) and Carhart (1997) models the estimated loadings on the IVOL factors are not large, ranging from 0.2134 (t-statistic 2.17) to 0.2560 (t-statistic 3.46). In the models formed by adding the IVOL factors to the other models the coefficients on the IVOL factor are small and not statistically significant.

[12] Available at: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/index.html
[13] We report the full set of results in Internet Appendix Table IA2.

Short interest (SI) has also been shown to predict negative abnormal returns (Akbas et al., 2017). Following the literature, each month we sort CRSP common stocks into short interest deciles based on SI_{t-1}, the ratio of short interest to shares outstanding. Then we define the SI factor as value-weighted average returns on a strategy that is long (short) the stocks in the top (bottom) short interest decile. The third set of columns in Table 3 report the regression results after controlling for this additional factor. The alphas range from −89 basis points to −119 basis points, and are statistically significant.

Another factor of interest is based on the findings that stocks with large ratio of current price to the highest price achieved within the past 52 weeks, the 52-week high, experience positive returns during the next 6 months. George and Hwang (2004) interpret this as a form of momentum. We follow the literature and first sort CRSP stocks in each month based on the ratio of current price to the highest price over the past 52 weeks, H52. Then we construct the H52 factor as the value-weighted average return on the strategy of buying the stocks with ratios of current price to the 52-week high falling in the top 30% and shorting the stocks with ratios in the bottom 30%. The last set of columns in Table 3 report the results of models that include this factor. Similar to the other results, we continue to obtain negative and statistically significant estimates of alpha when this additional factor is included in the regression specifications.

4.4 Fama-Macbeth regression analysis

We also use an alternative approach, Fama-Macbeth (1973) regressions, to study the ability of SEP sentiment to predict stock returns while controlling for other variables. Each month we estimate cross-sectional regressions in which we regress the next month's stock returns on the current month SEP sentiment measure $G_{t,n}$, the sentiment measure $G_{t,n}$ and each control variable separately, and the sentiment measure $G_{t,n}$ together with all of the control variables, and then report the time-series averages of the regression coefficients and *t*-statistics based on Newey-West standard errors. The control variables are the log of the market capitalization of the reference stocks (*FirmSize*), the Amihud (2002) illiquidity measure (*Illiquidity*), short interest normalized by the shares outstanding (*ShortInterestRatio*), the ratio of the current price to the highest price over the past 52 weeks (*52WeekHigh*), idiosyncratic volatility (*IVOL*), the Stambaugh, Yu, and Yuan (2015) mispricing measure (*Mispricing*), Google trends search volume (*SVI*), and the slope coefficient or "beta" (*betaBW*) from a regression of stock returns on the Baker-Wurgler index.[14] The construction of the Google trends search volume variable *SVI* is described in Appendix A. For each stock, we estimate *betaBW* by regressing monthly stock returns on the Baker-Wurgler index using all of the available data.

We report the Fama-Macbeth regression results in Table 4. The coefficient on the SEP-based sentiment measure $G_{t,n}$ is negative and significant in all specifications, regardless of the control variables included. These results are consistent with those from the calendar time regressions reported in Tables 1–3, and provide additional evidence consistent with the hypothesis that SEP-based sentiment predicts future stock return at the individual stock level. The Baker-Wurgler beta *betaBW* is significantly negatively related to returns in the two specifications in which this variable appears, and the short interest ratio is significant in the specification that includes all of the control variables.

4.5 Long-term performance

In the previous analyses we used monthly SEP sentiment, $G_{t,n}$, to predict returns during the subsequent month (t+1). Now we examine the ability of $G_{t,n}$ to predict returns during months $t + 2$ and $t + 3$ using the Fama and French (2015) five-factor model as a benchmark.

[14] The Baker-Wurgler index is a market-wide measure that does not differ across stocks for a given month, so the measure itself cannot be included in the cross-sectional Fama-Macbeth regressions.

Specifically, we repeat the regression analysis of SEP sentiment-weighted calendar time portfolio returns but now use returns during the second ($R_{t+2} = \sum_{n=1}^{N} G_{t,n} r_{t+2,n}$) and third ($R_{t+3} = \sum_{n=1}^{N} G_{t,n} r_{t+3,n}$) months after issuance and regress them on the factor realizations in months $t + 2$ and $t + 3$, respectively. We report the results of these regressions in Table 5. Using the 2004–2019 and 2006–2019 samples, for the regressions explaining R_{t+2} the estimates of alpha are −42 basis points (t-statistic −1.41) and −64 basis points (t-statistic −1.99), respectively. The estimated alphas in the regressions explaining R_{t+3} are smaller and insignificant. Thus, the results provide limited evidence that SEP sentiment during month t predicts returns in month t + 2, and no evidence that it predicts returns in month t + 3.

5. Additional evidence that SEP issuances reflect investor sentiment

We now describe other analyses that provide evidence consistent with the hypothesis that SEP issuance volumes proxy for retail investor sentiment. These other analyses include examining whether an aggregate sentiment measure constructed from SEP issuance volumes is correlated with the well known Baker-Wurgler (2006) sentiment index, exploring whether proxies for investor attention and sentiment explain SEP issues, and examing the returns around they earnings announcement dates shortly after the SEPs' pricing dates.

5.1 Are aggregate SEP issuance volumes correlated with the Baker-Wurgler sentiment index?

Our sentiment measure $G_{t,n}$ defined in Equation (1) is for individual stocks. For each month t, scaled SEP issuance volumes $Q_{t,n}/V_{t-1,n}$ are normalized so that the sum $\sum_{n=1}^{N} G_{t,n}$= 1. Thus, there is no reason to expect the individual stock measures $G_{t,n}$ to be correlated with the Baker-Wurgler (2006) sentiment index, which is an aggregate market-wide measure. However, we can construct an aggregate sentiment measure by summing each month's SEP issuance volumes across stocks. If SEP issuances reflect investor sentiment then one would expect these aggregate issuance volumes to be correlated with the Baker-Wurgler index.

To test this prediction, for each month we compute the cross-sectional sum of the unnormalized proceeds $Q_t = \sum_{n=1}^{N} Q_{t,n}$. We then construct a three-month moving average $\mathrm{MA}(Q)_t = (Q_t + Q_{t-1} + Q_{t-2})/3$ of the aggregate SEP proceeds and examine whether the time series of the moving average is related to the times series of the Baker-Wurgler index. Figure 5 plots this series along with the Baker-Wurgler sentiment index *SENT*, using SEP issuance data from 2004–2018.[15] The figure reveals a clear relation between aggregate SEPs proceeds and the Baker-Wurgler index. The correlations between $\mathrm{MA}(Q)_t$ and the Baker-Wurgler index are high, being 0.49 and 0.53 using the SEP data starting in January 2004 or 2006, respectively. The correlations are similar, 0.49 and 0.55, if we use a six-month moving average of aggregate SEP proceeds. These results that aggregate SEP issuance volumes are highly correlated with the Baker-Wurgler sentiment index validates our claim that SEP issuances reflect investor sentiment.

Given the high correlations between aggregate SEP issues as measured by $\mathrm{MA}(Q)$ and the Baker-Wurgler index, it is interesting to explore whether $\mathrm{MA}(Q)$ predicts aggregate stock market returns. We do this by estimating regression models in which we use $\mathrm{MA}(Q)_t$ (in billions of dollars), the month t value of the Baker-Wurgler index SENT_t, and the month t value of the orthogonalized Baker-Wurgler index $\mathrm{SENT}\perp_t$ to predict the three (six)-month returns on the value-weighted CRSP market index over the months $t + 1$ through $t + 3$ ($t + 6$). Letting $R_{t,t+3}^{VW}$ ($R_{t,t+6}^{VW}$) denote the three (six)-month return from the end of month t to the end of month $t + 3$ ($t +$ 6), we estimate the models

[15] The plots stop in 2018 rather than 2019 because the 2019 values of the Baker-Wurgler index *SENT* are not yet available. It begins with the March 2004 moving average computed from SEP issues in January–March 2004. The plot is similar if we use the othogonalized Baker-Wurgler index *SENT*⊥ .

$$R_{t,t+3}^{VW} = b_0 + b_1\mathrm{MA}(Q)_t + b_2\mathrm{SENT}_t + \epsilon_{t,t+3},$$

$$R_{t,t+6}^{VW} = b_0 + b_1\mathrm{MA}(Q)_t + b_2\mathrm{SENT}_t + \epsilon_{t,t+6,}$$

similar models in which $SENT_t$ is replaced by $SENT\perp_t$, and special cases of the models in which either b_1 or b_2 is set to zero. We estimate all models using both overlapping and non-overlapping three and six-month returns and data running from 2004 through 2019. For models that include $SENT_t$ or $SENT\perp_t$ the sample ends with December 2018 (and the last predicted return is for January through either March or June 2019) because *SENT* and *SENT*⊥ are only available through December 2018.

We report the results in Table 6. Columns (1)–(5) of Panel A (B) display the results of models estimated using monthly data and thus overlapping three (six)-month returns, while columns (6)–(10) show the results of models estimated quarterly (semi-annual) data and non-overlapping returns. The results indicate that MA(*Q*) is a significant predictor of future returns. The Baker-Wurgler indexes *SENT* and *SENT*⊥ also significantly predict returns in some of the univariate models. However, when we include both MA(*Q*) and either *SENT* or *SENT*⊥ in the specifications the coefficient on MA(*Q*) is significant and the coefficient on *SENT* or *SENT*⊥ is not. Thus, in this sample MA(*Q*) is a better predictor of returns than *SENT* and *SENT*⊥.

5.2 Tobit regression analyses of the determinants of SEP issuance volumes

We estimate Tobit regression models to study the relation between SEP monthly issuance volumes and a number of covariates, focusing on covariates that proxy for retail investor attention and mispricing. Approximately 98% of the retail SEPs are linked to stocks belonging to the top market capitalization quintile of CRSP stocks (Henderson, Pearson, and Wang, 2020). We form the sample for the Tobit regression analysis from this set of stocks. The unit of observation is a stock-month, for example Ford during July 2014. For month t (e.g., July 2014), the sample consists of CRSP stocks with share code of 10 or 11 that are in the top CRSP market capitalization quintile on the last trading day of month $t-1$ (e.g., the last trading day of June 2014). We use the pooled sample formed from the monthly samples during the period running from January 2004 through December 2018. We also use a smaller pooled sample starting from January 2006 because, as discussed earlier, the market grew rapidly between 2004 and 2006 and may not have been mature by January 2004.

We use issuance volume of SEPs based on stock n scaled by market capitalization, that is, $Q_{t,n}/V_{t-1,n}$, as the dependent variable rather than $G_{t,n}$ because $G_{t,n}$ is bounded from above by one and thus does not fit well with the normality assumption in the Tobit model.[16] The variable $Q_{t,n}/V_{t-1,n}$ is always non-negative and is zero when there is no SEP issue linked to stock n in month t, which occurs frequently. The explanatory variables include: Google trends search volume (SVI), the Stambaugh, Yu, and Yuan (2015) mispricing measure, the Baker-Wurgler sentiment index *SENT*, the option-to-stock trading volume (O/S) ratio, the option implied volatility (the average across call options with strike to stock price ratio between 0.9 and 1.1 and 60 days or less to expiration), past stock returns (over the preceding week, the second trailing week, the preceding month, the two-month window formed from the second and third trailing months, and the three-month window formed from the fourth through sixth trailing months) and past stock volatility (over the preceding month, the two-month window formed from the second and third trailing months, and the three-month window formed from the fourth through sixth trailing months). We also control for the firm size as measured by the natural logarithm of the

[16] Results using $G_{t,n}$ as the dependent variable are similar. We report them in Internet Appendix Table IA4.

previous month-end value of market capitalization, and include calendar month fixed effects and industry dummy variables based on the Fama-French 12-industry classification.

We present the results based on the two sample periods 2004–2018 and 2006–2018 in Table 7.[17] All specifications include lagged values of the Baker-Wurgler index. The first regressions for both sample periods (columns 1 and 6) do not include the SVI attention measure, the mispricing measure, and the O/S ratio. The second (columns 2 and 7) and third (columns 3 and 8) models for each sample period add the SVI and mispricing measures, respectively. The fourth (columns 4 and 9) specifications include both of these variables, and the last specifications (columns 5 and 10) add the O/S ratio. The results show that retail demand revealed in SEP issuances is positively related to retail investor attention as measured by SVI: the coefficient on this variable is positive and statistically significant in every specification in which it appears. This is consistent with the findings in early literature that retail investors tend to purchase attention-grabbing stocks (Barber and Odean, 2008). The estimate of the coefficient on the mispricing factor is also positive, and very highly statistically significant: across the specifications that include this variable the smallest *t*-statistic is 24.54. This result is consistent with the hypothesis that more overpriced stocks are associated with greater SEP sentiment. The coefficient on the Baker-Wurgler index is highly significant in every model, consistent with SEP issuances being related to sentiment.

Retail SEPs demand is also positively related to the O/S ratio in the results in columns 5 and 10, though the interpretation of this coefficient is unclear because the determinants of options trading volume are not well understood. Regardless, the significantly positive

[17] These samples do not include 2019 because the Baker-Wurgler index is not avaible for 2019. Internet Appendix Table IA4 presents results of panel Tobit regression models using $G_{t,n}$ as the dependent variable. Internet Appendix Table IA5 presents the average returns and average abnormal returns on the SEP reference stocks during each of the six months prior to the issuance month.

coefficients on the O/S ratio suggest that the factors that cause investors to trade options are correlated with the factors that lead to demand for SEPs.

The coefficient on implied volatility is positive and highly significant in all specifications. There is a natural supply side interpretation of this coefficient. Célérier and Vallée (2017) argue that issuers structure SEPs to cater to yield-chasing investors who focus on the SEPs' headline rates. Since the investor sells and the issuer retains the embedded options, it is easier for the issuer to structure a SEP with a high headline rate if the underlying stock volatility is high because then the embedded options are more valuable.

The significantly negative coefficient estimates before recent stock returns over the preceding one week, the preceding month, and the trailing three months show that short-term loser stocks are more likely to attract SEPs purchasers, possibly because these investors are short-term contrarian buyers. However, the economic magnitudes of the coefficient estimates are small. Beyond the past three months, there is no evidence of negative correlation and the magnitudes of the coefficients are very small. Lastly, SEPs issuances are also positively correlated with past volatility over the trailing one through six months. A demand side interpretation of this result is that SEP investors prefer SEPs on high volatility stocks, perhaps because volatility is correlated with investor attention. All these results together are consistent with the hypothesis that SEPs issuances reflect retail investor sentiment.

5.3 Returns during short windows around earnings announcement dates

If SEPs are overvalued because investors are overconfident about firms' prospects the overvaluation should be corrected when outcomes are realized (Bernard and Thomas, 1990). Because relevant information is often released on earnings announcement dates, any over (under) valuation due to biased expectations about future cash flows should be partially corrected by the subsequent earnings announcements, resulting in negative (positive) returns over a short window around the announcements (Lewellen, 2010). Since we hypothesize that the SEPs' underlying stocks are overvalued, we expect to observe such corrections during short windows around earnings announcements following the SEPs issuances. We investigate this prediction by examining the reference stock returns around earnings announcements within 60 days of the SEP issuances.

In Table 8 we report the average market-adjusted stock returns during short windows around the first earnings announcements within 15, 30, and 60 days following the SEPs issuances. The results in the table show that the average market-adjusted returns over three and four-day windows around earnings announcements within 15 days of SEPs issuances are negative, −36 bps and −32 bps, respectively, with *t*-statistics of −2.64 and −2.28. Thus, the returns around earnings announcements shortly after SEP issuances are a significant fraction of the abnormal return during the first month following the issuance month. This is consistent with the hypothesis that SEP issues reflect overvaluation, with some of the overvaluation being corrected when news is released.[18] The average returns around the more distant announcements within 30 and 60 days after SEPs issuances are also negative, with magnitudes and *t*-statistics that decline as the horizon lengthens.

