

Speculation Sentiment

Shaun William Davies 
University of Colorado Boulder Leeds School of Business
shaun.davies@colorado.edu

Abstract

I exploit the leveraged exchange-traded funds' (ETFs') primary market to measure aggregate, uninformed, gambling-like demand, that is, speculation sentiment. The leveraged ETFs' primary market is a novel setting that provides observable arbitrage activity attributed to correcting mispricing between ETFs' shares and their underlying assets. The arbitrage activity proxies for the magnitude and direction of speculative demand shocks and I use them to form the *Speculation Sentiment Index*. The measure negatively relates to contemporaneous market returns (e.g., it is bullish in down markets) and negatively predicts returns. The results are consistent with speculation sentiment causing market-wide price distortions that later reverse.

I. Introduction

Traders actively betting on a point of view believe they have special information, even if in reality they do not. The sources of these traders' "information" are numerous. Some traders may have material, nonpublic information that makes them truly informed. Others, however, may have a set of useless signals that give them a false sense of being informed (e.g., believing a spurious correlation between last week's weather and market returns is information). In this article, I focus on the demand from uninformed traders and I define speculation sentiment as a gambling-like, uninformed belief about the future direction of the market. Similar to the beliefs of the gambler who looks at the roulette wheel, conditions on its recent history of spins, and makes an "informed" wager on which color will result next,

I thank Turan Bali, Hendrik Bessembinder (the editor), Daniel Brown, David C. Brown, Roberto Caccia, Alex Chincio, Tony Cookson, Mike Cooper, Martijn Cremers, Zhi Da (the referee), Rob Dam, Andrew Detzel, Hui Dong, Rich Evans, Diego Garcia, David Ikenberry, Pab Jotikasthira, Andrew Karolyi, G. D. Knolls, Joni Kokkonen, Ryan Lewis, Nathalie Moyen, Peter Nolan, Matthew Ringgenberg, Charles Robbins, Shri Santosh, Clemens Sialm, Michael Stutzer, Brian T. Waters, Quan Wen, Ed Van Wesep, Jaime Zender, and conference and seminar participants at the 2019 China International Conference in Finance (CICF), 2019 Financial Intermediation Research Society (FIRS), UCLA Anderson Finance Conference in honor of Bhagwan Chowdhry, 2019 Front Range Finance Seminar, 2019 American Finance Association Annual Meeting (AFA), 8th Helsinki Finance Summit on Investor Behavior, 1st World Symposium on Investment Research, Arizona State University, Baylor University, the University of California, San Diego, Carnegie Mellon University, Chapman University, Cornell University, the University of Colorado, Boulder, the University of Utah, and the University of Virginia Darden School for their helpful insights and suggestions. All errors are my own. The Speculation Sentiment Index data are located on my website: <https://www.shaunwdavies.com/>.

speculation sentiment is the mood of an uninformed trader who develops a belief about the market's near-term performance and actively bets on it.¹

Does speculation sentiment from individual traders aggregate in a meaningful way? If it does, do changes in this speculative demand move asset prices away from fundamentals? To date, the answers to these questions have been elusive because speculation sentiment is difficult to identify. However, in this article, I provide a novel and direct means of measuring this nonfundamental (i.e., uninformed) demand and I provide credible evidence that speculation sentiment i) affects traders en masse and ii) leads to price distortions that subsequently reverse. The measure, which I coin the *Speculation Sentiment Index* (SSI), is based on observable arbitrage trades in correcting relative mispricing in the leveraged exchange-traded funds (ETFs) market. The index is negatively correlated with contemporaneous stock returns and it negatively predicts subsequent returns. Moreover, the return predictability is economically meaningful: a 1-standard-deviation increase in the monthly index is associated with a 1.14%–1.67% decline in broad market stock indices the following month.² The results are robust to the inclusion of other sentiment proxies and market controls and also to out-of-sample tests.

To understand why leveraged ETFs provide a novel laboratory for studying speculation sentiment, a short primer is helpful. A leveraged ETF's shares provide magnified, short-horizon exposure to a market benchmark (e.g., the Standard & Poor's 500 (S&P 500) index). The shares trade intraday in the secondary market and are characterized by high trade volume (relative to nonleveraged ETFs and single-name stocks). Leveraged ETF shares are primarily traded among individuals and short-horizon traders and, unlike margin accounts or option trading which require special approvals, any trader may purchase a leveraged ETF share in his or her brokerage account (and also in many retirement accounts).³ As a pooled investment vehicle, the intrinsic value of a leveraged ETF share is determined by the value of an underlying basket of derivative securities and cash holdings. The underlying derivative securities are traded primarily by institutions and for several purposes, such as risk management and hedging. Consequently, there are different investor clienteles trading the shares and trading the underlying derivative securities: "dumb" money trades the shares and "smart" money trades the underlying derivative securities. I argue that these two distinct clienteles cause there to be a difference in the demand for the leveraged ETF shares and the demand for the underlying derivative securities. In particular, my identifying assumption is that leveraged ETF share demand is *relatively* more sensitive to gambling-like, uninformed demand shocks than the underlying derivative security demand.

Under the identifying assumption that leveraged ETF share demand is relatively more sensitive to speculative demand shocks than the underlying derivative security demand, the realization of a shock gives rise to a relative mispricing (i.e., a violation of the law of one price). Importantly, remnants of mispricing are

¹Note an uninformed trader's beliefs may be correlated with fundamental news (e.g., over- and under-reaction to macroeconomic news).

²The Speculation Sentiment Index may be downloaded here: <https://www.shawnwdavies.com/research>.

³Frazzini and Pedersen (2014) argues that investors with leverage constraints (e.g., individual investors) are attracted to leveraged ETFs. See also Frazzini and Pedersen (2022).

observable in the leveraged ETF market unlike other settings in which mispricing may be quickly exploited by arbitrageurs leaving no evidence for the empiricist. Observable remnants are due to a unique feature of the ETF market: arbitrageurs exploit relative mispricing in a primary market by creating and redeeming ETF shares. This process allows the empiricist to observe arbitrage activity via changes in shares outstanding.⁴ Thus, leveraged ETFs provide a special setting to directly observe a proxy of speculative demand shocks.

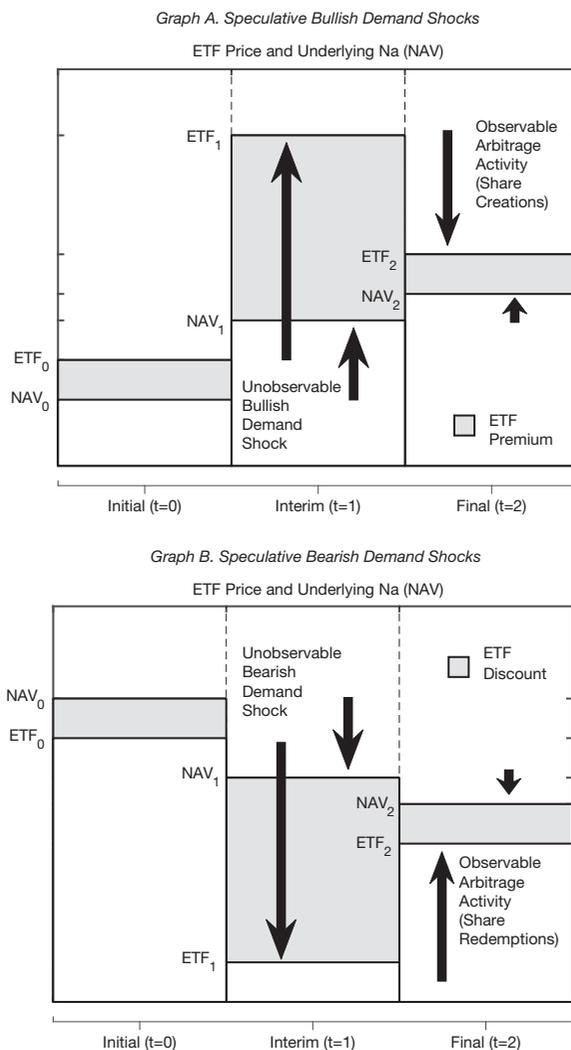
To see how relative mispricing between ETF shares and their underlying derivative securities leads to share creations or share redemptions, consider the two examples for a leveraged-long ETF in Figure 1. Graph A of Figure 1 depicts a setting in which ETF share prices increase relative to their underlying derivative securities' prices, leading to share creations. In that figure, at $t = 0$, a small relative mispricing exists between the leveraged-long ETF shares and their underlying net asset values (NAVs), but it is not large enough to attract arbitrageurs due to transaction costs. At $t = 1$, a latent *bullish* demand shock is realized, and the demands for the ETF shares and the underlying derivative securities are affected to different degrees, generating a larger relative mispricing in the form of an ETF premium (i.e., the share price exceeds its NAV). In response to the ETF premium, arbitrageurs short-sell the ETF shares and hedge that position with a long position in either the underlying asset or a derivative contract. The trades of arbitrageurs put downward price pressure on the ETF shares and upward price pressure on the NAV until the trade is no longer profitable, which can be seen at $t = 2$. At the end of the trading day, arbitrageurs close their trades by simultaneously unwinding their long hedge, purchasing newly minted ETF shares (priced at NAV), and covering their short position with the new ETF shares. As such, ETF premiums lead to share creations.

Graph B of Figure 1 depicts a setting in which ETF share prices decrease relative to their underlying derivative securities' prices, leading to share redemptions. Similar to the first example, at $t = 0$, a small relative mispricing exists between the leveraged-long ETF shares and the underlying NAVs, but it is not large enough to attract arbitrageurs due to transaction costs. At $t = 1$, a latent *bearish* demand shock is realized, and the demands for the ETF shares and the underlying derivative securities are affected to different degrees, generating a relative mispricing in the form of an ETF discount (i.e., the NAV exceeds its share price). In response to the ETF discount, arbitrageurs purchase the ETF shares and hedge the long position with a short position in either the underlying asset or a derivative contract. The trades of arbitrageurs put upward price pressure on the ETF shares and downward price pressure on the NAV until the trade is no longer profitable, which can be seen at $t = 2$. At the end of the trading day, arbitrageurs close their trades by unwinding their short hedge and delivering the purchased shares (which are priced at NAV and are subsequently destroyed) in exchange for cash. As such, ETF discounts lead to share redemptions. In the Supplementary

⁴In theory, if one had high frequency pricing data, he or she could also measure the realization of demand shocks (and arbitrage activity) via expansions and contractions in mispricing. Using net share change, which proxies for the aggregation of all mispricing corrected via arbitrage, has the advantage that it does not require high frequency data.

FIGURE 1
Speculative Demand Shocks in the Leveraged-Long ETF Market

Figure 1 displays the effect of speculative demand shocks on the leveraged-long ETF shares and the leveraged-long ETF underlying derivative securities. Graph A portrays a setting in which a speculative bullish demand shock leads to an ETF premium that is exploited via share creations. Graph B portrays a setting in which a speculative bearish demand shock leads to an ETF discount that is exploited via share redemptions.



Material, I provide a theoretical model which adds rigor and concreteness to the preceding discussion.⁵ The model echoes the insights from Figure 1 and also provides guidance for the empirical analysis.

⁵The creation and redemption process for leveraged-short ETFs is similar but with two notable differences: i) with leveraged-short ETFs, bullish demand shocks lead to share redemptions while bearish demand shocks lead to share creations and ii) the arbitrageurs' intraday hedging is in the opposite

I form the SSI using the first leveraged ETFs offered to traders, which were introduced by ProShares in the summer of 2006. Using the original leveraged ETFs, three that provide $2\times$ long exposure and three that provide $2\times$ short exposure to market indices, I calculate SSI at the monthly frequency. The index is calculated by taking the difference between share change in the $2\times$ leveraged-long ETFs and share change in the $2\times$ leveraged-short ETFs. SSI provides a glimpse into the mood of speculators; if the number is large and positive, speculators heavily demanded leveraged-long exposure, so much so that the leveraged-long ETF share prices drifted above NAVs leading to arbitrage opportunities (leveraged-short ETF share prices drifted below NAVs). If the number is large and negative, speculators heavily demanded leveraged-short exposure. Finally, if the number is near zero, the demand for leveraged-long and leveraged-short ETFs effectively canceled out or the speculative demand shock was small. Importantly, the measure does not require that leveraged ETF trading is the source of mispricing in the broad market (i.e., *this is not a price pressure story from leveraged ETF trading*). After all, the broad market mispricing identified later in this article is substantial relative to the size of the leveraged ETF market. Instead, I argue that leveraged ETFs are a unique setting to identify and measure market-wide speculative demand shocks.⁶

The SSI is depicted in Graph A of Figure 2. As can be seen, the index is considerably variable and exhibits little persistence. To gain some understanding of how SSI relates to aggregate stock returns, Graphs B and C provide anecdotal evidence using the S&P 500 index. Graph B depicts a scatter plot of SSI monthly values (the horizontal axis) versus contemporaneous S&P 500 index monthly returns (the vertical axis), along with a trend line. The scatter plot shows a strong *negative* relation between SSI and S&P 500 returns as the majority of the data points (66%) are in the graph's second and fourth quadrants. In other words, when the S&P 500 is performing well, SSI is generally bearish and when the S&P 500 is performing poorly, SSI is generally bullish. Graph C depicts a scatter plot of SSI monthly values (the horizontal axis) versus the following months' S&P 500 index returns (the vertical axis), along with a trend line. The scatter plot also shows a *negative* relation between SSI and future S&P 500 returns as the majority of the data points (57%) are in the graph's second and fourth quadrants. Said differently, bullish SSI predicts negative future returns and bearish SSI predicts positive future returns.⁷

Motivated by the scatter plots in Figure 2, I perform regression analysis to study the relation between SSI and the returns of three benchmark indices: i) the CRSP equal-weighted index, ii) the CRSP value-weighted index, and iii) the S&P 500 index. First, I regress contemporaneous index monthly returns on monthly SSI. A 1-standard-deviation increase in SSI is associated with a statistically significant 3.28% decline in the CRSP equal-weighted index, a 2.71% decline

direction of the leveraged-long ETF examples (e.g., to hedge being short leveraged-short ETF shares, arbitrageurs will short the underlying asset or a derivative contract).

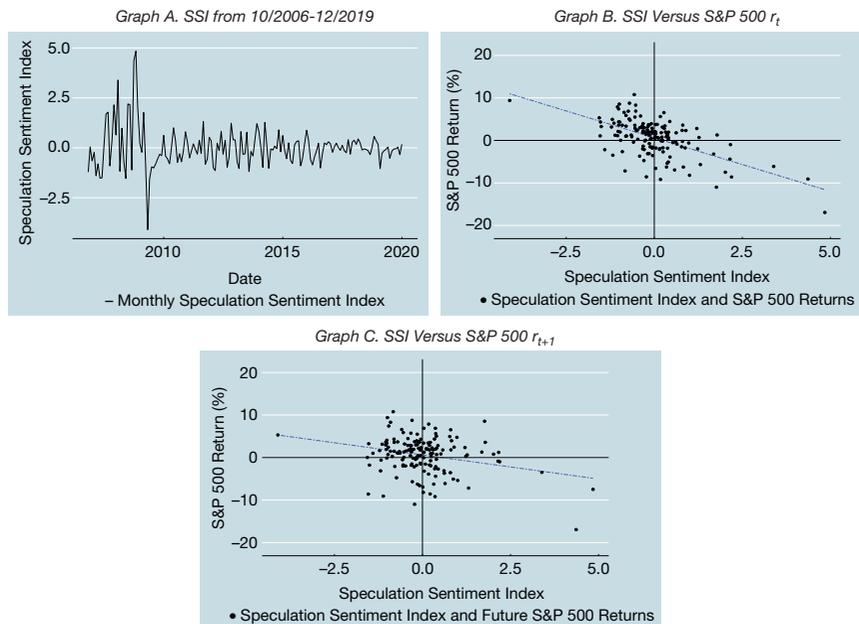
⁶The model in Section IA.2 of the Supplementary Material shows theoretically that SSI proxies for market-wide speculative demand shocks even without leveraged ETF primary market trades having price impact on the broad market. See Remark IA4 of the Supplementary Material.

⁷Scatter plots using the CRSP equal weighted index and the CRSP value weighted index depict similar relations between SSI and contemporaneous index returns/future index returns.

FIGURE 2

The Speculation Sentiment Index and Market Returns

Graph A of Figure 2 depicts the standardized monthly time series of SSI from Oct. 2006 through Dec. 2019. Graph B is a scatter plot of SSI versus the contemporaneous S&P 500 index monthly return. Graph C is a scatter plot of SSI versus the next's months S&P 500 return. The dotted lines in the scatter plots are the trend lines.



in the CRSP value-weighted index, and 2.52% decline in the S&P 500. The results confirm the evidence from Graph B of Figure 2: the negative relation between SSI and contemporaneous returns is statistically significant. Second, I perform predictive regressions with future monthly index returns as the dependent variable and monthly SSI as the independent variable. A 1-standard-deviation increase in SSI predicts a statistically significant 1.67% decline in the CRSP equal-weighted index, a 1.27% decline in the CRSP value-weighted index, and a 1.14% decline in the S&P 500 index. The predictive power of SSI is not driven by the 2008–2009 Financial Crisis; repeating the analysis beginning in Jan. 2010, the coefficients remain relatively stable in magnitude with statistically significant p -values.