6. Returns of portfolio formed using predicted SEP issuance volumes

[18] One should not expect all or most of the overvaluation to be corrected at the first earnings announcement because earnings announcements are only a small fraction of firm news. For example, Engelberg, Mclean, and Pontiff (2018) show that earnings announcements account for only 4% of the news in their 1979–2013 sample. Furthermore, the overvaluation may be due to incorrect beliefs about uncertain future outcomes (e.g., demands of other traders) that are not corrected by company news.

One limitation of using SEPs issuance volumes to predict underlying stock returns is that the non-negative issuance volumes only identify overvalued stocks that should be either avoided or sold short; they do not identify undervalued stocks that should be purchased. The Tobit regression models potentially can identify undervalued stocks because the underlying unobserved latent variable can be either negative or positive, and its predicted value can be computed from the coefficient estimates and covariates. In this section, we use the Tobit models for which results are reported in Table 7 to predict SEPs issuance volumes and then analyze returns on long and short portfolios formed using the predicted SEP issuance volumes. This analysis also sheds light on whether SEPs issuances contain information about returns beyond that in the covariates. If the results show that the predicted issuance volumes predict returns as well as the actual issuance volumes then this implies that the SEPs issuances do not contain additional information beyond that contained in the covariates.

Specifically, we use the Tobit model coefficient estimates and the covariates to compute the predicted value of the Tobit model latent variable for each stock and month. Then for each month, we form a long portfolio from the stocks referenced by SEPs with negative predicted values and a short portfolio from the stocks with positive predicted sentiment measures, following the same procedure as in Section 4.1. We then calculate the portfolio abnormal returns (regression intercepts) based on the Fama-French five-factor model.

Table 9 reports the point estimates and corresponding *t*-statistics of the abnormal returns of the long and short portfolios formed using the predicted SEPs issuance volumes. In Panel A, we restrict the long and short portfolios to include only the 50 stocks with the strongest negative and positive predicted issuances, respectively. The alphas for the long portfolio are all positive across the five regression models, ranging from 41 bps to 47 bps per month, with *t*-statistics ranging from 2.53 to 3.09. This is encouraging, because it indicates that one can also earn abnormal returns on the long side using predicted SEP issuances. Surprisingly, the alphas for the short portfolios are not consistently negative, and statistically insignificant. The alphas for the long-short portfolios based on the first four regression models are positive, between 42 bps and 110 bps. For Models 3 and 4 that include the Stambaugh, Yu, and Yuan (2015) mispricing measure, but not the O/S ratio, the alphas are weakly statistically significant (*t*-statistics of 1.66 and 1.64). The point estimate of the alpha of the long-short portfolio based on Model 5 that includes the O/S ratio is small and insignificant.

Since the Tobit regression models use publicly available variables and the predictions are not perfect, we also report the alphas for the portfolios formed using the actual SEP sentiment measure if there is a SEP issue during the month and the predicted value otherwise. For the long portfolio this involves removing stocks for which there was a SEP issue. For the short portfolio this involves replacing the predicted values of the latent variable with the actual value of $Q_{t,n}$ for stocks for which there was a SEP issue. Table 9 Panel A reports these results in the final three columns. The alphas for the long portfolios are again all positive, ranging from 37 bps to 49 bps, with *t*-statistics ranging from 2.32 to 2.98. For short portfolios, the alphas are all negative, between −58 bps and −68 bps, with *t*-statistics ranging from −1.92 to −2.26. The alphas for long-short portfolios are all positive, between 87 bps and 116 bps, and statistically significant, with *t*-statistics between 2.29 and 2.92. These estimates imply that the annualized returns to the long-short portfolio are around 10% to 14% per year.

Panel B removes the 50-stock constraint so that the long and short portfolios include all stocks with negative and positive predicted sentiment, respectively. Unsurprisingly, the results are similar but weaker. The alphas for the long portfolio are all positive across the five models, around 10 bps or 13 bps, but smaller than those in Panel A. The alphas for the short portfolios based on Models 1−4 are negative, ranging from −35 bps to −146 bps, but statistically insignificant. The alphas for the long-short portfolio are between 47 bps and 123 bps. We also report the alphas for the portfolios formed using actual SEP sentiment if there is one and using the predicted value otherwise in the last three columns of Panel B. The alphas for the long portfolio are 13 bps or 15 bps, with *t*-statistics ranging from 2.51 to 2.84. For the short portfolio, the alphas are between −59 bps and −65 bps, with *t*-statistics of −1.87 to −2.20. The alphas for long-short portfolios are between 60 bps and 80 bps, with sometimes marginal statistical significance (*t*-statistics ranging from 1.88 to 2.37).

These results that long portfolios formed using the Tobit model estimates yield significantly positive abnormal returns are consistent with the hypothesis that some of the information in SEPs issuances is already contained in the covariates. On the other hand, the finding that the abnormal performance of the short portfolios computed using the Tobit model estimates is worse than the performance of portfolios formed using the actual issuance volumes in Tables 1 and 2 indicates that the issuance volumes contain information beyond that in the covariates. Consistent with this, the abnormal performance when we use the actual SEP sentiment if there is an issue and the predicted value otherwise (the last three columns of both Panels A and B) is better than the performance achieved using only the predicted values (the first three columns).

The best performance would be achieved by combining a short position in one of the portfolios for which results are reported in Tables 1 and 2 with the long portfolios for which results are reported in Table 9 Panel A. Using the Fama-French (2015) five-factor model as the benchmark, the results in Table 1 show abnormal performance of −110 basis points per month for the 2004–2019 sample period. The results in Table 9 Panel A show abnormal performance on the long side of at least 41 bps, depending on the Tobit regression model used. Combining these two results, the annualized abnormal performance of the long-short portfolio would be 41 + 110 = 151 bps per month, which annualizes to about 18% per year.

7. Conclusion

This paper proposes a new cross-sectional measure of investor sentiment based on issuances of retail SEPs. The SEP sentiment measure strongly predicts negative abnormal stock returns. Using the 2004–2019 sample period and the Fama-French (1993, 2015) and Carhart (1997) models, the abnormal returns of calendar-time portfolios constructed using the sentiment measure are between −1.10% and −1.23% per month, with highly significant *t*-statistics between −3.90 and −4.52, respectively. We obtain similar results using the Hou, Xue, and Zhang (2015) and Hou et al. (2020) *q*-factor models. Using the Stambaugh and Yuan (2017) and Daniel, Hirshleifer and Sun (2020) mispricing and behavioral factor models the portfolio alphas remain large and statistically significant, though the magnitudes of the abnormal returns are smaller than those obtained using the other models. The results using the Stambaugh and Yu (2017) and Daniel, Hirshleifer, and Sun (2020) models as benchmarks indicate that the SEP sentiment measure provides incremental information about returns that is not captured by the mispricing and behavioral factors.

We add additional factors based on idiosyncratic volatility, short interest, and the 52-week high to the factor models. The estimated alphas remain large and statistically significant when the additional factors are included. Consistent with the results of the calendar-time regressions, we also find that the sentiment measure predicts returns in Fama-Macbeth (1973) regressions that control for a number of other variables known to predict returns.

We interpret the negative abnormal returns as reversals of overvaluation due to investor sentiment. Henderson and Pearson (2011) and Vokatá (2020) find that SEPs' markups are so large that under reasonable assumptions the expected returns of SEPs are less than the risk-free return. This makes it unlikely that many sophisticated investors purchase SEPs, and suggests the hypothesis that SEPs issuance volumes proxy for the demands of unsophisticated noise traders.

Aggregate SEP issuance volumes during each month are highly correlated with the Baker-Wurgler sentiment index, validating our claim that SEP issuances reflect investor sentiment. Tobit regression analysis show that proxies for investor attention, misevaluation, and sentiment help explain SEP issuance volumes. We find significantly negative market-adjusted returns during short windows around the first earnings announcement dates following SEPs issuances. This finding is consistent with some of the overvaluation of the SEPs' underlying stocks being corrected when earnings news is released.

We use the predicted values of the underlying latent variable in the Tobit regression models to form portfolios of stocks with negative predicted sentiment. Such stocks should have positive subsequent abnormal expected returns. Using calendar-time regressions, when we restrict attention to the 50 stocks with the largest negative estimated sentiment we find that the abnormal returns based on the Fama-French five-factor model on portfolios of such stocks are 41 to 47 basis points per month, and statistically significant.

Appendix. Construction of the Google trends search volume index (SVI)

This appendix provides the details of the construction of the Google SVI data used in the results reported in Tables 4 and 8. Google Trends provides aggregate search frequency over a monthly horizon from 2004 to the present. Da et. al. (2011) is the pioneering study on the relationship between Google SVI and investor attention, establishing that SVI is a proxy for retail investor attention. We mainly follow their steps to construct SVI for our sample covering top quintile market capitalization common stocks in CRSP (with security code 10 or 11) over a monthly horizon during our sample period, deviating from their approach only where necessary to construct a cross-sectional measure that is comparable across stocks.

Following Da et. al. (2011), we use the company stock ticker symbols (for example, “AMZN” for Amazon and “BTU” for Peabody) rather than company names in the queries to avoid gathering information on searches due to unrelated non-financial purposes and the issue of variation in company names. There are some exceptions. Ticker symbols that are combinations of letters that are frequently used in non-financial domains, for example, “ACT,” “ALL,” “AN,” “CAR,” “GAS,” “N,” and “SEE,” typically have very large SVIs. We manually identify these tickers and for them use the corresponding company name in place of the ticker symbol. This is necessary for 5.1% of the 2,195 sample stocks.

Next, Google Trends allow for at most five search terms for each search. For each term out of five the output number of searches is scaled by the peak value within all output numbers. For example, if the raw number of searches for stock i in month t is $SV_{i,t}$, then the output SVI from an one-term search would be defined as

$$SVI_{i,t}^{1} = \frac{SV_{i,t}}{SV_{i,MAX}},$$

where $SV_{i,MAX}$ is the time-series maximum number of searches for stock i.

The output SVI from a five-term search would be defined as

$$SVI_{i,t}^{j} = \frac{SV_{i,t}}{SV_{MAX}^{j}},$$

where SV_{MAX}^{j} is the maximum number of searches from the j^{th} five-term search. In other words, the SVIs produced by different searches are expressed using different "units" and are not comparable across stocks.

To make the SVIs comparable across different stocks, we employ a reference stock with ticker AACC in each search. Then all SVIs are further multiplied by the ratio of $\frac{SVI_{AACC,t}^{1}}{SVI_{AACC,t}^{j}}$, that is,

$$SVI_{i,t}^{j} \times \frac{SVI_{AACC,t}^{1}}{SVI_{AACC,t}^{j}} = \frac{SV_{i,t}}{SV_{MAX}^{j}} \times \frac{SV_{MAX}^{j}}{SV_{AACC,MAX}} = \frac{SV_{i,t}}{SV_{AACC,MAX}}$$

By doing this, all the search volume are presented in the same "unit," that is, scaled by the time-series maximum number of searches for AACC, which in our sample occurs in August 2011.

We fill missing values using interpolation, or when interpolation is not possible, extrapolation. Finally, we use the SVI for 1,290 and 2,195 tickers between the years of 2004 and 2019 in Tables 4 and 8, respectively.

References

Aboody, David, Omri Even-Tov, Reuven Lehavy, and Brett Trueman. 2018. Overnight returns and firm-specific investor sentiment. *Journal of Financial and Quantitative Analysis* 53: 485–505.

Akbas, Ferhat, Ekkehart Boehmer, Bilal Erturk, and Sorin Sorescu. 2017. Short interest, returns, and unfavorable fundamental information. *Financial Management* 56: 455–486.

Amihud, Yakov. 2002. Illiquidity and stock returns: cross-section and time-series effects. Journal of Financial Markets 5: 31–56.

Ang, A., R. Hodrick,Y. Xing, and X. Zhang. 2006. The Cross-section of volatility and expected returns. *Journal of Finance* 61:259–99.

Baker, Malcolm and Jeffrey Wurgler, 2006. Investor sentiment and the cross-section of stock returns. *Journal of Finance* 61(4): 1645–1680.

Barber, Brad M., Terrance Odean, and Ning Zhu, 2009. Do retail trades move markets? *Review of Financial Studies* 22:151–186.

Barber, Brad M. and Terrance Odean, 2008. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies* 21(2): 785–818.

Barrot, J.N., R. Kaniel, and D. Sraer. 2016. Are retail traders compensated for providing liquidity? *Journal of Financial Economics* 120 (1): 146–168.

Bernard, Victor L. and Jacob K. Thomas, 1990. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting and Economics* 13, 305–340.

Boehmer, Ekkhart, Charles M. Jones, Xiaoyan Zhang, and Xinran Zhang, 2019. Tracking retail investor activity. Working paper.

Carhart, M. M. 1997. On persistence in mutual fund performance. *Journal of Finance* 52, 57–82.

Célérier, Claire and Boris Vallée, 2017. Catering to investors through security design: Headline rate and complexity. *Quarterly Journal of Economics* 132 (3): 1469–1508.

Chen, H., P. De, Y. Hu, and B. Hwang, 2014. Wisdom of crowds: The value of stock opinions transmitted through social media. *Review of Financial Studies* 27(5):1367–1403.

Da, Zhi, Joseph Engelberg, and Pengjie Gao. 2011. In search of attention. *Journal of Finance* 66 (5): 1461–1499.

Daniel, K.D., D.A. Hirshleifer, and L. Sun, 2020. Short- and long-horizon behavioral factors. *Review of Financial Studies* 33: 1673–1736.

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

DeLong, J. B., A. Shleifer, L. H. Summers, and R. J.Waldmann. 1991. The survival of noise traders in financial markets. *Journal of Business* 64:1–20.

Egan, Mark. 2019. Brokers vs. Retail investors: Conflicting interests and dominated products. Journal of Finance 74 (3): 1217–1260.

Engelberg, Joseph, R. David Mclean, and Jefferey Pontiff. 2018. Anomalies and news. *Journal of Finance* 73(5): 1971−2001.

Fama, Eugene F. and Kenneth R. French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33: 3−56.

Fama, Eugene F. and Kenneth R. French. 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116:1−22.

Fama, Eugene F. and James D. Macbeth. 1973. Risk, return, and equilibrium: empirical tests. *Journal of Political Economy* 81: 607−636.

Frazzini, Andrea and Lamont, Owen, 2008. Dumb money: Mutual fund flows and the cross-section of stock returns. *Journal of Financial Economics* 88: 299–322.

George, Thomas J. and Chuan-Yang Hwang. 2004. The 52-week high and momentum investing. *Journal of Finance* 5: 2145−2176.

Henderson, Brian and Neil D. Pearson. 2011. The dark side of financial innovation: A case study of the pricing of a retail financial product. *Journal of Financial Economics* 100 (2): 227−247.

Henderson, Brian, Neil D. Pearson, and Li Wang. 2020. Pre-trade hedging: Evidence from the issuance of retail structured products. *Journal of Financial Economics*, 137:108−128.

Hou, K., C. Xue, C., and L. Zhang. 2015. Digesting anomalies: An investment approach. *Review of Financial Studies* 28(3): 650–705.

Hou, Kewei, Haitao Mo, Chen Xue, and Lu Zhang. 2020. An augmented *q*-factor model with expected growth. *Review of Finance*, forthcoming.