Given the strong return predictability results, I next provide evidence that the results are consistent with a sentiment interpretation as opposed to a rational explanation. First, it is possible that SSI relates to rational portfolio rebalancing and that those rebalancing trades have price impact that subsequently reverses. Such a possibility is consistent with the empirical evidence; SSI is negatively related to contemporaneous returns (i.e., the rebalancing) and SSI negatively predicts subsequent returns (i.e., the reversal of rebalancing trades' price impact). If this is indeed the case, then SSI simply proxies for realized returns and the return predictability is driven by autocorrelation in returns. Thus, to rule this possibility out, I examine

whether or not realized returns predict subsequent returns in my sample. I find no relation between realized returns and subsequent returns. Moreover, in bivariate predictive regressions using both realized returns and SSI as independent variables, the predictive power of SSI is stronger than in the univariate regressions (both in economic magnitude and statistical significance). These results suggest that the return predictability results are not driven by rational rebalancing.

Second, to provide additional evidence that the results are consistent with a sentiment interpretation, I examine the return predictability horizon. That is, one may be concerned about the sentiment interpretation if the monthly return predictability is concentrated in the first few days of the proceeding month (which would be more akin to price pressure reversals from rational trading). As such, I examine monthly SSI's ability to predict cumulative returns at shorter horizons (i.e., the first few days of the proceeding month) and longer horizons (i.e., the proceeding 6 months). Initially, there is little to no return predictability from SSI. However, a few days into the first month, return predictability becomes both economically meaningful and statistically significant. Throughout the first month, and into the proceeding months, cumulative return predictability grows and remains statistically significant; the CRSP equal-weighted index exhibits statistically significant return predictability out to 4 months, the CRSP value-weighted index exhibits statistically significant return predictability out to 4 months, and the S&P 500 index exhibits statistically significant return predictability out to 6 months. In other words, the evidence is more consistent with sentiment shocks that lead to dislocations that slowly reverse over multiple months as opposed to price impact from rational rebalancing that would likely reverse more quickly.

Finally, I examine the robustness of the return predictability results in two ways. First, given that there are many established sentiment proxies and aggregate return predictors, a reasonable concern is that SSI simply reflects one of these known measures. As such, I next perform bivariate predictive regressions with future monthly index returns as the dependent variable and monthly SSI along with a sentiment proxy or market control as the independent variables. The predictive power of SSI is robust with the inclusion of controls: a 1-standard-deviation increase in SSI predicts a 1.14%–1.86% decline in the CRSP equal-weighted index, a 0.85%–1.33% decline in the CRSP value-weighted index, and a 0.75%–1.18% decline in the S&P 500 index. Therefore, SSI is distinct and it is unlikely that SSI is measuring a known sentiment proxy or market control. Second, I examine the ability of SSI to predict returns out-of-sample in the spirit of Campbell and Thompson (2008) and Welch and Goyal (2008). I calculate out-of-sample R^2 s using several different starting dates and the results show that SSI has statistically significant out-of-sample predictive power. The results also imply that investors can benefit from a market-timing strategy that conditions on realized values of SSI.

The main contribution of this article is in providing a clean measure of speculation sentiment based on the arbitrage activity it generates and showing that the measure has substantial return predictability.⁸ That said, speculation sentiment

⁸Other empirical research also shows that demand for assets, unrelated to fundamentals, creates price dislocations that do not immediately revert. The sources of these nonfundamental demand shocks are

is just one dimension of broader *investor sentiment*.⁹ I argue that speculation sentiment is a gambling-like dimension of investor sentiment (as leveraged ETFs are uniquely tailored for short-horizon bets). In that regard, my measure is related to the closed-end fund discount: Zweig (1973), Lee, Shleifer, and Thaler (1991), and Neal and Wheatley (1998) argue that closed-end funds are disproportionately held by individual traders, much like leveraged ETF shares, and that the aggregate discount reflects individual traders' bearish or bullish beliefs. Nevertheless, I show that my results are robust to the inclusion of the aggregate closed-end fund discount.¹⁰ The robustness of the results implies that the nonfundamental demand identified in leveraged ETF share change is distinct from the nonfundamental demand identified in the closed-end fund discount.

While I focus on a single dimension of investor sentiment in this article, there are many other established proxies relating to a host of investor sentiment dimensions, for example, the mood of traders suffering from external disappointment (Edmans, Garcia, and Norli (2007)) or sadness (Saunders (1993)). With many dimensions to investor sentiment, there is not a shortage of sentiment proxies in the literature.¹¹ Importantly, sentiment measures need not compete with each other as "the sentiment" measure; given numerous dimensions to investor sentiment, observing many measures does not imply that most measures are wrong. Nevertheless, there are reasons why SSI is unique and important. First, the index is a strong sentiment proxy as it has significant predictive power even after controlling for other popular sentiment proxies. The index is also robust to alternative specifications and is robust to out-of-sample tests. Second, the index is constructed from the trades of arbitrageurs exploiting relative mispricing between leveraged ETFs' shares and the ETFs' underlying assets (i.e., the measure is based on the reliable law of one price). Thus, there is a natural economic interpretation to the measure; SSI proxies for realized disagreement between "dumb" and "smart" traders that

numerous: index rebalancing (Shleifer (1986)), liquidity needs (Coval and Stafford (2007)), mutual fund flows (Ben-Rephael, Kandel, and Wohl (2012), Lou (2012)), stale information (Huberman and Regev (2001), DellaVigna and Pollet (2007), and Hong, Torous, and Valkanov (2007)), and investor inattention (DellaVigna and Pollet (2009), Hirshleifer, Lim, and Teoh (2009)).

⁹Baker and Wurgler (2007) defines investor sentiment as "a belief about future cash flows and investment risk that is not justified by the facts at hand." As such, investor sentiment is inherently multidimensional as sentiment is related to the behaviors of individual traders and there are many well documented behavioral biases. See Hirshleifer (2001) and Barberis and Thaler (2003) for surveys of the behavioral finance literature.

¹⁰The results are also robust to the inclusion of the Baker–Wurgler Investor Sentiment Index (Baker and Wurgler (2006)) which aggregates several market sentiment measures, including the closed-end fund discount. While Baker and Wurgler (2006) shows that the Baker–Wurgler Investor Sentiment Index generates predictability in the cross section of returns, I show predictability in aggregate market returns. Huang, Jiang, Tu, and Zhou (2015) provides a modified measure of the Baker–Wurgler Investor Sentiment Index that utilizes a partial least squares (PLS) method to minimize noise in the index's input variables and shows that the modified measure predicts aggregate returns. My results are robust to the inclusion of this measure as a control variable.

¹¹See Baker and Wurgler (2007) for a survey of investor sentiment measures and DeVault, Sias, and Starks (2019) for an analysis of existing sentiment measures relation to institutional demand versus individual demand. See also Jiang, Lee, Martin, and Zhou (2018) for a measure of corporate manager sentiment based on the tone of financial disclosures and Da, Engelberg, and Gao (2014) for a measure of fear sentiment based on daily Internet search volume.

leads to mispricing. Third, the measure's input data is widely available, the measure is straightforward to construct, and the measure may easily be constructed at different frequencies. Fourth, the leveraged ETF market is vibrant and it is likely that the index will serve as a powerful sentiment measure in the foreseeable future. These reasons suggest that SSI will serve as an important sentiment proxy in future asset pricing and corporate finance studies.

This article also adds to a growing literature that uses ETFs as a laboratory to study nonfundamental demand. Ben-David, Franzoni, and Moussawi (2018) document the transmission of nonfundamental demand volatility for ETF shares to the ETF's underlying assets via the primary market mechanism. In a similar spirit, Brown, Davies, and Ringgenberg (2021) show theoretically and empirically that ETF share changes (i.e., ETF flows) provide informative signals of nonfundamental demand shocks and that conditioning on these signals yields cross-sectional return predictability.¹² Brown et al. (2021) also find that the predictability is, generally, stronger in the universe of leveraged ETFs (including equity, bond, and commodity funds). As such, my study complements that of Brown et al. (2021) in several ways. First, while Brown et al. (2021) are agnostic regarding the types of nonfundamental demand shocks identified in leveraged ETF flows, I focus exclusively on equity ETFs to isolate speculative demand shocks and to form SSI. Furthermore, as SSI is a sentiment measure, I show that it is distinct from other known proxies of nonfundamental demand (e.g., the Baker–Wurgler Investor Sentiment Index). As such, my analysis provides new insights to better appreciate the results in Brown et al. (2021). Second, Brown et al. (2021) forms long-short portfolios based on ETF flows and finds return predictability in the cross section of ETFs. Conversely, I show that SSI provides aggregate return predictability in the time series. In other words, my results imply that there is a market-wide component of nonfundamental demand identified in leveraged ETF flows.¹³

Finally, leveraged ETFs have also been of interest to academics. Cheng and Madhavan (2009) show that the daily rebalancing dynamics of leveraged ETFs (i.e., maintaining the target leverage exposure) supports the claim that leveraged ETFs lead to greater end-of-day market volatility. Empirically, however, there is a debate to how much excess volatility leveraged ETFs generate: Tuzun (2013) and Shum, Hejazi, Haryanto, and Rodier (2015) provide new evidence that leveraged ETF rebalancing exacerbates market volatility while Ivanov and Lenkey (2014) suggest excess volatility concerns are overblown. Furthermore, Bessembinder (2015) argues that end-of-day rebalancing leads to predictable order flow, which should have minimal effects on long-term prices. While I study a set of leveraged ETFs to formulate SSI, my focus is on the arbitrage activity associated with investor demand and not the daily rebalancing activities within leveraged ETFs. In addition to studies on the rebalancing dynamics of leveraged ETFs, Egan, MacKay, and Yang (2020) utilize leveraged ETFs, along with nonleveraged funds, to study

¹²See also Staer (2017) and Jiang, Xiao, and Yan (2020) which show that ETF arbitrage activity is associated with contemporaneous price pressure and subsequent return reversals. Furthermore, see Ben-David, Franzoni, and Moussawi (2017) for a survey of the ETF literature.

¹³There is little evidence that cross sectional return predictors make good time series return predictors, especially out of sample. See Engelberg, Mclean, Pontiff, and Ringgenberg (2022).

investors' expectations of market returns. The study estimates a model of investor expectations using investors' revealed preferences in choosing from a menu of nonleveraged and leveraged S&P 500 index products. While the study documents a weak negative correlation between investor expectations and returns the following year, the focal points of the study are documenting the heterogeneity of investor expectations, the path dependency of investor expectations, and the persistence of investor expectations.

II. Background

On June 21, 2006, ProShares announced a set of four ETFs designed to make it easier for investors to get magnified exposure to an index. The four ETFs' daily objective is to provide $2\times$ exposure to well-known indices like the S&P 500 and the Dow Jones Industrial Average (before fees and expenses). Three weeks later, on July 13, 2006, ProShares announced a set of four additional ETFs designed to provide magnified *short* exposure to well-known market indices. The set of leveraged ETFs announced during the summer of 2006 is provided in Table 1. These eight ETFs sponsored by ProShares represent the first set of leveraged ETFs offered in U.S. markets. Since the original eight launched in the summer of 2006, nearly 300 additional leveraged ETFs have been offered to investors. There are now leveraged ETFs providing magnified exposures to bond indices, commodities, currencies, emerging markets, and market volatility indices.

A. The ETF Market and Mechanism

The ETF market is large; with over three trillion dollars under management, ETFs collectively hold more assets than hedge funds (Madhavan (2016)). ETFs are also ingrained into nearly every asset market, both domestic and foreign.¹⁴ Furthermore, ETFs are accessible to novice and professional investors alike.¹⁵

TABLE 1
Leveraged ETFs Launched by ProShares During 2006

Table 1 reports the leveraged ETFs launched by ProShares during the summer of 2006. Panel A provides $2\times$ long exposure to prespecified indices, and Panel B provides $2\times$ short exposure to the same indices.

Fund Name	Daily Objective	Ticker
<i>Panel A. ETFs Announced on June 21, 2006</i>		
Ultra QQQ ProShares	Double the NASDAQ-100 Index	QLD
Ultra S&P 500 ProShares	Double the S&P 500 Index	SSO
Ultra Dow30 ProShares	Double the Dow Jones Industrial Average	DDM
Ultra MidCap400 ProShares	Double the S&P MidCap 400	MVZ
<i>Panel B. ETFs Announced on July 13, 2006</i>		
UltraShort QQQ ProShares	Double the inverse of the NASDAQ-100 Index	QID
UltraShort S&P 500 ProShares	Double the inverse of the S&P 500 Index	SDS
UltraShort Dow30 ProShares	Double the inverse of the Dow Jones Industrial Average	DXD
UltraShort MidCap400 ProShares	Double the inverse of the S&P MidCap 400	MZZ

ETFs are a pooled investment vehicle, like a mutual fund, which allows investors to buy a basket of assets at once.¹⁶ Like a closed-end mutual fund, investors can buy or sell an ETF share on a secondary market just as they would buy or sell a stock. However, unlike a closed-end mutual fund, shares in an ETF are added or removed in a primary market via the actions of third-party arbitrageurs called authorized participants (APs). APs who are prequalified by the fund sponsor (e.g., ProShares) are allowed to exchange shares of the ETF for shares of the underlying assets (an in-kind transaction) or for cash. Similarly, APs may deliver the underlying assets or cash in exchange for the ETF shares. This process, which is designed to equilibrate supply and demand for shares in the ETF, allows APs to enforce the law of one price. For example, if an ETF price gets too high relative to the value of the underlying assets, an AP short-sells the ETF shares and purchases the underlying assets. At the end of the day, the AP delivers the underlying assets (for in-kind transactions) or delivers cash in exchange for new ETF shares. The AP then covers the short position in the ETF with the new shares. The AP conducts the opposite trade if an ETF price gets too low relative to the value of the underlying assets, removing ETF shares from the market.

B. Leveraged ETFs

Leveraged ETFs are similar in most ways to traditional, nonleveraged ETFs, but they also have unique features. First, unlike most nonleveraged ETFs, leveraged ETFs replicate their intended benchmark via derivatives.¹⁷ For example, to obtain $2\times$ or $-2\times$ exposure to an index, the ETF sponsor enters into total return swaps, which are rolled on a regular basis. Second, while most nonleveraged ETFs adhere to a static policy of in-kind transactions, the creation and redemption process conducted between the leveraged ETF sponsor and APs always includes an element of cash in the exchange of shares.

Leveraged ETFs are designed for short-horizon trades as they effectively provide constant *daily* leveraged returns. To do so, a leveraged ETF must rebalance at the end of the day in the direction of that day's return. For example, to maintain constant $2\times$ exposure, a leveraged-long ETF would have to buy if the underlying benchmark index increased in value or sell if the index decreased in value. Importantly, the rebalancing activities of a leveraged ETF are within the fund and are distinct from the primary market activities of share creations and redemptions.¹⁸

¹⁵ETFs are a popular investment choice within individual retirement plans (e.g., 401Ks) and also a popular investment for professional managers to "equitize" cash in their funds' benchmarks (Antoniewicz and Heinrichs (2014)).

¹⁵ETFs are a popular investment choice within individual retirement plans (e.g., 401Ks) and also a popular investment for professional managers to "equitize" cash in their funds' benchmarks (Antoniewicz and Heinrichs (2014)).

¹⁶Like mutual funds, most ETFs are formally registered with the SEC as investment companies under the Investment Company Act of 1940.

¹⁷All ETFs replicate their intended benchmark via one of three methods: full replication, optimized replication, and derivative replication. Many ETFs are fully replicated, meaning that the ETF physically holds the underlying assets in the intended benchmark. Optimized replication is similar, but does not require the ETF to hold every asset. Instead, the ETF sponsor may hold a representative sample that minimizes tracking error while avoiding difficult to obtain or illiquid securities.

While leveraged ETFs are effective in their goal of providing constant multiples of daily returns, longer-term compounded returns do not share these constant multiples. Borrowing an example from Cheng and Madhavan (2009), consider a leveraged ETF that intends to provide $2\times$ exposure to a particular index. The ETF begins with an initial NAV of \$100. The benchmark index that starts at 100, falls by 10% 1 day and then goes up by 10% the next. Over the 2-day period, the index declines by 1% (down to 90 and then up to 99). One might expect that the leveraged ETF would provide a return of -2% . Instead, it declines by 4%; doubling the index's 10% fall pushes the ETF's NAV to \$80 on the first day. The next day, the fund's NAV climbs to \$96. The preceding example highlights that leveraged ETFs' provision of daily constant leverage multiples does not translate to longer-term constant leverage multiples. As such, an investor with a different objective from achieving a constant multiple of daily returns will require regular portfolio rebalancing. For example, if an investor is utilizing leveraged ETFs to achieve a target leverage quantity (i.e., a constant leverage multiple over several days), she is required to rebalance her holdings on a daily frequency in a trading pattern that resembles contrarian trading. Specifically, if her position increases in value, she must sell and if her position decreases in value, she must buy. In Section IA.6 of the Supplementary Material, I analytically solve for the quantity of rebalancing an investor must engage in, given that day's index return and the leverage quantity provided by the ETF.

Consistent with being designed for short-horizon trades, leveraged ETFs exhibit greater trade volume than their nonleveraged counterparts based on average turnover, which is measured as monthly volume divided by end-of-month shares outstanding. Table 2 compares the ProShares leveraged ETFs to their largest, nonleveraged, comparable ETFs. Share turnover in the leveraged ETFs SSO and SDS, which provide $2\times$ and $-2\times$ exposure to the S&P 500 index, were 1.38 and 1.54 times more than that in the nonleveraged ETF SPY (which is the largest nonleveraged ETF providing exposure to the S&P 500). To put this in perspective, if all shares in SPY were to transact once during a period of time, all shares in SSO would have transacted 1.38 times and all shares in SDS would have transacted 1.54 times during that same period. For the other ProShares leveraged ETFs, the numbers are slightly larger.