Hvidkjaer, Soeren. 2008, Small trades and the cross-section of stock returns. *Review of Financial Studies* 21, 1123–1151.

Kelley, E. K., and P. C. Tetlock. 2013. How wise are crowds? Insights from retail orders and stock returns. *Journal of Finance* 68:1229–65.

Kelley, E. K., and P. C. Tetlock. 2017. Retail short selling and stock prices. *Review of Financial Studies* 30(3):810−834.

Keswani, Aneel, and David Stolin, 2008. Which money is smart? Mutual fund buys and sells of individual and institutional investors. *Journal of Finance* 63(1): 85−118.

Lewellen, Jonathan. 2010. Accounting anomalies and fundamental analysis: An alternative view. *Journal of Accounting and Economics* 50: 455−466.

Shin, Donghwa. 2019. Extrapolation and complexity. Working paper, University of North Carolina.

Shleifer, A. and L. Summers. 1990. The noise trader approach to finance. *Journal of Economic Perspectives* 4:19–33.

Stambaugh, Robert F., Jianfeng Yu, and Yu Yuan, 2015, Arbitrage asymmetry and the idiosyncratic volatility puzzle, *Journal of Finance* 70, 1903–1948.

Stambaugh, R.F., and Yuan, Y. 2017. Mispricing factors, *Review of Financial Studies* 30(4):1270–1315.

U.S. Securities and Exchange Commission, 2011, Staff summary report on issues identified in examinations of certain structured securities products sold to retail investors. https://www.sec.gov/news/studies/2011/ssp-study.pdf.

Vokatá, Petra. 2020. Engineering lemons. Working paper, Ohio State University.

Figure 1
Monthly Structured Equity Product (SEP) issuances and proceeds
This figure shows monthly number of SEP issues (left scale) and total proceeds (principal amounts) in million dollars (right scale) from January 2000 to December 2019. SEPs issuances are identified by searching the SEC's EDGAR website as detailed in Section 2.

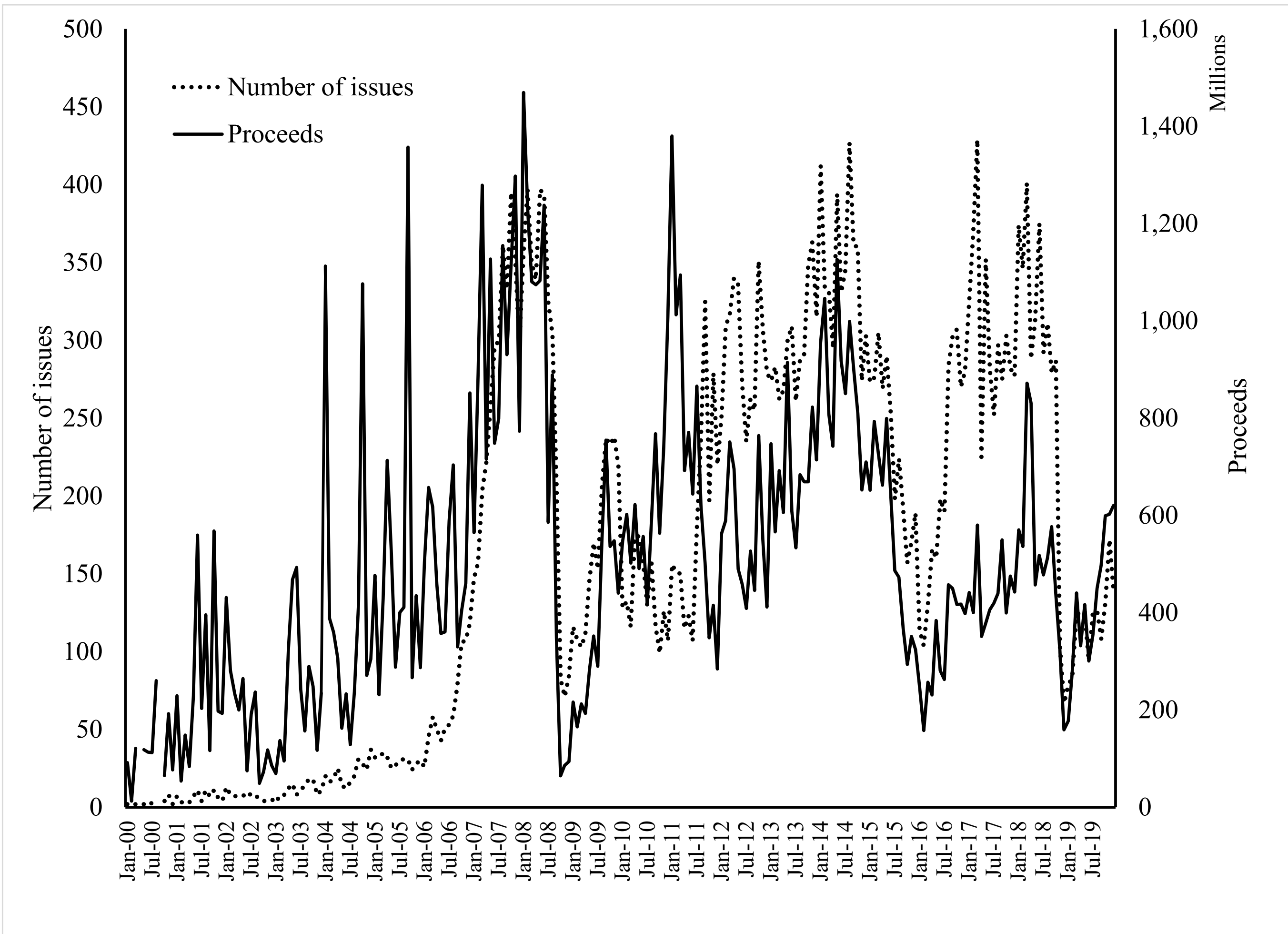

Figure 2
Monthly proceeds of SEPs based on stocks in top three industries
This figure shows the monthly SEP proceeds in millions of dollars for the three most frequently referenced industries (Business Equipment, Energy, and Finance) during the time period running from January 2000 through December 2019.

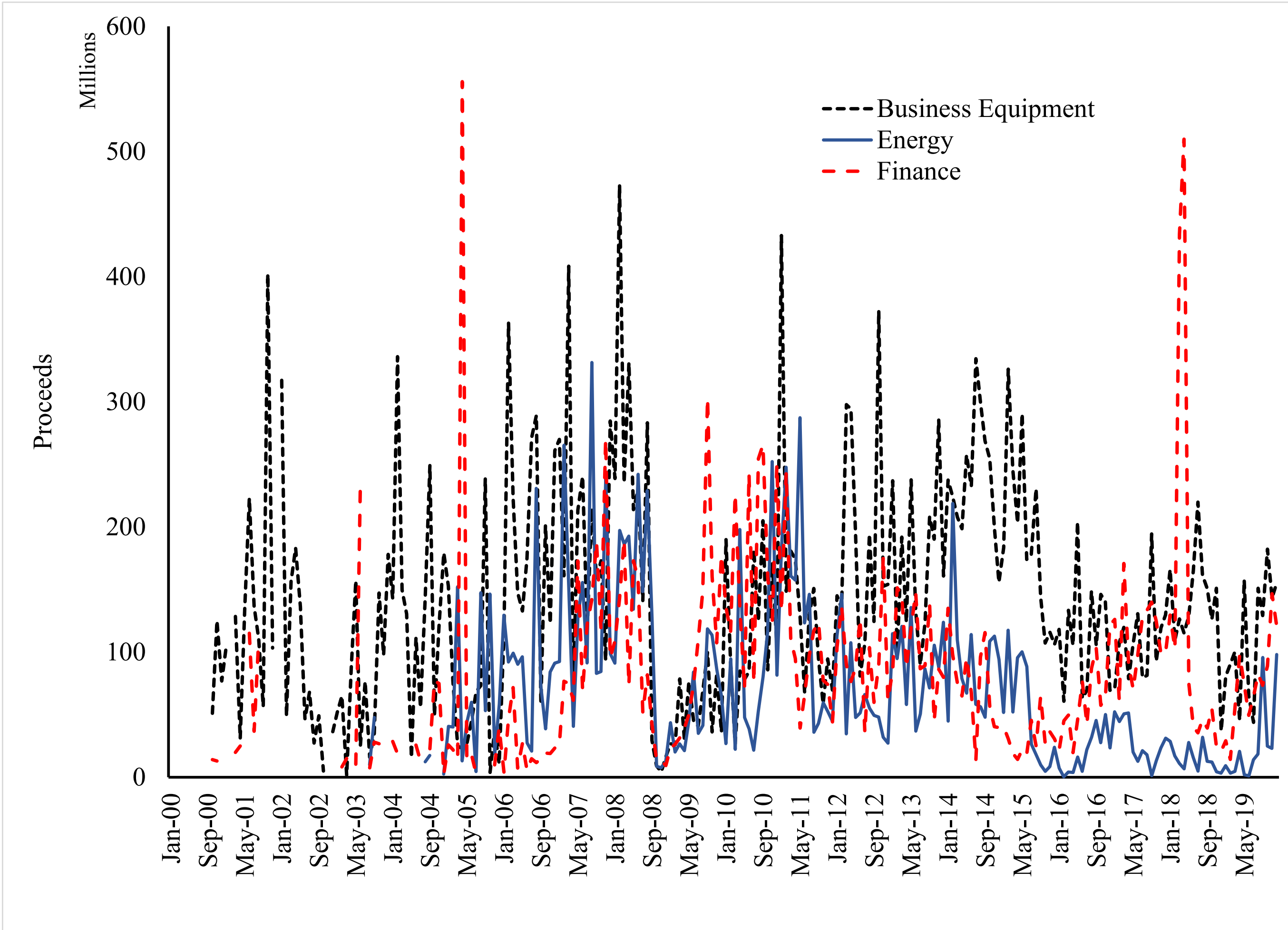

Figure 3
SEP sentiment measures
Time-series of the SEP sentiment measure $G_{t,n}$ during the period running from January 2004 through December 2019. Panel A reports $G_{t,n}$ for the five stocks with greatest average sentiment measures subject to the restriction of having at least 60 months' data in the sample period: Chesapeake Energy (CHK), Delta Air Lines (DAL), Under Armour (UAA), United Rentals (URL), and US Steel (X). Panel B displays the time series of $G_{t,n}$ for five stocks with average levels of $G_{t,n}$ closest to the median: Caterpillar (CAT), Celgene (CELG), General Motor (GM), Micron Tech (MU), and Schlumberger (SLB).

Panel A: Five stocks with largest average values of $G_{t,n}$

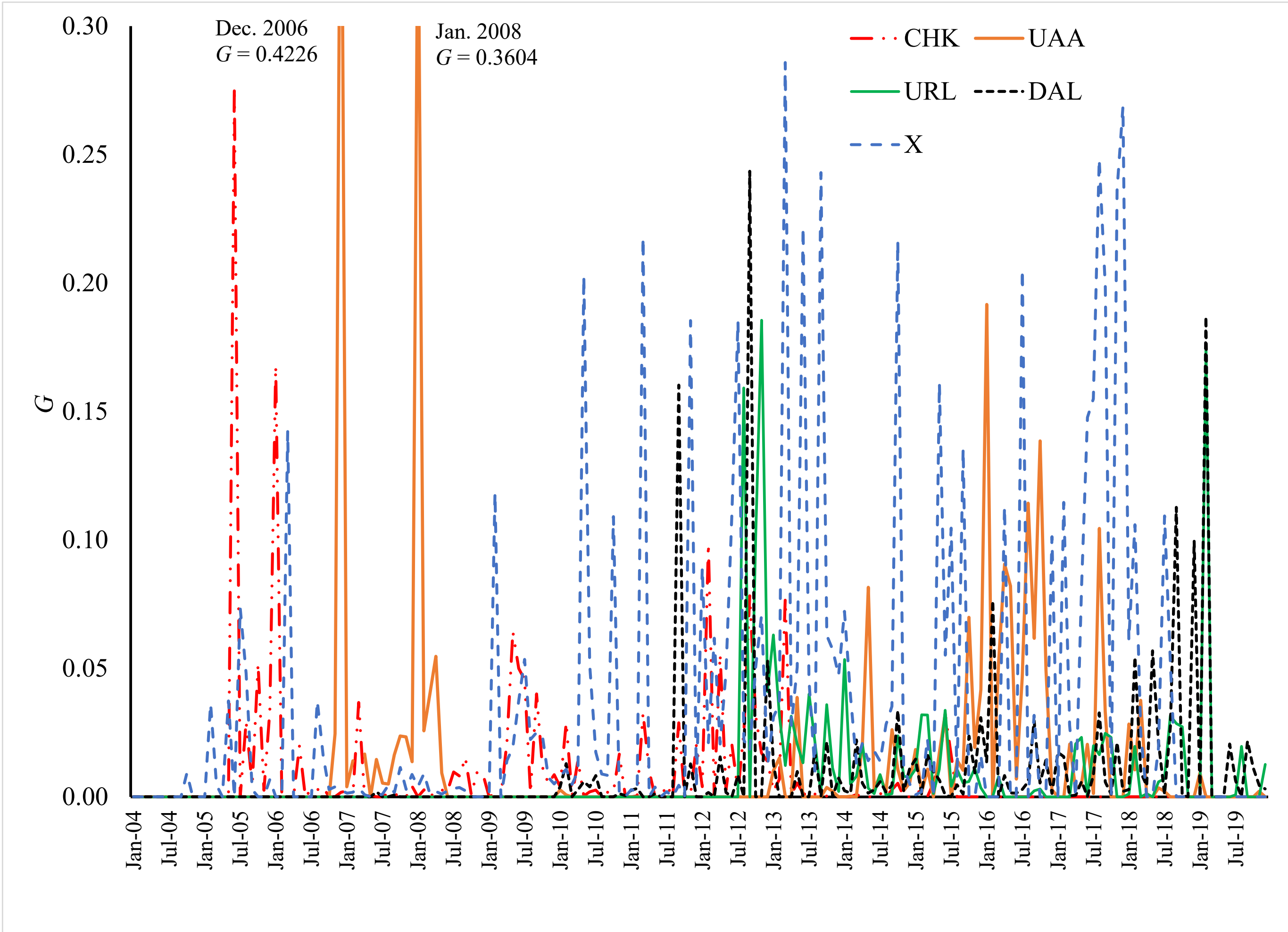

Figure 3 (continued)
Panel B: Five stocks with average G values near the median average value