Leveraged ETFs are also traded among retail investors relatively more than nonleveraged ETFs or single-name stocks. For example, institutional ownership in leveraged ETFs relative to nonleveraged counterparts is low. Table 2 also provides the ratio of percent of shares held by institutional investors in the ProShares leveraged ETFs as compared to their largest, nonleveraged, comparable ETFs.¹⁹ For example, SSO and SDS exhibit only 34% and 15% of the percent of shares held by institutional investors in SPY. For the other original six leveraged ETFs, the ratios are comparable.

¹⁸For additional discussion, see Bessembinder (2015).

¹⁹Data for percent of shares held by institutions comes from Bloomberg. Institutional ownership is defined as *Percentage of Shares Outstanding held by institutions. Institutions include 13Fs, US and International Mutual Funds, Schedule Ds (US Insurance Companies) and Institutional stake holdings that appear on the aggregate level. Based on holdings data collected by Bloomberg.*

TABLE 2
Leveraged ETFs Compared to Nonleveraged ETFs

Table 2 compares leveraged ETFs to nonleveraged ETFs along the dimensions of monthly share turnover, institutional ownership, and end-of-month assets under management (AUM). For leveraged ETFs SSO and SDS, the nonleveraged comparable is SPY. For leveraged ETFs QLD and QID, the nonleveraged comparable is QQQ. For leveraged ETFs DDM and DXD, the nonleveraged comparable is DIA. For leveraged ETFs MVV and MZZ, the nonleveraged comparable is IJH. TURNOVER_RATIO is calculated as the ratio of a leveraged ETF's average monthly turnover (monthly volume divided by end-of-month shares outstanding) to its nonleveraged comparable's average monthly turnover. INST_OWN_RATIO is calculated as the ratio of a leveraged ETF's average institutional ownership to its nonleveraged comparable's average institutional ownership. AUM_Ratio is calculated as the ratio of a leveraged ETF's average AUM to its nonleveraged comparable's average AUM. TURNOVER_RATIO data covers Jan. 31, 2007–Dec. 31, 2019. INST_OWN_RATIO data covers Mar. 28, 2010–Dec. 31, 2019. AUM_Ratio data covers Jan. 31, 2007–Dec. 31, 2019.

	SPY		QQQ		DIA		IJH	
	SSO (%)	SDS (%)	QLD (%)	QID (%)	DDM (%)	DXD (%)	MVV (%)	MZZ (%)
TURNOVER_RATIO	138	154	210	305	161	234	1,104	1,615
INST_OWN_RATIO	34	15	31	27	22	16	56	33
AUM_RATIO	1	1	3	2	3	3	1	0

Finally, leveraged ETFs' are small in size as compared to their nonleveraged counterparts. Returning to Table 2, SSO represents just 1% of AUM as compared to SPY and SDS also represents only 1% of SPY. Across the other ProShares leveraged ETFs, the ratios are similar.

III. Data and Index Construction

A. Data

To construct and study SSI, I combine data from Bloomberg, ProShares, CRSP, Jeffrey Wurgler's website, Guofu Zhou's website, Hao Zhou's website, Robert Stambaugh's website, Asaf Manela's website, Matthew Ringgenberg, Robert Shiller's website, Kenneth French's website, Turan Bali's website, the University of Michigan Survey of Consumer's website, and the U.S. Treasury's website. From Bloomberg, I get daily data on ETF shares outstanding, share changes, prices, and trade volumes.²⁰ From Bloomberg I also get weekly data on ETF institutional ownership and ETF characteristics, such as stated benchmarks, leverage quantities, and leverage directions (i.e., long or short). From ProShares, I get ETF shares outstanding data which are used to crosscheck the Bloomberg data and for tests using daily share change. From CRSP, I get return data on the CRSP equal-weighted, CRSP value-weighted, and S&P 500 indices.

To control for broader investor sentiment, I use the Baker–Wurgler Investor Sentiment Index (Baker and Wurgler (2006)) and the closed-end fund discount, which are both obtained from Jeffrey Wurgler's website and I use the Survey of Consumer Confidence, which is taken from the University of Michigan Survey of Consumer's website. I also use the aligned investor sentiment level (Huang et al. (2015)), which exploits information in the Baker–Wurgler Investor Sentiment Index using a partial least squares (PLS) method. The measure is designed to predict aggregate stock returns and the data is obtained from Guofu Zhou's website.

²⁰Ben-David et al. (2018) shows that Bloomberg provides the most accurate ETF data.

To control for market conditions, I use VIX index data, which is obtained from Bloomberg. I also control for the variance risk premium (Bollerslev, Tauchen, and Zhou (2009)) using data from Hao Zhou's website. I control for aggregate liquidity using the Pastor–Stambaugh liquidity series (Pastor and Stambaugh (2003)), which is obtained from Robert Stambaugh's website and intermediary liquidity using the He–Kelly–Manela intermediary liquidity series (He, Kelly, and Manela (2017)), which is obtained from Asaf Manela's website. I control for short interest as a proxy for ETF arbitrageur liquidity using the Short Interest Index (Rapach, Ringgenberg, and Zhou (2016)), which is obtained from Matthew Ringgenberg. Additionally, I control for other predictors of returns including aggregate dividends-to-price and cyclically adjusted earnings-to-price ratios, which are obtained from Robert Shiller's website. Term spread and short-rate data are obtained from the U.S. Treasury's website. I add information on the 3-factor pricing model (Fama and French (1993)) from Kenneth French's website.

As discussed earlier, leveraged ETFs cater to short-horizon traders that desire amplified exposure to market benchmarks. Moreover, as discussed in Section II.B, leveraged ETFs are primarily held by individual investors. It is well-established that there is investor demand for lottery-like assets; Kumar (2009) and Han and Kumar (2013) show that speculative individual traders demonstrate a propensity to gamble with lottery-like stocks (e.g., low-priced stocks with high idiosyncratic volatility and idiosyncratic skewness). Motivated by these findings, Bali, Brown, Murray, and Tang (2017) form a measure of investor lottery demand using stocks' largest (smallest) daily returns the previous month.²¹ The measure, MAX factor, is formed using a strategy that goes short the stocks with the five largest daily returns and goes long the stocks with the five smallest daily returns. The MAX factor earns subsequent excess returns that cannot be explained by traditional risk factors and the measure also explains the beta anomaly (Black, Jensen, and Scholes (1972), Frazzini and Pedersen (2014), and Baker, Hoeyer, and Wurgler (2016)). As such, to control for investor lottery demand I use the MAX factor, which is obtained from Turan Bali's website.

From each data source, I obtain data series from 2006 through the end of 2019, with the exception of the intermediary liquidity series, which is only available through Nov. 2018, and the Baker–Wurgler Investor Sentiment Index, the aligned investor sentiment index, and the closed-end fund discount, which are only available through Dec. 2018.

B. SSI Construction

I construct the index using six of the eight original leveraged ETFs offered by ProShares: three leveraged-long ETFs (QLD, SSO, and DDM) and three leveraged-short ETFs (QID, SDS, and DXD). Each long-short pair tracks an intended index: SSO and SDS provide 2× exposure to the S&P 500 index, QLD and QID provide 2× exposure to the NASDAQ-100 index, and DDM and DXD provide 2× exposure to the Dow Jones Industrial Average. The two excluded ETFs, MVV and MZZ, are a long-short pair that provide exposure to the S&P MidCap 400 Index. MVV and

²¹See also Bali, Cakici, and Whitelaw (2011).

MZZ are excluded due to their inability to gain traction among investors from 2006 through 2019, in particular MZZ. Aside from excluding MVV and MZZ, I use the remaining six original leveraged ETFs (three $2\times$ and three $-2\times$) to avoid cherry-picking based on realized outcomes.

The index is constructed in the following manner. Of the six leveraged ETFs, J denotes the set of leveraged-long ETFs and K denotes the set of leveraged-short ETFs. In each month t , ETF i 's percent share change is computed as

$$(1) \quad \Delta_{i,t} = \frac{SO_{i,t}}{SO_{i,t-1}} - 1,$$

in which $SO_{i,t}$ is the ETF's shares outstanding in month t and $t-1$ denotes the previous month. $\Delta_{i,t}$ can be negative valued (ETF shares are redeemed in net) or $\Delta_{i,t}$ can be positive valued (ETF shares are created in net). Both negative and positive values of $\Delta_{i,t}$ imply net arbitrage activity, with the sign on $\Delta_{i,t}$ providing the direction.

Once percent share changes are computed, month t 's SSI level is computed as the net difference in share changes for leveraged-long ETFs and leveraged-short ETFs

$$(2) \quad SSI_t = \sum_{i \in J} \Delta_{i,t} - \sum_{i \in K} \Delta_{i,t}.$$

Equation (2) represents the net demand shock in the set of leveraged ETFs. For example, if SSI is near zero then the implicit demand shock that generates mispricing is either small or it affects leveraged-long and leverage-short ETFs equally. Conversely, if SSI is large and positive, the demand shock favors leveraged-long products. If SSI is large and negative, the demand shock favors leveraged-short products. While SSI is an intuitive measure, I provide a model in Section IA.2 of the Supplementary Material that provides theoretical motivation for using it. Moreover, by netting the leveraged-long ETFs' share change and the leveraged-short ETFs' share change, other nonfundamental demand shocks are attenuated. For example, if there is a shock to arbitrageurs' liquidity, the shock should affect leveraged-long and leveraged-short ETF share change in the same direction. Thus, netting share change would also net out the shock to arbitrageur liquidity.²² See Remark IA3 in Section IA.2 of the Supplementary Material for additional discussion.