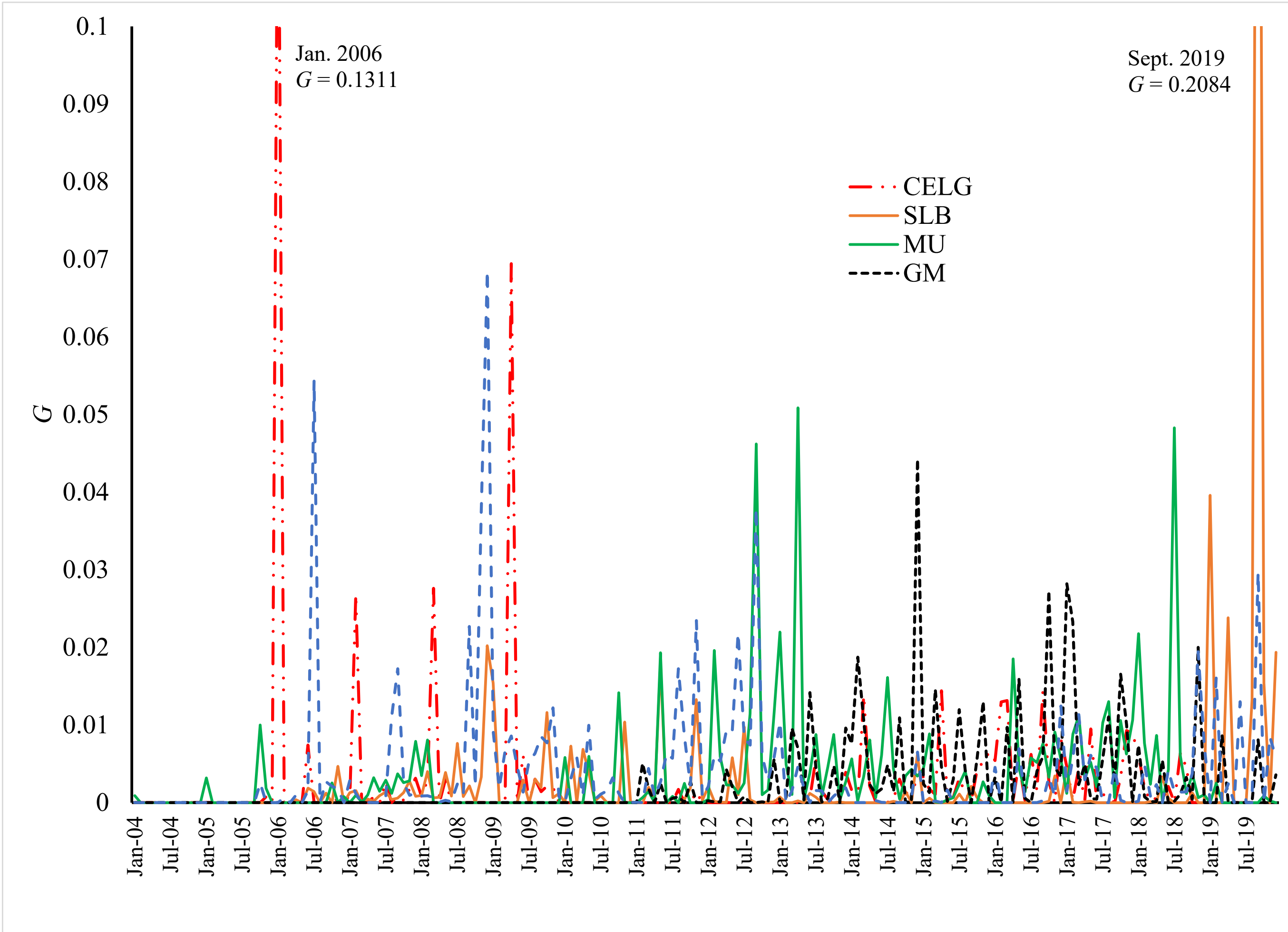

Figure 4: Realized abnormal returns

Time-series of realized abnormal returns from January 2004 to December 2019. Realized abnormal return (dashed line) is defined as sum of the intercept and residual from the Fama-French three-factor model for which estimates are reported Table 1, Panel A. The solid line plots the 12-month moving average of the realized abnormal returns.

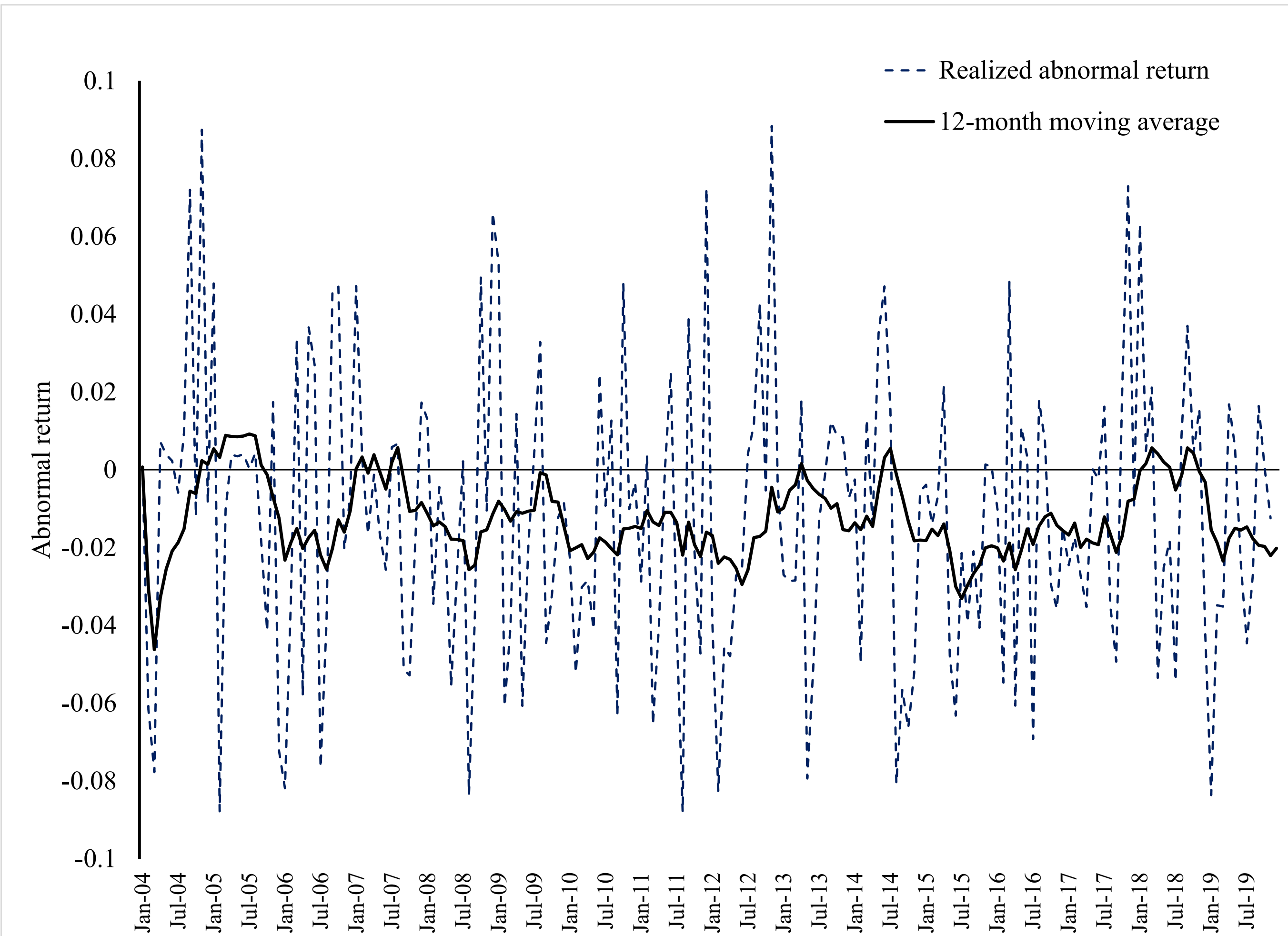

Figure 5
The figure plots the three-month moving average of monthly aggregate SEP proceeds along with the Baker-Wurgler (2006) sentiment index SENT. The monthly aggregate SEP proceeds are computed as the cross-sectional sum of the proceeds of all SEP issues during the month. The sample of SEP issues begins in January 2004 so that the three-month moving average begins in March 2004. The sample ends in December 2018 because the Baker-Wurgler index is not available for 2019.

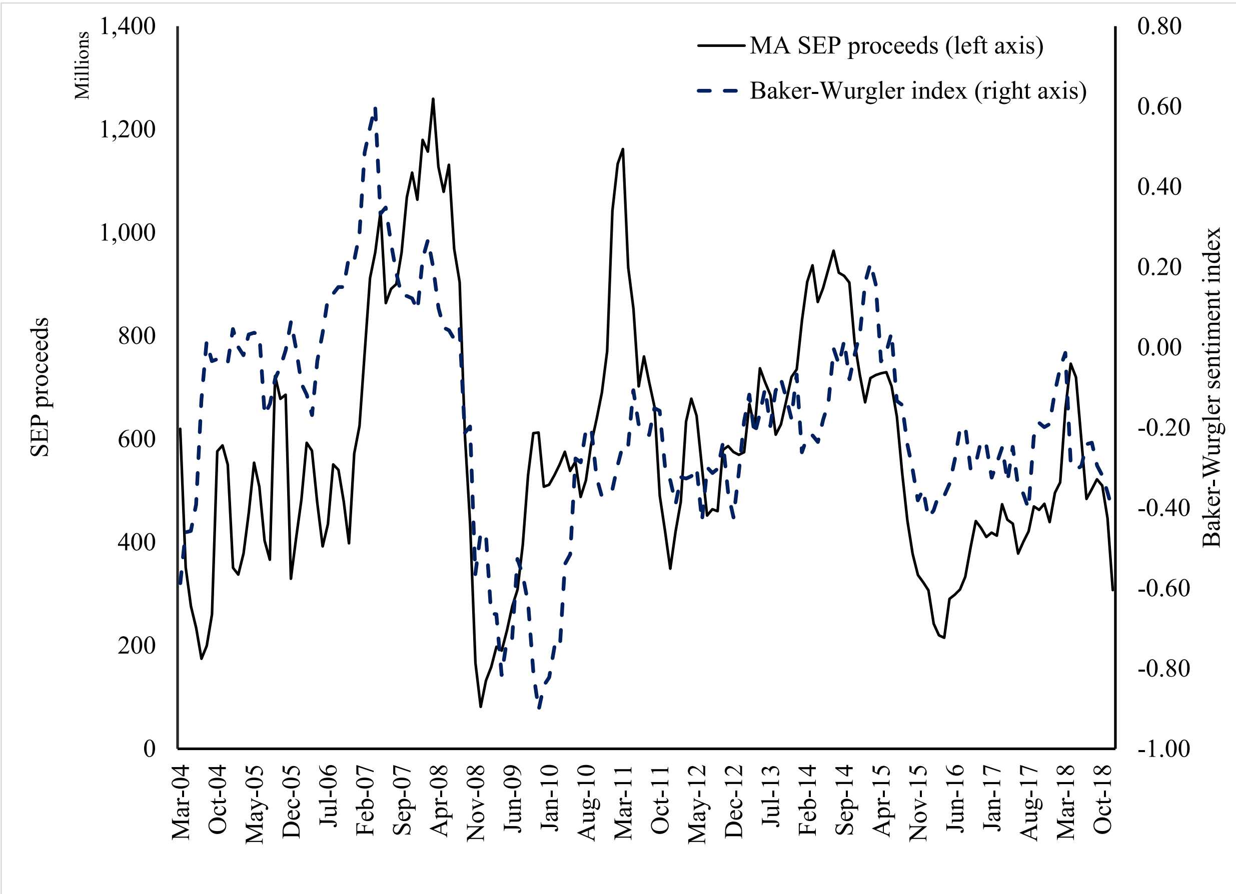

Table 1
Regressions of SEP sentiment-weighted portfolio excess returns on market factors
This table reports results of regressions in which monthly excess returns of portfolios formed using the sentiment measure $G_{t,n}$ are regressed on the factor returns of the Fama-French (1993, 2015) three and five-factor models and the Carhart (1997) four-factor model. Each month, we form a portfolio in which the weight for each stock is the principal amount of SEP issues based on the stock during month t, normalized by the stock's market capitalization at the previous month-end, and divided by the aggregate normalized SEP issue principal over the trailing month. We then estimate regressions in which the dependent variable is the SEP-sentiment-weighted portfolio excess return in month $t + 1$ and the covariates are the factor portfolio returns over the same month, and report the results in the table. Panels A through C report the results using Fama-French three, Carhart four, and Fama-French five-factor models, respectively. The left two columns of each panel present the coefficient estimates and corresponding t-statistics over the full sample period running from January 2004 through December 2019. The right two columns report the results using a sample beginning in January 2006.

	Sample period: 2004−2019		Sample period: 2006−2019	
Variable	Coefficient estimate	*t*-statistic	Coefficient estimate	*t*-statistic
Panel A: Fama-French (1993) three-factor model				
Intercept	**–0.0123**	**(–4.52)**	**–0.0131**	**(–4.58)**
MKTRF	1.4063	(19.12)	1.3795	(18.83)
SMB	0.2808	(2.22)	0.3112	(2.36)
HML	–0.0337	(–0.31)	–0.0499	(–0.45)
Panel B: Carhart (1997) four-factor model				
Intercept	**–0.0121**	**(–4.45)**	**–0.0130**	**(–4.56)**
MKTRF	1.3876	(18.18)	1.3540	(17.87)
SMB	0.2854	(2.26)	0.3080	(2.34)
HML	–0.0750	(–0.63)	–0.1119	(–0.92)
MOM	–0.0647	(–0.92)	–0.0919	(–1.28)
Panel C: Fama-French (2015) five-factor model				
Intercept	**–0.0110**	**(–3.90)**	**–0.0122**	**(–4.11)**
MKTRF	1.3785	(17.44)	1.3590	(17.18)
SMB	0.2185	(1.68)	0.2537	(1.87)
HML	–0.0902	(–0.70)	–0.1225	(–0.90)
RMW	–0.3424	(–1.73)	–0.3415	(–1.62)
CMA	0.1969	(0.88)	0.2040	(0.86)

Table 2
Regressions of SEP sentiment-weighted portfolio excess returns on alternative market factors
This table reports results of regressions in which monthly excess returns of portfolios formed using the sentiment measure $G_{t,n}$ are regressed on the factor returns of an alternative set of factor models. Each month, we form a portfolio in which the weight for each stock is the principal amount of SEP issues based on the stock during month t, normalized by the stock's market capitalization at the previous month-end, divided by the aggregate normalized SEP issue principal over the trailing month. We then estimate regressions in which the dependent variable is the SEP-sentiment-weighted portfolio excess return in month $t + 1$ and the covariates are the factor portfolio returns over the same month, and report the results in the table. Panels A through Panel D reports the using the Hou, Xue, and Zhang (2015) *q* factors, the Hou et al. (2020) *q*5 factors, the Stambaugh and Yuan (2017) mispricing factors, and the Daniel, Hirshleifer, and Sun (2020) behaviorally motivated factors, respectively. The left two columns of each panel present the coefficient estimates and corresponding *t*-statistics over the full sample period running from January 2004 through December 2019. The right two columns report results for a sample beginning in January 2006. Returns for the Stambaugh, Yu, and Yuan (2017) and Daniel, Hirshleifer, and Sun (2020) factors are available only through 2016 and 2018, respectively.

	Sample period: 2004–2019		Sample period: 2006–2019	
Variable	Coefficient estimate	*t*-statistic	Coefficient estimate	*t*-statistic
Panel A: Hou, Xue, and Zhang (2015) q-factor model				
Intercept	**–0.0109**	**(–4.06)**	**–0.0112**	**(–4.01)**
MKT	1.3657	(17.38)	1.3203	(16.91)
ME	0.2105	(1.73)	0.2549	(1.98)
IA	–0.1206	(–0.68)	–0.2100	(–1.17)
ROE	–0.2042	(–1.40)	–0.2359	(–1.58)
Panel B: Hou et al. (2020) q5-factor model				
Intercept	**–0.0092**	**(–3.31)**	**–0.0094**	**(–3.26)**
MKT	1.3159	(16.36)	1.2727	(15.91)
ME	0.1448	(1.17)	0.1972	(1.52)
IA	–0.2176	(–1.20)	–0.3110	(–1.69)
ROE	–0.0958	(–0.63)	–0.1291	(–0.83)
EG	–0.4507	(–2.36)	–0.4300	(–2.24)

Table 2 (continued)
Regressions of SEP sentiment-weighted portfolio excess returns on alternative market factors

	Sample period: 2004–2019		Sample period: 2006–2019	
Variable	Coefficient estimate	*t*-statistic	Coefficient estimate	*t*-statistic
Panel C: Stambaugh and Yuan (2017), factors available through 2016				
Intercept	**–0.0112**	**(–3.66)**	**–0.0114**	**(–3.56)**
MKTRF	1.3710	(14.77)	1.3181	(14.47)
SMB	0.1835	(1.23)	0.2411	(1.51)
MGMT	–0.5840	(–3.92)	–0.6766	(–4.55)
PERF	–0.1267	(–1.63)	–0.1374	(–1.76)
Panel D: Daniel, Hirshleifer, and Sun (2020), factors available through 2018				
Intercept	**–0.0085**	**(–3.09)**	**–0.0091**	**(–3.19)**
MKTRF	1.2989	(17.42)	1.2615	(16.99)
FIN	–0.5046	(–4.71)	–0.5278	(–4.74)
PEAD	–0.3760	(–2.69)	–0.4200	(–3.02)

Table 3
Regressions to explain SEP sentiment-weighted portfolio excess returns including additional factors

This table reports results of regressions in which monthly excess returns of portfolios formed using the sentiment measure $G_{t,n}$ are regressed on the factor returns of the single-factor model and all factor models in Tables 1 and 2, each with one additional factor included. The additional factors are based on idiosyncratic volatility (IVOL), short interest (SI), or the 52-week high (H52). Each month, we form a portfolio in which the weight of each stock is the principal amount of SEP issues based on the stock during month *t*, normalized by the stock's market capitalization at the previous month-end, divided by the aggregate normalized SEP issue amounts over the trailing month. We then estimate regressions in which the dependent variable is the SEP sentiment-weighted portfolio excess return in month *t* + 1 and the covariates are the factor returns over the same month. The estimates of the intercept and the coefficient on the additional factor are reported in the table. (Full results are in Internet Appendix Table IA3.) Panel A reports the results for the single-factor model with the additional factor. Panels B–H repeat the specifications for which results are reported in Tables 1 and 2 but including the additional factor. The first set of columns report the intercept and coefficient estimates for the FF_IVOL factor. The FF_IVOL factor is the return on the strategy of buying a value-weighted portfolio of CRSP stocks in the top residual variance quintile and seling a value-weighted portfolio of stocks in the bottom residual variance quintile. These portfolio returns are taken from the French data library, where residual variance is computed using the daily residuals from the Fama-French three-factor over 60 days. The second set of columns report the estimates using the SYY_IVOL factor. SYY_IVOL differs from FF_IVOL is that residual variance is computed following Stambaugh, Yu, and Yuan (2015) using the residuals from the Fama-French three-factor model over the previous month. The third set of columns report the results of regressions that include the SI factor. The SI factor is the return on the strategy of buying a value-weighted portfolio of CRSP stocks in the largest short interest ratio (short interest divided by outstanding shares) quintile and selling a value-weighted portfolio of CRSP stocks in the smallest short interest ratio quintile. The last set of columns report the results for regressions that include the H52 factor. The H52 factor is the return on the strategy of buying a value-weighted portfolio of the 30% of CRSP stocks with the largest ratios of current price to the highest price over the past 52 weeks and shorting a value-weighted portfolio of the 30% of CRSP stocks with the smallest ratios. Results are reported for both the full sample running from 2004 through 2019 and a sample starting in 2006. Returns for the Stambaugh, Yu, and Yuan (2017) and Daniel, Hirshleifer, and Sun (2020) factors are available only through 2016 and 2018, respectively.

	FF_IVOL (High–Low)			SYY_IVOL (High–Low)			SI (High–Low)			H52 (High-Low)	
	2004–2019	2006–2019		2004–2019	2006–2019		2004–2019	2006–2019		2004–2019	2006–2019
Panel A: Single-factor model with additional factor											
Intercept	–0.0113	–0.0124	Intercept	–0.0112	–0.0120	Intercept	–0.0115	–0.0125	Intercept	–0.0121	–0.0131
	(–3.73)	(–3.80)		(–3.68)	(–3.70)		(–3.67)	(–3.75)		(–3.99)	(–4.08)
FF_IVOL	0.2560	0.2541	SYY_IVOL	0.2821	0.3029	SI	0.2262	0.2504	H52	–0.2518	–0.2742
	(3.46)	(3.36)		(3.61)	(3.78)		(2.40)	(2.65)		(–3.14)	(–3.45)
Panel B: Fama-French (1993) three-factor model with additional factor											
Intercept	–0.0115	–0.0128	Intercept	–0.0114	–0.0123	Intercept	–0.0119	–0.0131	Intercept	–0.0121	–0.0133
	(–3.77)	(–3.92)		(–3.71)	(–3.78)		(–3.78)	(–3.92)		(–4.03)	(–4.20)
FF_IVOL	0.2315	0.2146	SYY_IVOL	0.2534	0.2605	SI	0.1668	0.1853	H52	–0.2469	–0.2710
	(2.70)	(2.49)		(2.76)	(2.78)		(1.54)	(1.70)		(–2.98)	(–3.3)

Table 3 (continued)
Regressions to explain SEP sentiment-weighted portfolio excess returns including additional factors

	FF_IVOL (High–Low)			SYY_IVOL (High–Low)			SI (High–Low)			H52 (High–Low)	
	2004–2019	2006–2019		2004–2019	2006–2019		2004–2019	2006–2019		2004–2019	2006–2019
Panel C: Carhart (1997) four-factor model with additional factor											
Intercept	–0.0115	–0.0128	Intercept	–0.0113	–0.0123	Intercept	–0.0119	–0.0132	Intercept	–0.0118	–0.0125
	(–3.74)	(–3.89)		(–3.69)	(–3.74)		(–3.78)	(–3.94)		(–3.97)	(–4.00)
FF_IVOL	0.2444	0.2134	SYY_IVOL	0.2661	0.2657	SI	0.1535	0.1617	H52	–0.4585	–0.5064
	(2.57)	(2.17)		(2.63)	(2.49)		(1.36)	(1.42)		(–3.63)	(–3.85)
Panel D: Fama-French (2015) five-factor model with additional factor											
Intercept	–0.0110	–0.0126	Intercept	–0.0107	–0.0120	Intercept	–0.0109	–0.0123	Intercept	–0.0116	–0.0131
	(–3.47)	(–3.72)		(–3.37)	(–3.55)		(–3.31)	(–3.54)		(–3.68)	(–3.97)
FF_IVOL	0.2355	0.2224	SYY_IVOL	0.2600	0.2772	SI	0.1643	0.1896	H52	–0.2422	–0.2717
	(2.55)	(2.39)		(2.66)	(2.78)		(1.48)	(1.69)		(–2.79)	(–3.16)
Panel E: Hou, Xue, and Zhang (2015) q-factor model with additional factor											
Intercept	–0.0116	–0.0120	Intercept	–0.0115	–0.0118	Intercept	–0.0112	–0.0114	Intercept	–0.0120	–0.0121
	(–3.83)	(–3.77)		(–3.79)	(–3.71)		(–3.65)	(–3.58)		(–4.03)	(–3.90)
FF_IVOL	0.1452	0.0890	SYY_IVOL	0.1526	0.1251	SI	0.1448	0.1420	H52	–0.1952	–0.2115
	(1.39)	(0.82)		(1.40)	(1.08)		(1.34)	(1.31)		(–1.94)	(–2.08)
Panel F: Hou et al. (2020) q5-factor model with additional factor											
Intercept	–0.0098	–0.0102	Intercept	–0.0097	–0.0101	Intercept	–0.0089	–0.0093	Intercept	–0.0101	–0.0103
	(–3.21)	(–3.19)		(–3.16)	(–3.13)		(–2.85)	(–2.86)		(–3.33)	(–3.26)
FF_IVOL	0.1075	0.0484	SYY_IVOL	0.1273	0.0960	SI	0.1766	0.1705	H52	–0.1901	–0.2083
	(1.03)	(0.45)		(1.18)	(0.83)		(1.66)	(1.60)		(–1.92)	(–2.09)
Panel G: Stambaugh and Yuan (2017) mispricing factor model with additional factor, SYY factors available through 2016											
Intercept	–0.0111	–0.0114	Intercept	–0.0109	–0.0113	Intercept	–0.0104	–0.0107	Intercept	–0.0117	–0.0122
	(–3.60)	(–3.55)		(–3.55)	(–3.52)		(–3.36)	(–3.31)		(–3.83)	(–3.81)
FF_IVOL	0.0849	0.0011	SYY_IVOL	0.0990	0.0447	SI	0.1490	0.1551	H52	–0.1971	–0.2112
	(0.79)	(0.01)		(0.88)	(0.38)		(1.41)	(1.48)		(–1.71)	(–1.79)

Table 3 (continued)
Regressions to explain SEP sentiment-weighted portfolio excess returns including additional factors

	FF_IVOL (High–Low) 2004–2019	FF_IVOL (High–Low) 2006–2019		SYY_IVOL (High–Low) 2004–2019	SYY_IVOL (High–Low) 2006–2019		SI (High–Low) 2004–2019	SI (High–Low) 2006–2019		H52 (High–Low) 2004–2019	H52 (High–Low) 2006–2019
Panel H: Daniel, Hirshleifer, and Sun (2020) behavioral factor model with additional factor, DHS factors available through 2018											
Intercept	–0.0094	–0.0102	Intercept	–0.0094	–0.0102	Intercept	–0.0090	–0.0097	Intercept	–0.0095	–0.0103
	(–3.11)	(–3.21)		(–3.12)	(–3.20)		(–2.95)	(–3.04)		(–3.18)	(–3.27)
FF_IVOL	0.0958	0.0852	SYY_IVOL	0.0938	0.1148	SI	0.1063	0.1279	H52	–0.1390	–0.1627
	(1.11)	(0.98)		(0.97)	(1.18)		(1.12)	(1.36)		(–1.50)	(–1.77)

Table 4
Fama-Macbeth regression analyses
This table reports results from Fama-Macbeth (1973) regressions of monthly stock returns on the SEP sentiment measure (*G*). The control variables are the log of market capitalization *FirmSize*, the Amihud (2002) *Illiquidity* measure, the short interest normalized by the outstanding shares (*ShortInterestRatio*), the current price to the maximum price over the past 52 weeks (*52WeekHigh*), idiosyncratic volatility (*IVOL*), the Stambaugh, Yu, and Yuan (2015) mispricing measure (*Mispricing*), and Google trends search volume (*SVI*). The unit of observation is a stock-month. For all models except (7) and (10) the sample consists of the reference stocks of SEP issues during the period running from 2004 through 2018. For models (7) and (10) the sample is limited by the availability of mispring measure and ends with December 2016. Newey-West *t*-statistics are in parentheses below the average coefficient estimates.

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)	Model (7)	Model (8)	Model (9)	Model (10)
Intercept	0.0071	–0.0116	0.0081	0.0094	–0.0045	0.0092	0.0165	0.0077	0.0066	0.0294
	(1.62)	(–0.43)	(1.87)	(2.38)	(–0.42)	(2.66)	(2.56)	(1.79)	(1.55)	(0.57)
G	–0.1699	–0.1476	–0.2304	–0.1479	–0.1480	–0.1441	–0.2347	–0.1866	–0.2024	–0.1652
	(–3.38)	(–3.12)	(–4.07)	(–2.76)	(–3.41)	(–3.00)	(–3.23)	(–3.82)	(–4.06)	(–2.59)
FirmSize		0.0008								–0.0006
		(0.74)								(-0.34)
Illiquidity			–15.965							29.7014
			(–1.62)							(1.20)
ShortInterestRatio				–0.4298						–2.1129
				(–0.48)						(–2.29)
52WeekHigh					0.0127					–0.0076
					(1.21)					(–0.46)
IVOL						–0.1549				0.1597
						(–0.97)				(0.66)
Mispricing							–0.0211			–0.0102
							(–1.29)			(-0.89)
SVI								–0.0039		0.0166
								(–0.49)		(0.91)
BetaBW									–0.0476	–0.0792
									(–3.13)	(–2.82)
Number of obs.	16,046	16,044	16,044	13,834	16,046	16,044	11,055	16,046	15,401	11,050

Table 5
Results of regressions of portfolio excess returns based on lagged sentiment
This table reports results of regressions in which excess returns during months $t+2$ and $t+3$ of portfolios formed using the month t sentiment measures $G_{t,n}$ are regressed on the factor returns of the Fama-French (2015) five-factor model. For the results in Panel A, the dependent variable is the excess return $R_{t+2} - r_f$, where $R_{t+2} = \sum_{n=1}^{N} G_{t,n} r_{t+2,n}$ is computed using the sentiment measures from month t and the returns from month $t+2$ and r_f is the risk-free rate. For the results in Panel B, the dependent variable is the excess return $R_{t+3} - r_f$, where $R_{t+3} = \sum_{n=1}^{N} G_{t,n} r_{t+3,n}$ is computed using the sentiment measures from month t and the returns from month $t+3$. The left two columns of each panel present the coefficient estimates and corresponding t-statistics using the full sample running from January 2004 through December 2019. The right two columns report the results using a sample beginning in January 2006.

	Sample period: 2004–2019		Sample period: 2006–2019	
Variable	Coefficient estimate	t-statistic	Coefficient estimate	t-statistic
Panel A: Dependent variable $R_{t+2} - r_f = \sum_{n=1}^{N} G_{t,n} r_{t+2,n} - r_f$				
Intercept	**–0.0042**	**(–1.41)**	**–0.0064**	**(–1.99)**
MKTRF	1.3626	(16.18)	1.3719	(16.08)
SMB	0.4647	(3.34)	0.4354	(2.98)
HML	0.0474	(0.35)	0.0968	(0.66)
RMW	–0.3711	(–1.76)	–0.1403	(–0.62)
CMA	–0.3124	(–1.31)	–0.3861	(–1.50)
Panel B: Dependent variable $R_{t+3} - r_f = \sum_{n=1}^{N} G_{t,n} r_{t+3,n} - r_f$				
Intercept	**–0.0003**	**(–0.09)**	**–0.0028**	**(–0.91)**
MKTRF	1.4149	(16.34)	1.4463	(17.42)
SMB	0.2482	(1.74)	0.1380	(0.97)
HML	0.1933	(1.37)	0.0864	(0.61)
RMW	–0.3008	(–1.38)	–0.4278	(–1.93)
CMA	–0.2154	(–0.88)	0.0232	(0.09)

Table 6
Results of regressions that use sentiment measures to explain future three and six-month aggregate stock market returns
This table reports the results of regressions in which a moving average of SEP issuances MA(Q), the BW sentiment index *SENT*, and the othogonalized BW sentiment index *SENT*⊥ are used to explain future three and six-month returns on the CRSP value-weighted index. The variable MA(Q) = $(Q_t + Q_{t-1} + Q_{t-2})/3$ is the three-month moving average of aggregate SEP proceeds, where Q_t is the sum across stocks of SEP proceeds during month t. *SENT* and *SENT*⊥ are the BW index and the orthogonalized BW index during month t. The dependent variable in Panel A (B) is the three (six)-month return on the CRSP value-weighted index over months $t + 1$ through $t + 3$ ($t + 6$). The models in columns (1)–(5) of Panel A (B) are estimated using monthly observations of overlapping three (six)-month returns; those in columns (6)–(10) use quarterly (semi-annual) observations of three (six)-month returns. All models are estimated using SEP data starting in January 2004; thus for example in model (1) the first observation consists of MA(Q) in March 2004 (computed from SEP issuances in January–March 2004) and the three (six)-month return over April–June (September) 2004. Models (1) and (6) use return data through December 2019; thus for example in Panel A (B) model (1) the last observation consists of MA(Q) in September (June) 2019 and the three (six)-month return over October (July)–December 2019. In models (2)–(5) and (7)–(10) the last observation of the dependent variable is the return over January-March (June) 2019 because SENT and *SENT*⊥ are only available through December 2018. In Panel A (B) the t-statistics for models (1)–(5) are computed from Newey-West standard errors using three (six) lags.