While the economics of SSI are examined in the subsequent sections, it is worth highlighting one feature of the index here as it relates to the methodology. SSI is basic to construct as one only needs to observe monthly shares outstanding for six ETFs. While simple, the method appears to capture the main driver of share change in the set of ETFs; a more sophisticated method using a principal components analysis (PCA) yields nearly identical results. If one performs PCA on monthly

²²While this article primarily relies on a monthly construction of SSI, it is possible to compute the index at a daily and weekly frequency because share change data are available daily. However, in Section IA.3 of the Supplementary Material, I provide a discussion about potential shortcomings in daily and weekly measures due to stale data and strategic delay by APs in creating new ETF shares. The monthly measure does not suffer from these shortcomings.

percent share changes in the six ETFs, the first principal component explains over 50% of the joint variation (if share changes across the six ETFs were independent, the first principal component would explain one-sixth of the joint variation or 16.7%). Furthermore, the linear weights associated with forming the first principal component from the original data include three that are positive valued and three that are negative valued. The three positive valued linear weights are assigned to the leveraged-long ETFs and the three negative valued linear weights are assigned to the leveraged-short ETFs. The first principal component has a correlation coefficient of 0.94 with SSI. Because PCA is agnostic to economic interpretation, in many settings it is difficult to explain which economic force a particular principal component embodies. In this setting, however, the interpretation is straightforward: the net bullish/bearish sentiment measured by the difference between leveraged-long and leveraged-short ETFs' share change is the primary driver of fund-level arbitrage activity.

IV. Speculation Sentiment and Aggregate Returns

Under my identifying assumption that leveraged ETF share demand is relatively more sensitive to gambling-like, uninformed demand, SSI proxies for market-wide speculative demand shocks. In this section, I examine the relation between i) SSI and contemporaneous aggregate stock returns and ii) SSI and future aggregate stock returns. Under the null, SSI should not be related to contemporaneous returns nor should SSI predict returns. However, I find that SSI is negatively related to contemporaneous returns and it has substantial predictive power. The results are consistent with SSI measuring speculative demand shocks, which distort stock prices.

To begin, I perform regressions using SSI and one of three benchmark indices: i) the CRSP equal-weighted index, ii) the CRSP value-weighted index, and iii) the S&P 500 index. The regressions examine the contemporaneous relation between monthly SSI and that month's return in each of the three indices

$$(3) \quad r_t = a + \beta \text{SSI}_t + \epsilon_t,$$

in which r_t is either the CRSP equal-weighted index monthly return, the CRSP value-weighted index monthly return, or the S&P 500 index monthly return in month t , a is the regression intercept, SSI_t is the contemporaneous value of SSI, β is the regression coefficient, and ϵ_t is the regression error term. The results for the regressions are reported in Panel A of Table 3. Results for the CRSP equal-weighted index are reported as regression 1, results for the CRSP value-weighted index are reported as regression 2, and results for the S&P 500 index are reported as regression 3. The sample's index returns are from Oct. 2006 through Nov. 2019. SSI is standardized and index returns are reported as percentages so that β may be interpreted as the effect of a 1-standard-deviation increase in SSI on contemporaneous returns (throughout the article all variables, other than returns, are standardized, unless otherwise stated). In regressions 1–3, the coefficients are statistically significant at a 1% p -value threshold; for the CRSP equal-weighted index, a 1-standard-deviation increase in SSI is associated with a

TABLE 3
Contemporaneous Regressions and Predictive Regressions with SSI

In Panel A of Table 3, the CRSP equal-weighted, CRSP value-weighted, or S&P 500 index monthly returns are regressed on the contemporaneous Speculation Sentiment Index value: $r_t = a + \beta SSI_t + \epsilon_t$ in which r_t is the index monthly return, SSI_t is the contemporaneous Speculation Sentiment Index monthly value, β is the estimated coefficient on SSI_t , and ϵ_t is the error term. In Panel B, the future CRSP equal-weighted, CRSP value-weighted, or S&P 500 index monthly returns are regressed on the Speculation Sentiment Index value: $r_{t+1} = a + \beta SSI_t + \epsilon_{t+1}$ in which r_{t+1} is the future index monthly return, SSI_t is the Speculation Sentiment Index value, β is the estimated coefficient on SSI_t , and ϵ_{t+1} is the error term. In addition to the predictive regression results, BIAS and PVAL report the coefficient β 's bias and the coefficient β 's p-value from a parametric bootstrap. 90th-PCT, 95th-PCT, and 99th-PCT report the critical values of adjusted R^2 from a parametric bootstrap. For Panel A, in regressions 1–3, the sample returns run from Oct. 2006 to Nov. 2019 and in regressions 4–6, the sample returns run from Dec. 2009 to Nov. 2019. For Panel B, in regressions 1–3, the sample returns run from Nov. 2006 to Dec. 2019 and in regressions 4–6, the sample returns run from Jan. 2010 to Dec. 2019. White standard errors are used to account for heteroscedasticity and t -statistics are reported, in parenthesis, below each estimated coefficient. All variables, except for returns, are standardized. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Full Sample			Post-2009		
	EW	VW	SP500	EW	VW	SP500
	1	2	3	4	5	6
<i>Panel A. Contemporaneous Regressions r_t</i>						
SSI_t	-3.28*** (-9.19)	-2.71*** (-9.34)	-2.52*** (-9.02)	-2.98*** (-4.13)	-2.75*** (-4.39)	-2.66*** (-4.46)
Adj. R^2	0.44	0.39	0.36	0.16	0.16	0.16
N	158	158	158	120	120	120
<i>Panel B. Predictive Regressions r_{t+1}</i>						
SSI_t	-1.67*** (-3.08)	-1.27*** (-2.67)	-1.14** (-2.57)	-1.65** (-2.49)	-1.31** (-2.08)	-1.19* (-1.93)
Adj. R^2	0.11	0.08	0.07	0.04	0.03	0.02
N	158	158	158	120	120	120
BIAS	0.03	0.03	0.03	0.03	0.02	0.02
PVAL	0.00	0.01	0.02	0.02	0.04	0.06
90th-PCT	0.01	0.01	0.01	0.01	0.01	0.01
95th-PCT	0.02	0.02	0.02	0.02	0.02	0.02
99th-PCT	0.04	0.04	0.04	0.05	0.05	0.05

contemporaneous 3.28% decline in the index. For the CRSP value-weighted index, the effect is slightly smaller with a decline of 2.71%. For the S&P 500, the effect is also smaller with a decline of 2.52%.

As can be seen in Graph A of Figure 2, the index exhibits the most pronounced swings just prior, during, and immediately after the 2008–2009 Financial Crisis. Thus, to ensure that the results are not driven by the 2008–2009 Financial Crisis, regressions 4–6 in Panel A of Table 3 use a data sample that begins in Jan. 2010 (SSI is not restandardized). The coefficients remain relatively stable in magnitude and also remain statistically significant. Regressions 1–6 in Panel A of Table 3 provide evidence that SSI is largely contrarian. In Section IA.5 of the Supplementary Material, I show that the negative relation between SSI and contemporaneous returns also holds at higher frequencies using daily data.

Next, I perform predictive regressions to examine the ability of monthly SSI to predict the next month's return in each of the three indices

$$(4) \quad r_{t+1} = a + \beta SSI_t + \epsilon_{t+1},$$

in which r_{t+1} is either the CRSP equal-weighted index monthly return, the CRSP value-weighted index monthly return, or the S&P 500 index monthly return in month $t + 1$, a is the regression intercept, SSI_t is the monthly value of SSI, β is the regression coefficient, and ϵ_{t+1} is the regression error term. The results for the regressions are reported in Panel B of Table 3. Results for the CRSP equal-weighted

index are reported as regression 1, results for the CRSP value-weighted index are reported as regression 2, and results for the S&P 500 index are reported as regression 3. In regressions 1–3, the coefficients are statistically significant at a 5% p -value threshold or lower. Moreover, the results are large in economic magnitude; for the CRSP equal-weighted index, a 1-standard-deviation increase in SSI predicts a 1.67% decline in the index the following month. For the CRSP value-weighted index, the effect is slightly smaller with a predicted decline of 1.27%. For the S&P 500, the effect is also smaller with a predicted decline of 1.14%. To ensure that the results are not driven by the 2008–2009 Financial Crisis, regressions 4–6 in Panel B of Table 3 use a data sample that begins in Jan. 2010; the coefficients are nearly identical and remain statistically significant. Furthermore, in regressions 1–6, the coefficients on SSI with CRSP equal-weighted index returns are the largest in magnitude and the coefficients on SSI with S&P 500 index returns are the smallest in magnitude. This rank order of coefficients appears consistently throughout the article and it suggests that speculative demand shocks disproportionately affect smaller capitalization stocks.