Panel A. Regression models that explain three-month returns

	Monthly observations (overlapping three-month returns)					Quarterly observations (non-overlapping returns)				
Covariate	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
MA(Q)	−0.0903			−0.0837	−0.0874	−0.0886			−0.0899	−0.0856
	(−2.79)			(−2.58)	(−2.63)	(−2.39)			(−2.00)	(−1.98)
SENT		−0.0553		−0.0169			−0.0415		0.0003	
		(−1.73)		(−0.63)			(−1.11)		(0.01)	
SENT⊥			−0.0363		−0.0086			−0.0338		−0.0068
			(−1.32)		(−0.33)			(−1.18)		(−0.22)
intercept	0.0760	0.0128	0.0222	0.0692	0.0744	0.0753	0.0153	0.0221	0.0761	0.0733
	(3.98)	(1.21)	(2.56)	(3.46)	(4.16)	(3.23)	(1.27)	(2.23)	(2.34)	(2.65)
R^2	0.0914	0.0376	0.0265	0.0948	0.0933	0.0857	0.0209	0.0233	0.0853	0.0861
No.of obs.	187	178	178	178	178	63	60	60	60	60

Table 6 (continued)

Panel B. Regression models that explain six-month returns

Covariate	Monthly observations (overlapping six-month returns) (1)	(2)	(3)	(4)	(5)	Semi-annual observations (non-overlapping returns) (6)	(7)	(8)	(9)	(10)
MA(*Q*)	–0.2445			–0.2366	–0.2426	–0.2325			–0.2545	–0.2558
	(–4.23)			(–4.11)	(–3.57)	(–2.97)			(–2.71)	(–2.81)
SENT		–0.1383		–0.0297			–0.0939		0.0172	
		(–2.10)		(-0.74)			(–1.12)		(0.20)	
SENT⊥			–0.0931		–0.0163			–0.0623		0.0182
			(–2.17)		(–0.42)			(–0.91)		(0.27)
intercept	0.1908	0.0223	0.0457	0.1820	0.1906	0.1907	0.0319	0.0472	0.2099	0.208
	(6.60)	(0.98)	(2.79)	(5.96)	(5.70)	(3.68)	(1.18)	(2.07)	(2.99)	(3.42)
R^2	0.3009	0.106	0.0785	0.3116	0.3009	0.2330	0.0430	0.0284	0.2474	0.2483
No.of obs.	184	178	178	178	178	31	30	30	30	30

Table 7
Tobit regression analyses of the determinants of SEP sentiment

Coefficient estimates and t-statistics from Tobit regressions explaining the SEP sentiment measure. The data consist of monthly observations from the top quintile U.S. stocks as of the end of the previous month. Results for the 2004–2018 (2006–2018) sample are in the left-hand (right-hand) side of the table. For each stock n in month t, the dependent variable is the normalized SEP sentiment, $G_{t,n}$, which is defined in Section 2. It is positive if there is any SEP linked to stock n is issued during month t, and zero otherwise. The independent variables are Google trends search volume at the end of the previous month ($SVI_{t-1,n}$), the Stambaugh, Yu, and Yuan (2015) mispricing measure in the previous month ($SYY\_Misp_{t-1,n}$), the option-to-stock trading volume ratio in the previous month ($O/S_{t-1,n}$), the option implied volatility at end of previous month ($IV_{t-1,n}$), lagged one week return (*Ret*(*w*–2, *w*–1)), the second trailing week return (*Ret*(*w*–3, *w*–2)), lagged one month return (*Ret*(*t*–2, *t*–1)), two-month return over the second and third trailing months (*Ret*(*t*–4, *t*–2)), three-month return from the forth to sixth trailing month (*Ret*(*t*–7, *t*–4)), realized volatility over the preceding month (*RealizedVol*(*t*–2, *t*–1)), realized volatility over the second and third trailing months (*RealizedVol*(*t*–4, *t*–2)), realized volatility over the forth to sixth trailing month (*RealizedVol*(*t*–7, *t*–4)), natural logarithm of market capitalization at end of previous month (ln(*MarketCap*(*t*–1))), calendar month fixed effects, and industry fixed effects based on the Fama-French 12-industry classification. The samples end in 2018 rather than 2019 because the Baker-Wurgler sentiment index is not available for 2019.

	Sample period: 2004–2018					Sample period: 2006–2018				
Explanatory variable	1	2	3	4	5	6	7	8	9	10
$SVI_{t-1,n}$		0.0001		0.0001	0.0001		0.0002		0.0001	0.0002
		(3.68)		(2.65)	(3.28)		(5.33)		(4.08)	(4.72)
$SYY\_Misp_{t-1,n}$			0.0028	0.0028	0.0025			0.0026	0.0026	0.0024
			(26.65)	(26.64)	(24.28)			(27.10)	(27.08)	(24.75)
$O/S_{t-1,n}$					0.0008					0.0007
					(23.78)					(22.75)
$BWt–1$	0.0013	0.0013	0.0012	0.0012	0.0012	0.0010	0.0010	0.0011	0.0011	0.0010
	(28.60)	(28.61)	(27.42)	(27.42)	(26.59)	(28.78)	(28.78)	(26.49)	(26.50)	(25.78)
$IV_{t-1,n}$	0.0016	0.0016	0.0014	0.0014	0.0013	0.0010	0.0010	0.0009	0.0009	0.0009
	(12.51)	(12.49)	(11.32)	(11.31)	(10.49)	(9.83)	(9.78)	(8.28)	(8.24)	(7.65)
Ret(*w*–2, *w*–1)	–0.0030	–0.0030	–0.0028	–0.0028	–0.0026	–0.0025	–0.0025	–0.0027	–0.0027	–0.0025
	(–12.94)	(–12.92)	(–11.94)	(–11.92)	(–11.2)	(–13.71)	(–13.68)	(–12.45)	(–12.42)	(–11.74)
Ret(*w*–3, *w*–2)	0.0002	0.0002	0.0003	0.0003	0.0003	0.0001	0.0001	0.0002	0.0002	0.0002
	(1.30)	(1.31)	(1.36)	(1.37)	(1.54)	(1.05)	(1.07)	(0.93)	(0.94)	(1.14)

Table 7
Tobit regression analyses of the determinants of SEP sentiment (continued)

	Sample period: 2004–2018					Sample period: 2006–2018				
Explanatory variable:	1	2	3	4	5	6	7	8	9	10
Ret(*t*–2, *t*–1)	–0.0007	–0.0007	–0.0008	–0.0008	–0.0009	–0.0006	–0.0006	–0.0007	–0.0007	–0.0008
	(–5.68)	(–5.66)	(–5.42)	(–5.41)	(–6.10)	(–6.24)	(–6.23)	(–5.60)	(–5.60)	(–6.21)
Ret(*t*–4, *t*–2)	0.0000	0.0000	0.0001	0.0001	0.0000	0.0000	0.0000	0.0001	0.0001	0.0000
	(0.03)	(0.07)	(1.17)	(1.19)	(–0.10)	(–0.23)	(–0.17)	(0.89)	(0.93)	(–0.23)
Ret(*t*–7, *t*–4)	0.0003	0.0003	0.0004	0.0004	0.0003	0.0003	0.0003	0.0004	0.0004	0.0003
	(5.52)	(5.57)	(6.76)	(6.79)	(5.28)	(6.85)	(6.92)	(7.90)	(7.94)	(6.54)
RealizedVol(*t*–2, *t*–1)	0.0117	0.0117	0.0114	0.0114	0.0107	0.0098	0.0098	0.0112	0.0112	0.0106
	(10.49)	(10.48)	(9.77)	(9.77)	(9.20)	(11.40)	(11.39)	(10.46)	(10.45)	(9.85)
RealizedVol(*t*–4, *t*–2)	0.0130	0.0130	0.0108	0.0108	0.0104	0.0101	0.0101	0.0099	0.0099	0.0094
	(10.94)	(10.92)	(9.00)	(9.00)	(8.62)	(11.03)	(11.01)	(8.95)	(8.94)	(8.52)
RealizedVol(*t*–7, *t*–4)	0.0225	0.0224	0.0165	0.0165	0.0159	0.0170	0.0169	0.0141	0.0140	0.0135
	(20.82)	(20.78)	(15.73)	(15.7)	(15.12)	(20.47)	(20.40)	(14.66)	(14.61)	(14.04)
ln(*MarketCap*(*t*–1))	0.0009	0.0009	0.0009	0.0009	0.0008	0.0007	0.0007	0.0008	0.0008	0.0007
	(69.17)	(68.82)	(61.63)	(61.34)	(54.28)	(62.62)	(62.23)	(58.21)	(57.88)	(50.87)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cal. month fixed effs.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. of obs.	156,729	156,729	133,087	133,087	132,901	134,674	134,674	112,136	112,136	111,950

Table 8
Returns in short windows around earnings announcement dates
This table reports the average market-adjusted returns of SEP reference stocks during narrow windows around the first earnings announcement date following the SEP issuances. The market adjustment is done using the return on the CRSP value-weighted index. The earnings announcements are categorized by their distance from the SEP issuances: 15 days, 30 days, and 60 days after the SEP issuances. The market-adjusted returns over days [−1, +1] and [−1, +2] are reported for each category, where the earnings announcement date is day 0. *t*-statistics are in parentheses below the average returns.

Earnings announcement category	Mkt.-adj. return over days [−1,1]	Mkt.-adj. return over days [−1,2]	No. of obs.
First earnings announcement within 15 days after SEP issuance	−0.0036 (−2.64)	−0.0032 (−2.28)	3,432
First earnings announcement within 30 days after SEP issuance	−0.0025 (−2.23)	−0.0022 (−1.86)	5,213
First earnings announcement within 60 days after SEP issuance	−0.0021 (−2.21)	−0.0016 (−1.68)	7,401

Table 9
Returns of long and short portfolios formed using predicted SEP sentiment
This table reports the abnormal returns based on the Fama-French five-factor model on the long and short portfolios formed using predicted SEPs sentiment. The predicted SEP sentiment measures are computed using the Tobit regression models for which results are reported in Table 7. The last three columns repeat the first three but using actual SEP sentiment for stock *n* if there is there is a SEP issue based on stock *n*, and the predicted sentiment otherwise. The portfolios in Panel A include only the 50 stocks with the largest predicted sentiment measures while the portfolios in Panel B include all stocks. The sample period is 2004–2018.

		Portfolios formed using predicted SEP sentiment			Portfolios formed using actual SEP sentiment if it is available and predicted value otherwise		
Panel A: Portfolios formed using SEP sentiment estimated from Tobit regression models in Table 7 (≤ 50 stocks)							
		Long	Short	Long-Short	Long	Short	Long-Short
Model 1	Intercept	0.0041	0.0011	0.0042	0.0037	–0.0062	0.0087
	t-statistic	(2.55)	(0.14)	(0.61)	(2.32)	(–1.94)	(2.29)
Model 2	Intercept	0.0041	0.0020	0.0033	0.0038	–0.0063	0.0088
	t-statistic	(2.53)	(0.25)	(0.48)	(2.37)	(–1.95)	(2.32)
Model 3	Intercept	0.0047	–0.0068	0.0110	0.0047	–0.0068	0.0115
	t-statistic	(3.09)	(–0.95)	(1.66)	(3.00)	(–2.25)	(2.91)
Model 4	Intercept	0.0046	–0.0071	0.0110	0.0047	–0.0068	0.0116
	t-statistic	(3.02)	(–1.10)	(1.64)	(2.98)	(–2.26)	(2.92)
Model 5	Intercept	0.0047	0.0026	0.0018	0.0049	–0.0058	0.0104
	t-statistic	(2.91)	(0.36)	(0.27)	(2.98)	(–1.92)	(2.60)
Panel B: Portfolios formed using SEP sentiment estimated from Tobit regression models in Table 7 (all stocks)							
		Long	Short	Long-Short	Long	Short	Long-Short
Model 1	Intercept	0.0010	0.0011	0.0011	0.0012	–0.0058	0.0061
	t-statistic	(1.99)	(0.14)	(0.16)	(2.41)	(–1.87)	(1.91)
Model 2	Intercept	0.0010	0.0020	0.0003	0.0012	–0.0059	0.0062
	t-statistic	(2.00)	(0.25)	(0.04)	(2.41)	(–1.88)	(1.92)
Model 3	Intercept	0.0013	–0.0068	0.0076	0.0016	–0.0068	0.0084
	t-statistic	(2.47)	(–0.95)	(1.17)	(2.91)	(–2.29)	(2.46)
Model 4	Intercept	0.0013	–0.0071	0.0076	0.0016	–0.0068	0.0084
	t-statistic	(2.47)	(–1.10)	(1.17)	(2.92)	(–2.29)	(2.46)
Model 5	Intercept	0.0013	0.0026	–0.0016	0.0015	–0.0061	0.0074
	t-statistic	(2.47)	(0.36)	(–0.49)	(2.81)	(–1.85)	(1.97)

Internet Appendix to

Retail Derivatives and Sentiment:

A Sentiment Measure Constructed from Issuances of Retail Structured Equity Products

Brian J. Henderson,[a] Neil D. Pearson,[b, c] and Li Wang[d]

August 15, 2020

[a] George Washington University, Funger Hall #502, 2201 G Street NW, Washington, DC 20052. Tel.: +1 (202) 994-3669. E-mail: bjhndrsn@gwu.edu.

[b] Department of Finance, University of Illinois at Urbana-Champaign, 1206 South Sixth Street, Champaign, Illinois 61820. Tel.: +1 (217) 244-0490. E-mail: pearson2@illinois.edu.

[c] CDI Research Fellow, Canadian Derivatives Institute, 3000, Chemin de la Côte-Sainte-Catherine, Montréal, Québec H3T 2A7, Canada.

[d] Department of Banking and Finance, Weatherhead School of Management, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, Ohio 44106. Tel.: +1 (216) 368-0802. E-mail: lxw429@case.edu.

* Corresponding author: Department of Finance, University of Illinois at Urbana-Champaign, 1206 South Sixth Street, Champaign, Illinois 61820. Tel.: +1 (217) 244-0490. Fax: +1 (217) 244-3102.

E-mail: pearson2@illinois.edu (N.D. Pearson).

This internet appendix provides additional results that supplement the results in the main text.

Table IA1 reports the results of calendar-time portfolio return regressions in which we form equal-weighted portfolios of SEP stocks in each month and analyze the returns of those equal-weighted portfolios. These regressions are similar to those using the returns of value-weighted portfolios reported in Tables 1–3 except that they analyze equal-weighted rather than value-weighted returns.

Table IA2 reports the factor loadings for the calendar-time portfolio return regressions for which only intercepts and selected factor loadings are reported in Table 3.

In the results discussed in the main text we use the SEP sentiment measure based on issuances during month t to predict returns during the subsequent three months. Because the SEP issues occur throughout calendar month t, analysis of the subsequent monthly returns does not capture the return from the SEPs issue dates through the end of month t. The best way to capture these intra-month returns is not obvious, precisely because the SEP issues are occur throughout the issue month. Here we outline the approach we use to measure and benchmark returns to SEPs sentiment stocks for the remaining portion of the calendar month following their issuance.

Let h be the number of trading dates in month t, and use $t - j/h$ to denote the trading date in month t that is j trading dates prior to the end of month t. Thus $t - 0/h = t$ is the last trading date of month t. Consider SEP n that is priced and sold j_n dates prior to the month-end, on date $t - j_n/h$. We want to measure the return during the part of month subsequent to the issue. Henderson, Pearson, and Wang (2020) find that SEPs issuers' delta-hedge trades impact underlying stock prices on the pricing date $t - j_n/h$., followed by reversals on the next trading date $t - (j_n - 1)/h$. Because we do not want the returns to include these reversals on date $t - (j_n -$

$1)/h$, we cumulate the daily returns over the dates from $t - (j_n - 2)/h$ through the last trading date t. Specifically, we compute the daily abnormal returns based on the Fama-French (1993) three-factor model estimate using data from the previous year and then sum them over the trading dates from $t - (j_n - 2)/h$ through the last trading date t. Then because the intra-month returns subsequent to different SEP issues cover different periods, we scale each of them up to a monthly interval by multiplying by $h/(j_n - 2)$. SEPs priced on one of the last two trading dates (for which $j_n - 2$ is either 0 or −1) are not included in this calculation.

In Table IA3 we report the average SEP issue size-weighted and equal-weighted average abnormal returns scaled to the monthly horizon for the two sample periods 2004–2019 and 2006–2019. In the average SEP issue size-weighted returns, for SEP n the abnormal return from $t - (j_n - 2)/h$ to t is weighted by the ratio $Q_{t,n}/V_{t-1,n}$, scaled by the sum of the same ratios of all SEPs included in the return calculations for the month. In the average equal-weighted returns the equal-weighting is by SEP issue, where for each SEP issue the abnormal return of the reference stock is included with same weight.[1] The point estimates of SEP issue size-weighted returns of −101 and −136 basis points are consistent with the abnormal returns reported in Table 1, although the −101 basis point estimate for the 2004–2019 sample period is insignificant and the −136 basis point estimate for the 2006–2019 sample period is significant at only the 10% level. The estimates of the equal-weighted abnormal returns are significant at either the 10% or 1% level and are consistent with the equal-weighted abnormal returns reported in Table A1. Thus, even though not all of the estimates of average returns in Table 6 are significant, the point estimates are consistent with those reported for month $t + 1$ in Tables 2 and IA1.

[1] Thus, the equal-weighting by SEP issue differs from the equal-weighting by stock used in Table A1. With equal-weighting by SEP issue the weight on each stock is proportional to the number of issues, while in the returns used in Table A1 each stock receives equal weight.

Table IA4 presents results of panel Tobit regression models using $G_{t,n}$ as the dependent variable in place of the variable $Q_{t,n}/V_{t-1}$,n used in the Tobit regressions reported in Table 7. The regressions for which results are reported in Table IA4 also do not include the Baker-Wurgler index as a covariate because, due to the normalization used in constructing $G_{t,n}$, there is no important time-series variation in $G_{t,n}$ that might be explained by the Baker-Wurgler index.

Table IA5 presents the average returns and average abnormal returns on the SEP reference stocks during each of the six months prior to the issuance month.

Table IA1
Abnormal returns on equal-weighted portfolios of SEPs' reference stocks

This table reports results of regressions in which monthly excess returns of equal-weighted portfolios formed from SEPs' reference stocks are regressed on the factor returns of various factor models. The regression models are those for which results are reported in Tables 1 and 2 of the main text, except that here the dependent variables are the excess returns on equal-weighted rather than SEP sentiment value-weighted portfolios. The left two columns of each panel present the coefficient estimates and corresponding *t*-statistics over the full sample period running from January 2004 through December 2019. The right two columns report the results using a sample beginning in January 2006. Returns for the Stambaugh, Yu, and Yuan (2017) and Daniel, Hirshleifer, and Sun (2020) factors are available only through 2016 and 2018, respectively.

	Sample period: 2004–2019		Sample period: 2006–2019	
Variable	Coeff. est.	*t*-stat.	Coeff. est.	*t*-stat.
Panel A: Fama-French (1993) three-factor model				
Intercept	**–0.0050**	**(–3.38)**	**–0.0054**	**(–3.42)**
MKTRF	1.2968	(32.12)	1.2907	(31.90)
SMB	0.2430	(3.50)	0.1816	(2.50)
HML	–0.0937	(–1.55)	–0.0951	(–1.54)
Panel B: Carhart (1997) four-factor model				
Intercept	**–0.0050**	**(–3.31)**	**–0.0053**	**(–3.40)**
MKTRF	1.2866	(30.71)	1.2686	(30.55)
SMB	0.2455	(3.54)	0.1789	(2.48)
HML	–0.1162	(–1.78)	–0.1486	(–2.23)
MOM	–0.0353	(–0.92)	–0.0793	(–2.02)
Panel C: Fama-French (2015) five-factor model				
Intercept	**–0.0042**	**(–2.68)**	**–0.0046**	**(–2.79)**
MKTRF	1.2676	(29.25)	1.2631	(28.97)
SMB	0.2044	(2.86)	0.1457	(1.95)
HML	–0.0856	(–1.21)	–0.0898	(–1.20)
RMW	–0.2235	(–2.05)	–0.2237	(–1.92)
CMA	–0.0217	(–0.18)	–0.0278	(–0.21)
Panel D: Hou, Xue, and Zhang (2015) q-factor model				
Intercept	**–0.0040**	**(–2.87)**	**–0.0037**	**(–2.52)**
MKTRF	1.2598	(30.77)	1.2376	(30.4)
SMB	0.2001	(3.15)	0.1445	(2.15)
MGMT	–0.2095	(–2.26)	–0.2439	(–2.60)
PERF	–0.1139	(–1.50)	–0.1615	(–2.08)
Panel E: Hou et al. (2020) q5-factor model				
Intercept	**–0.0031**	**(–2.15)**	**–0.0025**	**(–1.69)**
MKT	1.2341	(29.43)	1.2074	(29.16)
ME	0.1662	(2.58)	0.1079	(1.60)
IA	–0.2596	(–2.76)	–0.3080	(–3.24)
ROE	–0.0579	(–0.73)	–0.0938	(–1.17)
EG	–0.2327	(–2.34)	–0.2728	(–2.75)

Table IA1 (Continued)
Abnormal returns on equal-weighted portfolios of SEPs' reference stocks

	Sample period: 2004–2019		Sample period: 2006–2019	
Variable	Coeff. est.	*t*-stat.	Coeff. est.	t-stat.
Panel F: Stambaugh and Yuan (2017) factor model				
Intercept	**–0.0042**	**(–2.58)**	**–0.0034**	**(–2.01)**
MKT	1.2662	(25.77)	1.2382	(26.04)
ME	0.1959	(2.48)	0.0961	(1.15)
IA	–0.4948	(–6.28)	–0.5673	(–7.31)
ROE	–0.0369	(–0.90)	–0.0771	(–1.89)
Panel G: Daniel, Hirshleifer, and Sun (2020) factor model				
Intercept	**–0.0028**	**(–1.89)**	**–0.0028**	**(–1.87)**
MKTRF	1.2178	(30.71)	1.1859	(30.59)
FIN	–0.4087	(–7.18)	–0.4207	(–7.24)
PEAD	–0.1522	(–2.05)	–0.1944	(–2.68)

Table IA2
Full results for Table 3
This table reports the estimates of all of the factor loadings for the regression models for which only the estimates of the intercepts and loadings on the additional factors are reported in Table 3 of the main text.

	FF_IVOL (High–Low)			SYY_IVOL (High–Low)			SI (High–Low)			H52 (High–Low)	
	2004–2019	2006–2019		2004–2019	2006–2019		2004–2019	2006–2019		2004–2019	2006–2019
Panel A: Single-factor model with additional factor											
Variable	2004-2019	2006-2019		2004-2019	2006-2019		2004-2019	2006-2019		2004-2019	2006-2019
Intercept	–0.0113	–0.0124	Intercept	–0.0112	–0.0120	Intercept	–0.0115	–0.0125	Intercept	–0.0121	–0.0131
	(–3.73)	(–3.80)		(–3.68)	(–3.70)		(–3.67)	(–3.75)		(–3.99)	(–4.08)
MKTRF	1.3173	1.2947	MKTRF	1.3345	1.2951	MKTRF	1.3918	1.3493	MKTRF	1.3478	1.2961
	(13.68)	(13.44)		(14.61)	(14.20)		(14.56)	(14.23)		(14.25)	(13.68)
FF_IVOL	0.2560	0.2541	SYY_IVOL	0.2821	0.3029	SI	0.2262	0.2504	H52	–0.2518	–0.2742
	(3.46)	(3.36)		(3.61)	(3.78)		(2.40)	(2.65)		(–3.14)	(–3.45)
Panel B: Fama-French (1993) three-factor model with additional factor											
Intercept	–0.0115	–0.0128	Intercept	–0.0114	–0.0123	Intercept	–0.0119	–0.0131	Intercept	–0.0121	–0.0133
	(–3.77)	(–3.92)		(–3.71)	(–3.78)		(–3.78)	(–3.92)		(–4.03)	(–4.2)
MKTRF	1.3333	1.3158	MKTRF	1.3436	1.3096	MKTRF	1.4033	1.3684	MKTRF	1.3266	1.2832
	(13.55)	(13.36)		(14.05)	(13.62)		(14.43)	(14.17)		(13.76)	(13.46)
FF_IVOL	0.2315	0.2146	SYY_IVOL	0.2534	0.2605	SI	0.1668	0.1853	H52	–0.2469	–0.2710
	(2.70)	(2.49)		(2.76)	(2.78)		(1.54)	(1.70)		(–2.98)	(–3.3)
HML	–0.1173	–0.1500	HML	–0.0634	–0.0875	HML	–0.1298	–0.1597	HML	–0.1981	–0.2396
	(–0.95)	(–1.19)		(–0.50)	(–0.68)		(–1.03)	(–1.25)		(–1.59)	(–1.9)
SMB	0.1023	0.1809	SMB	0.0902	0.1418	SMB	0.1976	0.2332	SMB	0.2367	0.2726
	(0.63)	(1.07)		(0.55)	(0.83)		(1.20)	(1.34)		(1.65)	(1.81)

Table IA2 (Continued)
Full results for Table 3

Panel C: Carhart (1997) four-factor model with additional factor											
Intercept	–0.0115	–0.0128	Intercept	–0.0113	–0.0123	Intercept	–0.0119	–0.0132	Intercept	–0.0118	–0.0125
	(–3.74)	(–3.89)		(–3.69)	(–3.74)		(–3.78)	(–3.94)		(–3.97)	(–4)
MKTRF	1.3324	1.3160	MKTRF	1.3438	1.3092	MKTRF	1.3994	1.3627	MKTRF	1.2640	1.2086
	(13.49)	(13.28)		(14.01)	(13.55)		(14.30)	(14.05)		(12.72)	(12.15)
FF_IVOL	0.2444	0.2134	SYY_IVOL	0.2661	0.2657	SI	0.1535	0.1617	H52	–0.4585	–0.5064
	(2.57)	(2.17)		(2.63)	(2.49)		(1.36)	(1.42)		(–3.63)	(–3.85)
HML	–0.1002	–0.1515	HML	–0.0445	–0.0801	HML	–0.1523	–0.2004	HML	–0.1003	–0.1275
	(–0.74)	(–1.08)		(–0.32)	(–0.54)		(–1.13)	(–1.45)		(–0.77)	(–0.96)
SMB	0.0869	0.1820	SMB	0.0756	0.1370	SMB	0.2118	0.2518	SMB	0.1342	0.1793
	(0.51)	(1.04)		(0.44)	(0.77)		(1.26)	(1.43)		(0.90)	(1.17)
MOM	0.0260	–0.0022	MOM	0.0246	0.0088	MOM	–0.0356	–0.0606	MOM	0.2446	0.2682
	(0.32)	(–0.03)		(0.30)	(0.10)		(–0.46)	(–0.77)		(2.20)	(2.27)
Panel D: Fama-French (2015) five-factor model with additional factor											
Intercept	–0.0110	–0.0126	Intercept	–0.0107	–0.0120	Intercept	–0.0109	–0.0123	Intercept	–0.0116	–0.0131
	(–3.47)	(–3.72)		(–3.37)	(–3.55)		(–3.31)	(–3.54)		(–3.68)	(–3.97)
MKTRF	1.3235	1.3061	MKTRF	1.3291	1.2940	MKTRF	1.3709	1.3390	MKTRF	1.3172	1.2781
	(13.15)	(12.95)		(13.47)	(13.02)		(13.43)	(13.18)		(13.24)	(12.97)
FF_IVOL	0.2355	0.2224	SYY_IVOL	0.2600	0.2772	SI	0.1643	0.1896	H52	–0.2422	–0.2717
	(2.55)	(2.39)		(2.66)	(2.78)		(1.48)	(1.69)		(–2.79)	(–3.16)
HML	–0.1732	–0.2225	HML	–0.1215	–0.1704	HML	–0.1613	–0.2124	HML	–0.2387	–0.2993
	(–1.23)	(–1.48)		(–0.86)	(–1.14)		(–1.13)	(–1.39)		(–1.66)	(–1.97)
SMB	0.0769	0.1509	SMB	0.0579	0.0983	SMB	0.1583	0.1915	SMB	0.2155	0.2530
	(0.46)	(0.87)		(0.35)	(0.56)		(0.94)	(1.07)		(1.47)	(1.65)
RMW	–0.1181	–0.1435	RMW	–0.1464	–0.1781	RMW	–0.2555	–0.2637	RMW	–0.1317	–0.1226
	(–0.52)	(–0.59)		(–0.65)	(–0.75)		(–1.15)	(–1.10)		(–0.59)	(–0.51)
CMA	0.2139	0.2363	CMA	0.2284	0.2838	CMA	0.1203	0.1555	CMA	0.1592	0.1920
	(0.83)	(0.84)		(0.89)	(1.00)		(0.47)	(0.55)		(0.63)	(0.71)

Table IA2 (Continued)
Full results for Table 3

Panel E: Hou, Xue, and Zhang (2015) q-factor model with additional factor											
Intercept	–0.0116	–0.0120	Intercept	–0.0115	–0.0118	Intercept	–0.0112	–0.0114	Intercept	–0.0120	–0.0121
	(–3.83)	(–3.77)		(–3.79)	(–3.71)		(–3.65)	(–3.58)		(–4.03)	(–3.9)
MKTRF	1.3589	1.3317	MKTRF	1.3696	1.3285	MKTRF	1.3680	1.3196	MKTRF	1.3392	1.2825
	(14.07)	(13.98)		(14.66)	(14.36)		(14.49)	(14.24)		(14.04)	(13.66)
FF_IVOL	0.1452	0.0890	SYY_IVOL	0.1526	0.1251	SI	0.1448	0.1420	H52	–0.1952	–0.2115
	(1.39)	(0.82)		(1.40)	(1.08)		(1.34)	(1.31)		(–1.94)	(–2.08)
ME	0.1504	0.2789	ME	0.1513	0.2559	ME	0.1614	0.2359	ME	0.2138	0.2882
	(0.95)	(1.69)		(0.96)	(1.54)		(1.04)	(1.42)		(1.55)	(1.97)
IA	–0.2553	–0.4212	IA	–0.2279	–0.3709	IA	–0.3123	–0.4431	IA	–0.3373	–0.4646
	(–1.15)	(–1.81)		(–1.00)	(–1.52)		(–1.47)	(–2.03)		(–1.61)	(–2.18)
ROE	–0.0182	–0.0946	ROE	–0.0293	–0.0747	ROE	–0.1060	–0.1348	ROE	0.0512	0.0541
	(–0.10)	(–0.47)		(–0.16)	(–0.38)		(–0.65)	(–0.81)		(0.27)	(0.27)
Panel E: Hou et al. (2020) q5-factor model with additional factor											
Intercept	–0.0098	–0.0102	Intercept	–0.0097	–0.0101	Intercept	–0.0089	–0.0093	Intercept	–0.0101	–0.0103
	(–3.21)	(–3.19)		(–3.16)	(–3.13)		(–2.85)	(–2.86)		(–3.33)	(–3.26)
MKTRF	1.3110	1.2869	MKTRF	1.3133	1.2777	MKTRF	1.2852	1.2444	MKTRF	1.2750	1.2229
	(13.47)	(13.42)		(13.84)	(13.61)		(13.21)	(13.03)		(13.13)	(12.78)
FF_IVOL	0.1075	0.0484	SYY_IVOL	0.1273	0.0960	SI	0.1766	0.1705	H52	–0.1901	–0.2083
	(1.03)	(0.45)		(1.18)	(0.83)		(1.66)	(1.6)		(–1.92)	(–2.09)
ME	0.0905	0.2251	ME	0.0791	0.1944	ME	0.0384	0.1239	ME	0.1220	0.2053
	(0.57)	(1.37)		(0.50)	(1.17)		(0.24)	(0.73)		(0.87)	(1.39)
IA	–0.3045	–0.4796	IA	–0.2728	–0.4249	IA	–0.3304	–0.4617	IA	–0.3626	–0.4914
	(–1.39)	(–2.08)		(–1.21)	(–1.76)		(–1.59)	(–2.16)		(–1.76)	(–2.34)
ROE	0.0731	–0.0114	ROE	0.0799	0.0248	ROE	0.0539	0.0193	ROE	0.1811	0.1814
	(0.39)	(–0.06)		(0.43)	(0.13)		(0.32)	(0.11)		(0.93)	(0.9)
EG	–0.5665	–0.5456	EG	–0.5774	–0.5409	EG	–0.6483	–0.6048	EG	–0.5954	–0.5571
	(–2.35)	(–2.23)		(–2.42)	(–2.23)		(–2.72)	(–2.52)		(–2.52)	(–2.35)