To provide additional support for the results in Panel B of Table 3, I also perform a parametric bootstrap, which allows me to estimate the bias in β .²³ The biases of β in regressions 1–6, which are reported in Panel B of Table 3, are not only close to zero, they are also *positive*. Thus, concerns about a biased estimate appear minimal.²⁴

The parametric bootstrap also yields a simulated distribution of the t -statistics, which allows the ability to construct p -values using the simulated distribution. The p -values from the simulated distribution are also reported in Panel B of Table 3. In the full sample, each of the coefficients is statistically significant at a 5% p -value threshold or lower and in the post-2009 sample, each of the coefficients is statistically significant at a 10% p -value threshold or lower. The p -values from the simulated distribution provide additional evidence to those using asymptotic standard errors that the relation between SSI and future returns is statistically significant.

The regressions in Panel B of Table 3 also provide economically meaningful adjusted R^2 s: for the CRSP equal-weighted index, the adjusted R^2 in regression 1 is 11%. For the CRSP value-weighted index, the adjusted R^2 in regression 2 is 8%. For the S&P 500 index, the adjusted R^2 in regression 3 is 7%. Using the simulated data from the parametric bootstrap, the 90th, 95th, and 99th percentile adjusted R^2 s are also included in Panel B of Table 3. Accordingly, the realized adjusted R^2 s exceed the 99th percentile in regressions 1–3. Similarly, the adjusted R^2 s exceed the 95th percentile in regressions 4–6, which correspond to

²³Details of the parametric bootstrap procedure are provided in Section IA.4 of the Supplementary Material.

²⁴Naturally, one might be concerned about persistence in SSI and the possibility of a Stambaugh-bias (Stambaugh (1999)). I document the autocorrelation of SSI in Section IA.7.1 of the Supplementary Material. While the 1-month lagged value SSI_{t-1} has predictive power, the coefficient is estimated to be only 0.29. Nevertheless, for robustness, I provide an alternative index by estimating SSI as an AR(1) process and by forming the new index from the process's innovations. Table IA4 of the Supplementary Material shows that the return predictability results are qualitatively identical when using the autocorrelation corrected index SSI^{AR} in place of SSI.

the post-2009 sample. Therefore, SSI explains a substantial proportion of future return variation in each of the three benchmark indices. The results of Panel B of Table 3 provide strong evidence that SSI measures nonfundamental demand, which distorts prices and gives rise to return predictability.²⁵

Two special disclaimers here are in order. First, it is worthwhile to reiterate that this is not a price pressure story from leveraged ETF trading. After all, leveraged ETFs are tiny relative to the size of the broad market and the mispricing shown in Table 3. Rather, I argue that leveraged ETFs provide a unique setting to identify and measure market-wide speculative demand shocks. As additional evidence of this, the model in Section IA.2 of the Supplementary Material demonstrates that SSI theoretically contains information about market-wide speculative demand shocks, even if trades in the leveraged ETF primary market have no price impact on the broad market. In particular, see Remark IA4 of the Supplementary Material.

Second, throughout the empirical analysis, a negative relation between SSI and contemporaneous returns is documented (i.e., the empirical *measure* of SSI is contrarian). While the negative relation is consistent with speculative demand shocks *themselves* being contrarian, one must recognize that there is no identification to show this. For example, SSI may jointly measure two pieces: speculative demand and rational rebalancing demand. While Section V.A shows that SSI's ability to predict returns is driven by speculative demand shocks and not rational rebalancing demand, the same type of analysis cannot be done with contemporaneous returns. Specifically, rational rebalancing demand is likely highly correlated (perhaps even collinear) with contemporaneous returns making it nearly impossible to show that speculative demand shocks are negatively correlated with contemporaneous returns. As such, throughout the article, I am careful to not classify speculative demand itself as a contrarian.

In Section V, I provide additional results to speak to the robustness of the return predictability results. First, I show that SSI is distinct from rational rebalancing, which strengthens the interpretation that SSI is a measure of sentiment. Second, I examine the return predictability at both shorter horizons and longer horizons. The shorter horizon results provide evidence that SSI measures sentiment shocks as opposed to price pressures from rational trading. The longer horizon results provide new evidence that speculation sentiment shocks take several months to correct. Third, I consider the robustness of the return predictability results with the addition of other market and sentiment controls and I also provide evidence that SSI has out-of-sample predictability.

V. Economic Insights and Robustness

The results in Section IV provide evidence that speculative demand shocks i) affect traders en masse and ii) lead to price distortions that subsequently reverse. While the preceding results are both economically meaningful and

²⁵Note the model in Section IA.2 of the Supplementary Material highlights that a negative relation between SSI and future returns is a symptom that SSI measures nonfundamental demand. See Remark IA1 of the Supplementary Material.

statistically significant, they demand additional investigation. First, one may be concerned that the results are driven by rational trading (namely, portfolio rebalancing), that SSI simply measures this rational trading, and that the sentiment interpretation is misleading. As such, it is important to show that SSI is distinct from a measure of rational rebalancing. Second, it is worthwhile to document the horizon at which speculative demand shocks resolve themselves to help distinguish between rational price pressure interpretations versus sentiment interpretations. Third, it is important to evaluate the robustness of SSI's ability to predict returns both with control variables and in out-of-sample tests.

A. Speculation Sentiment Versus Rational Rebalancing

In this subsection, I provide evidence that the return predictability results in [Section IV](#) are consistent with a sentiment interpretation as opposed to a price pressure story from rational trading. Specifically, rather than measuring sentiment, it is possible that SSI relates to rational portfolio rebalancing and that those rebalancing trades have price impact that subsequently reverses. Such a possibility is consistent with the evidence in [Table 3](#); SSI is negatively related to contemporaneous returns (i.e., the rebalancing) and SSI negatively predicts subsequent returns (i.e., the reversal of rebalancing trades' price impact).²⁶ Thus, it is possible that the contrarian nature and return predictability of SSI is driven by an omitted variable that relates to rational trading.

Rational portfolio rebalancing is highly correlated (if not collinear) with returns. As such, one can examine whether or not SSI proxies for rational rebalancing with additional tests using returns. The first natural test is to examine whether or not aggregate returns predict subsequent returns during the sample period. That is, it is possible that a spurious *positive* autocorrelation in returns exists during the years used in my sample. If such a relation exists in the time series of returns, a reasonable concern is that SSI is simply a proxy for realized returns and the results from [Section IV](#) reflect this. Consequently, I perform predictive regressions of the form

$$(5) \quad r_{t+1} = a + \beta r_t + \epsilon_{t+1},$$

in which r_{t+1} is either the CRSP equal-weighted index monthly return, the CRSP value-weighted index monthly return, or the S&P 500 index monthly return in month $t + 1$, a is the regression intercept, r_t is the monthly index return in month t , β is the regression coefficient, and ϵ_{t+1} is the regression error term. The results are reported in Panel A of [Table 4](#) and results for the CRSP equal-weighted index are reported as regression 1, results for the CRSP value-weighted index are reported

²⁶As an example, in Section IA.6 of the Supplementary Material, I show that an investor using leveraged ETFs to obtain a leveraged portfolio must rebalance her portfolio daily to retain the target leverage quantity. The daily rebalancing is mechanically contrarian; it sells leveraged-long ETFs (buys leveraged-short ETFs) in up markets and buys leveraged-long ETFs (sells leveraged-short ETFs) in down markets. As another example, SSI may reflect market-wide asset allocation shifts as investors sell equities after good performance and buy equities after poor performance to maintain portfolio weights. If rational actions, like leverage rebalancing and asset reallocations, have price impact these actions could give rise to return reversals.