Table IA2 (Continued)
Full results for Table 3

Panel G: Stambaugh and Yuan (2017) mispricing factor model with additional factor											
Intercept	–0.0111	–0.0114	Intercept	–0.0109	–0.0113	Intercept	–0.0104	–0.0107	Intercept	–0.0117	–0.0122
	(–3.60)	(–3.55)		(–3.55)	(–3.52)		(–3.36)	(–3.31)		(–3.83)	(–3.81)
MKT	1.3423	1.3178	MKT	1.3441	1.3069	MKT	1.3230	1.2728	MKT	1.3213	1.2696
	(13.44)	(13.51)		(13.73)	(13.60)		(13.41)	(13.3)		(13.66)	(13.46)
FF_IVOL	0.0849	0.0011	SYY_IVOL	0.0990	0.0447	SI	0.1490	0.1551	H52	–0.1971	–0.2112
	(0.79)	(0.01)		(0.88)	(0.38)		(1.41)	(1.48)		(–1.71)	(–1.79)
SMB	0.1288	0.2405	SMB	0.1202	0.2151	SMB	0.0942	0.1494	SMB	0.1457	0.2108
	(0.78)	(1.41)		(0.72)	(1.23)		(0.58)	(0.88)		(0.97)	(1.32)
MGMT	–0.4893	–0.6753	MGMT	–0.4737	–0.6246	MGMT	–0.5574	–0.6438	MGMT	–0.4510	–0.5229
	(–2.55)	(–3.41)		(–2.43)	(–3.09)		(–3.72)	(–4.30)		(–2.70)	(–3.07)
PERF	–0.0876	–0.1369	PERF	–0.0934	–0.1209	PERF	–0.0875	–0.0921	PERF	0.0080	0.0161
	(–0.95)	(–1.41)		(–1.08)	(–1.35)		(–1.06)	(–1.1)		(0.07)	(0.14)
Panel H: Daniel, Hirshleifer, and Sun (2020) behavioral factor model with additional factor											
Intercept	–0.0094	–0.0102	Intercept	–0.0094	–0.0102	Intercept	–0.0090	–0.0097	Intercept	–0.0095	–0.0103
	(–3.11)	(–3.21)		(–3.12)	(–3.20)		(–2.95)	(–3.04)		(–3.18)	(–3.27)
MKTRF	1.2863	1.2528	MKTRF	1.3023	1.2539	MKTRF	1.2861	1.2366	MKTRF	1.2697	1.2157
	(13.62)	(13.4)		(14.46)	(14.00)		(13.64)	(13.3)		(13.58)	(13.04)
FF_IVOL	0.0958	0.0852	SYY_IVOL	0.0938	0.1148	SI	0.1063	0.1279	H52	–0.1390	–0.1627
	(1.11)	(0.98)		(0.97)	(1.18)		(1.12)	(1.36)		(–1.50)	(–1.77)
FIN	–0.4303	–0.4668	FIN	–0.4277	–0.4376	FIN	–0.4798	–0.5020	FIN	–0.4614	–0.4813
	(–3.13)	(–3.27)		(–2.95)	(–2.92)		(–4.03)	(–4.04)		(–3.84)	(–3.85)
PEAD	–0.3047	–0.3594	PEAD	–0.3030	–0.3449	PEAD	–0.3024	–0.3402	PEAD	–0.2028	–0.2227
	(–1.95)	(–2.29)		(–1.92)	(–2.18)		(–1.93)	(–2.16)		(–1.14)	(–1.24)

Table IA3
Intra-month returns on SEP reference stocks

This table reports the average abnormal returns on the SEP reference stocks from the second day after the pricing date through the end of the month. For SEP issue n in month t, we compute the daily abnormal returns on the reference stock using the Fama-French three-factor model and then cumulate them over the dates from $t - (j_n - 2)/h$ through the last trading date of month t, where h is the number of trading dates in the month, the pricing date of the nth SEP is j_n trading dates prior to the last trading date of the month, $t - j_n/h$ is the pricing date of the nth SEP, and $t - (j_n - 2)/h$ is the date that is two trading days after the SEP pricing date. Each return is then scaled up to a monthly interval by multiplying by $h/(j_n - 2)$, where the returns for which $j_n - 2$ is 0 or −1 are not used. In the average SEP issue size-weighted returns, for SEP n the reference stock abnormal return from $t - (j_n - 2)/h$ to t is weighted by the ratio $Q_{t,n}/V_{t-1,n}$, scaled by the sum of the same ratios of all SEPs included in the return calculations for the month. In the average equal-weighted returns the equal-weighting is by SEP issue, so the reference stock return underlying each SEP issue receives the same weight.

	Sample period 2004–2019	Sample period 2006–2019
SEP issue size-weighted returns	−0.0101 (−1.44)	−0.0136 (−1.77)
Equal-weighted returns	−0.0050 (−1.71)	−0.0073 (−2.34)

Table IA4
Tobit regression analysis of the determinants of SEP sentiment

Coefficient estimates and *t*-statistics from Tobit regressions explaining the SEP sentiment measure $G_{t,n}$. The data consist of monthly observations from the top quintile U.S. stocks as of the end of the previous month. Results for the 2004–2019 (2006–2019) sample are in the left-hand (right-hand) side of the table. For each stock *n* in month *t*, the dependent variable is the normalized SEP sentiment, $G_{t,n}$, which is defined in Section 2. It is positive if there is any SEP linked to stock *n* is issued during month *t*, and zero otherwise. The independent variables are Google trends search volume at the end of the previous month ($SVI_{t-1,n}$), the Stambaugh, Yu, and Yuan (2015) mispricing measure in the previous month ($SYY\_Misp_{t-1,n}$), the option-to-stock trading volume ratio in the previous month ($O/S_{t-1,n}$), the option implied volatility at end of previous month ($IV_{t-1,n}$), lagged one week return (*Ret*(*w*–2, *w*–1)), the second trailing week return (*Ret*(*w*–3, *w*–2)), lagged one month return (*Ret*(*t*–2, *t*–1)), two-month return over the second and third trailing months (*Ret*(*t*–4, *t*–2)), three-month return from the forth to sixth trailing month (*Ret*(*t*–7, *t*–4)), realized volatility over the preceding month (*RealizedVol*(*t*–2, *t*–1)), realized volatility over the second and third trailing months (*RealizedVol*(*t*–4, *t*–2)), realized volatility over the forth to sixth trailing month (*RealizedVol*(*t*–7, *t*–4)), natural logarithm of market capitalization at end of previous month (ln(*MarketCap*(*t*–1))), month fixed effects, and industry fixed effects based on the Fama-French 12-industry classification. These specifications do not include the Baker-Wurgler index because the dependent variable $G_{t,}$is normalized so that in each month its sum is one.

	Sample period: 2004–2019					Sample period: 2006–2019				
Explanatory variable	1	2	3	4	5	6	7	8	9	10
$SVI_{t-1,n}$		0.0024		0.0017	0.0021		0.0032		0.0025	0.0029
		(3.53)		(2.54)	(3.21)		(5.13)		(4.12)	(4.8)
$SYY\_Misp_{t-1,n}$			0.0502	0.0502	0.0456			0.0434	0.0433	0.0394
			(27.09)	(27.08)	(24.54)			(28.22)	(28.2)	(25.67)
$O/S_{t-1,n}$					0.0164					0.0128
					(25.24)					(24.4)
$IV_{t-1,n}$	0.0208	0.0207	0.0188	0.0188	0.0171	0.0133	0.0132	0.0104	0.0103	0.0093
	(10.12)	(10.1)	(8.5)	(8.48)	(7.71)	(7.52)	(7.48)	(5.67)	(5.63)	(5.09)
Ret(*w*–2, *w*–1)	–0.0502	–0.0501	–0.0460	–0.0459	–0.0427	–0.0446	–0.0445	–0.0403	–0.0402	–0.0375
	(–13.12)	(–13.1)	(–11.1)	(–11.08)	(–10.31)	(–13.48)	(–13.45)	(–11.63)	(–11.6)	(–10.86)
Ret(*w*–3, *w*–2)	0.0064	0.0064	0.0127	0.0127	0.0130	0.0049	0.0049	0.0089	0.0089	0.0093
	(3.17)	(3.18)	(3.82)	(3.83)	(3.93)	(2.86)	(2.88)	(3.23)	(3.25)	(3.39)

Table IA4
Tobit regression analysis of the determinants of SEP sentiment (continued)

	Sample period: 2004–2019					Sample period: 2006–2019				
Explanatory variables:	1	2	3	4	5	6	7	8	9	10
Ret(*t*–2, *t*–1)	–0.0126	–0.0125	–0.0148	–0.0148	–0.0166	–0.0118	–0.0118	–0.0130	–0.0130	–0.0143
	(–6.14)	(–6.12)	(–5.83)	(–5.82)	(–6.52)	(–6.78)	(–6.76)	(–6.2)	(–6.19)	(–6.82)
Ret(*t*–4, *t*–2)	–0.0015	–0.0015	–0.0001	–0.0001	–0.0017	–0.0019	–0.0018	–0.0006	–0.0006	–0.0018
	(–1.26)	(–1.21)	(–0.09)	(–0.06)	(–1.37)	(–1.81)	(–1.75)	(–0.57)	(–0.53)	(–1.72)
Ret(*t*–7, *t*–4)	0.0040	0.0040	0.0053	0.0054	0.0038	0.0043	0.0043	0.0052	0.0053	0.0041
	(4.24)	(4.29)	(5.38)	(5.4)	(3.86)	(5.27)	(5.34)	(6.43)	(6.47)	(5.01)
RealizedVol(*t*–2, *t*–1)	0.2097	0.2095	0.2150	0.2149	0.2031	0.1898	0.1896	0.1895	0.1894	0.1789
	(11.17)	(11.16)	(10.27)	(10.26)	(9.7)	(11.87)	(11.85)	(10.98)	(10.97)	(10.37)
RealizedVol(*t*–4, *t*–2)	0.2356	0.2354	0.2142	0.2140	0.2057	0.2000	0.1997	0.1736	0.1734	0.1657
	(11.78)	(11.77)	(9.89)	(9.88)	(9.5)	(11.75)	(11.74)	(9.78)	(9.77)	(9.35)
RealizedVol(*t*–7, *t*–4)	0.4050	0.4043	0.3213	0.3208	0.3097	0.3415	0.3405	0.2545	0.2537	0.2445
	(22.3)	(22.26)	(16.99)	(16.96)	(16.38)	(22.23)	(22.16)	(16.55)	(16.5)	(15.92)
ln(*MarketCap*(*t*–1))	0.0162	0.0162	0.0164	0.0164	0.0145	0.0135	0.0134	0.0133	0.0132	0.0117
	(77.17)	(76.78)	(70.2)	(69.86)	(61.33)	(74.15)	(73.68)	(67.27)	(66.87)	(58.37)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cal. month fixed effs.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. of obs.	160,938	160,938	133,087	133,087	132,901	138,883	138,883	112,136	112,136	111,950

Table IA5
Returns prior to SEP issuances
This table reports the average abnormal returns on the SEPs' reference stocks during the six month prior to the SEPs issuance months and during the first part of the issuance month for the two samples 2004–2019 and 2006–2019. For months −6 through −1, we form calendar-time portfolio returns of the reference stocks of the SEPs issued during the issue month 0. The portfolio returns are either weighted by the sentiment measures $G_{t,n}$ from the issue month or equally weighted. The first two sets of results in each panel report the average SEP-weighted and equal-weighted returns. For the second two sets of results in each panel, we regress the calendar-time portfolio returns on the Fama-French (2015) five factors and report the regression intercepts. The returns during the first part of the issuance month are the returns from the first trading day of the month through the trading day before the pricing date, then scaled up to a monthly interval similar to the scaling used in Table 7 of the main text. The computation of abnormal returns and weighting of returns during the first part of the issuance month also follows the approach used in Table 7.

	Month relative to pricing month						First part of
	−6	−5	−4	−3	−2	−1	issuance month
Panel A: 2004–2019							
SEP-weighted returns	0.0128	0.0144	0.0097	0.0103	0.0094	0.0077	0.0095
	(2.22)	(2.48)	(1.65)	(1.82)	(1.83)	(1.51)	(1.03)
Equal-weighted returns	0.0139	0.0144	0.0139	0.0134	0.0110	0.0084	0.0027
	(3.05)	(3.28)	(3.07)	(3.04)	(2.60)	(2.06)	(0.38)
SEP-weighted abnormal returns	0.0028	0.0032	−0.0023	−0.0006	−0.0005	−0.0018	0.0015
	(0.74)	(0.87)	(−0.68)	(−0.19)	(−0.16)	(−0.48)	(0.29)
Equal-weighted abnormal returns	0.0036	0.0042	0.0036	0.0036	0.0016	−0.0009	−0.0053
	(1.91)	(2.32)	(1.91)	(2.06)	(0.84)	(−0.52)	(−2.28)
Panel B: 2006–2019							
SEP-weighted returns	0.0123	0.0130	0.0084	0.0092	0.0082	0.0062	0.0085
	(1.91)	(2.08)	(1.33)	(1.52)	(1.48)	(1.13)	(0.84)
Equal-weighted returns	0.0136	0.0141	0.0128	0.0123	0.0112	0.0079	0.0016
	(2.72)	(2.91)	(2.57)	(2.55)	(2.45)	(1.79)	(0.21)
SEP-weighted abnormal returns	0.0031	0.0022	−0.0029	−0.0007	−0.0017	−0.0029	0.0011
	(0.75)	(0.59)	(−0.83)	(−0.21)	(−0.50)	(−0.76)	(0.21)
Equal-weighted abnormal returns	0.0039	0.0047	0.0031	0.0032	0.0019	−0.0015	−0.0065
	(1.92)	(2.42)	(1.56)	(1.63)	(0.95)	(−0.78)	(−2.61)