TABLE 4
 Predictive Regressions with Returns and SSI

In Panel A of Table 4, the future CRSP equal-weighted, CRSP value-weighted, or S&P 500 index monthly returns are regressed on their contemporaneous monthly returns: $r_{t+1} = a + \beta r_t + \epsilon_{t+1}$ in which r_{t+1} is the future index monthly return, r_t is the contemporaneous index monthly return, β is the estimated coefficient on r_t , and ϵ_{t+1} is the error term. In Panel B, the future CRSP equal-weighted, CRSP value-weighted, or S&P 500 index monthly returns are regressed on SSI and their contemporaneous monthly returns: $r_{t+1} = a + \beta_{SSI} SSI_t + \beta_r r_t + \epsilon_{t+1}$ in which r_{t+1} is the future index monthly return, SSI_t is the monthly value of SSI, r_t is the contemporaneous index monthly return, β_{SSI} is the estimated coefficient on SSI, β_r is the estimated coefficient on r_t and ϵ_{t+1} is the error term. For both panels, in regressions 1–3, the sample returns run from Nov. 2006 to Dec. 2019 and in regressions 4–6, the sample returns run from Jan. 2010 to Dec. 2019. White standard errors are used to account for heteroscedasticity and *t*-statistics are reported, in parenthesis, below each estimated coefficient. All variables, except for returns, are standardized. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Full Sample			Post-2009		
	EW 1	VW 2	SP500 3	EW 4	VW 5	SP500 6
<i>Panel A. Univariate Predictive Regressions r_{t+1}</i>						
r_t	0.19 (1.49)	0.10 (0.83)	0.09 (0.79)	-0.08 (-0.66)	-0.16 (-1.39)	-0.17 (-1.54)
Adj. R^2	0.03	0.00	0.00	0.00	0.02	0.02
<i>N</i>	158	158	158	120	120	120
<i>Panel B. Bivariate Predictive Regressions r_{t+1}</i>						
SSI_t	-1.87*** (-2.80)	-1.65*** (-2.66)	-1.43** (-2.45)	-2.27*** (-3.07)	-2.11*** (-2.93)	-1.97*** (-2.79)
r_t	-0.06 (-0.48)	-0.14 (-1.09)	-0.12 (-0.87)	-0.21* (-1.82)	-0.29*** (-2.65)	-0.30*** (-2.69)
Adj. R^2	0.10	0.09	0.07	0.07	0.09	0.09
<i>N</i>	158	158	158	120	120	120

as regression 2, results for the S&P 500 index are reported as regression 3 and regressions 4–6 use the data sample that begins in Jan. 2010. The results show no predictive power from returns. In other words, it does not appear that SSI is simply a proxy for realized returns.

To further examine the relation between SSI and contemporaneous returns, Panel B of Table 4 reports the results of the bivariate predictive regression

$$(6) \quad r_{t+1} = a + \beta_{SSI} SSI_t + \beta_r r_t + \epsilon_{t+1},$$

in which r_{t+1} is either the CRSP equal-weighted index monthly return, the CRSP value-weighted index monthly return, or the S&P 500 index monthly return in month $t + 1$, a is the regression intercept, SSI_t is the monthly value of SSI, r_t is the monthly index return in month t , β_{SSI} is the regression coefficient on SSI_t , β_r is the regression coefficient on r_t , and ϵ_{t+1} is the regression error term. A key insight emerges from Panel B of Table 4: the return predictability from SSI, as compared to the results in Table 3, is *stronger* once one controls for contemporaneous returns (both in economic magnitude and statistical significance).²⁷ The results in Table 4

²⁷In Section IA.6 of the Supplementary Material, I solve analytically for the rebalancing required by a trader who wants to retain a target leverage quantity. Using the analytic solution, I compute a measure of implied rebalancing and I replicate Table 4 with the measure in place of r_t . The results, which are reported in Table IA2 of the Supplementary Material, are qualitatively the same as those reported in Table 4.

show that the return predictability from SSI is distinct from a rational trading story and, instead, more consistent with a sentiment interpretation.

B. Return Predictability Horizons

While I focus primarily on monthly return predictability, there is no obvious time period over which speculative demand shocks should resolve themselves. In this subsection, I examine monthly SSI's ability to predict cumulative returns at both shorter horizons (i.e., the first few days of the preceding month) and longer horizons (i.e., the preceding 6 months).

Beginning with shorter horizons, [Table 5](#) reports univariate regression results in which SSI in month t predicts daily cumulative returns over the first 20 days following month t .²⁸ Results for the CRSP equal-weighted index are reported in Panel A of [Table 5](#), results for the CRSP value-weighted index are reported in Panel B of [Table 5](#), and results for the S&P 500 index are reported in Panel C of [Table 5](#). The results in the panels show a consistent theme; initially, there is little to no return predictability following month t . However, after a few days, return predictability is negative, statistically significant, and increasing in magnitude through day 20. The results complement those from [Table 3](#) and provide additional evidence that SSI measures sentiment shocks as opposed to measuring price pressures from leveraged ETF trading. That is, since the return predictability is not concentrated in the first few days following month t , it helps to further rule out a price pressure story (e.g., rational rebalancing or reallocation).

Turning to longer horizons, [Table 6](#) provides univariate regression results in which SSI predicts cumulative returns over 1, 2, 3, 4, 5, and 6 months in each of the three indices. In [Table 6](#), results for the CRSP equal-weighted index are reported in Panel A, results for the CRSP value-weighted index are reported in Panel B, and results for the S&P 500 index are reported in Panel C. The sample's index returns run from Nov. 2006 through July 2019 (the later dates in 2019 are excluded so that there are an equal number of observations for each return horizon). Hodrick (1992) standard errors are reported because of the mechanical autocorrelation introduced by the dependent variable's overlapping periods.

SSI predicts economically and statistically significant returns out 4–6 months in the three indices in [Table 6](#). However, the vast majority of the predicted return is earned in the first 4 months and a significant fraction is earned in the first month.²⁹

²⁸Note, while there is an average of 21 trading days in each month of the 2006–2019 data sample, the estimated coefficients of β in [Table 5](#) for r_{t+20d} slightly differ from the estimates in [Table 3](#) for r_{t+1} (i.e., the 20-day coefficients differ from the monthly coefficients). There are two reasons for the apparent differences. First, the number of trading days in the monthly returns of [Table 3](#) range from 19 to 23. Second, CRSP monthly returns are not compounded from daily returns. Instead, ordinary dividends are reinvested at month-end for monthly returns and are reinvested on the ex-distribution date for daily returns. Nevertheless, in unreported analysis, reconstructing [Table 3](#) from compounded daily returns yields nearly identical results to those in [Table 3](#).

²⁹Note, for the regressions in which the dependent variable is r_{t+1} , the t -statistics differ from those in [Table 3](#) because Hodrick (1992) standard errors are used.

TABLE 5
Return Predictability with SSI at Short Horizons

Each column in Table 5 represents a regression in which the CRSP equal-weighted, CRSP value-weighted, or S&P 500 index cumulative future returns are regressed on the Speculation Sentiment Index value: $r_{t+i} = a + \beta SSI_t + \epsilon_{t+i}$ in which r_{t+i} is the index cumulative return over the first i days following month t , SSI_t is the Speculation Sentiment Index value in month t , β is the estimated coefficient on SSI_t , and ϵ_{t+i} is the error term. The cumulative return r_{t+i} is measured at horizons of 1–10 days, 15 days, and 20 days. Panel A reports the results for the CRSP equal-weighted index, Panel B reports the results for the CRSP value-weighted index, and Panel C reports the results for the S&P 500 index. White standard errors are used to account for heteroscedasticity and t -statistics are reported, in parenthesis, below each estimated coefficient. The sample returns run from Nov. 2006 to Dec. 2019. All variables, except for returns, are standardized. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	r_{t+1d}	r_{t+2d}	r_{t+3d}	r_{t+4d}	r_{t+5d}	r_{t+6d}	r_{t+7d}	r_{t+8d}	r_{t+9d}	r_{t+10d}	r_{t+15d}	r_{t+20d}
	1	2	3	4	5	6	7	8	9	10	11	12
<i>Panel A. EW CRSP</i>												
SSI_t	-0.21 (-1.42)	-0.46* (-1.91)	-0.68*** (-3.77)	-1.10*** (-4.28)	-1.12*** (-3.15)	-1.34*** (-3.23)	-1.58*** (-3.15)	-1.69*** (-3.20)	-1.32*** (-3.67)	-1.55*** (-4.19)	-1.90*** (-3.16)	-1.91*** (-2.83)
Adj. R^2	0.02	0.07	0.10	0.18	0.14	0.15	0.20	0.19	0.14	0.16	0.15	0.12
N	158	158	158	158	158	158	158	158	158	158	158	158
<i>Panel B. VW CRSP</i>												
SSI_t	-0.25 (-1.62)	-0.46* (-1.84)	-0.67*** (-4.09)	-1.05*** (-4.57)	-1.01*** (-2.91)	-1.13*** (-2.98)	-1.35*** (-2.69)	-1.50*** (-2.75)	-1.06*** (-2.99)	-1.33*** (-3.56)	-1.60*** (-2.79)	-1.60** (-2.61)
Adj. R^2	0.02	0.07	0.12	0.19	0.14	0.13	0.17	0.17	0.11	0.14	0.12	0.10
N	158	158	158	158	158	158	158	158	158	158	158	158
<i>Panel C. S&P 500</i>												
SSI_t	-0.25* (-1.69)	-0.42* (-1.79)	-0.62*** (-4.14)	-0.99*** (-4.56)	-0.96*** (-2.92)	-1.07*** (-2.96)	-1.28*** (-2.61)	-1.43*** (-2.70)	-1.00*** (-3.00)	-1.24*** (-3.58)	-1.47*** (-2.73)	-1.43*** (-2.54)
Adj. R^2	0.02	0.06	0.11	0.18	0.13	0.12	0.16	0.16	0.10	0.13	0.11	0.09
N	158	158	158	158	158	158	158	158	158	158	158	158

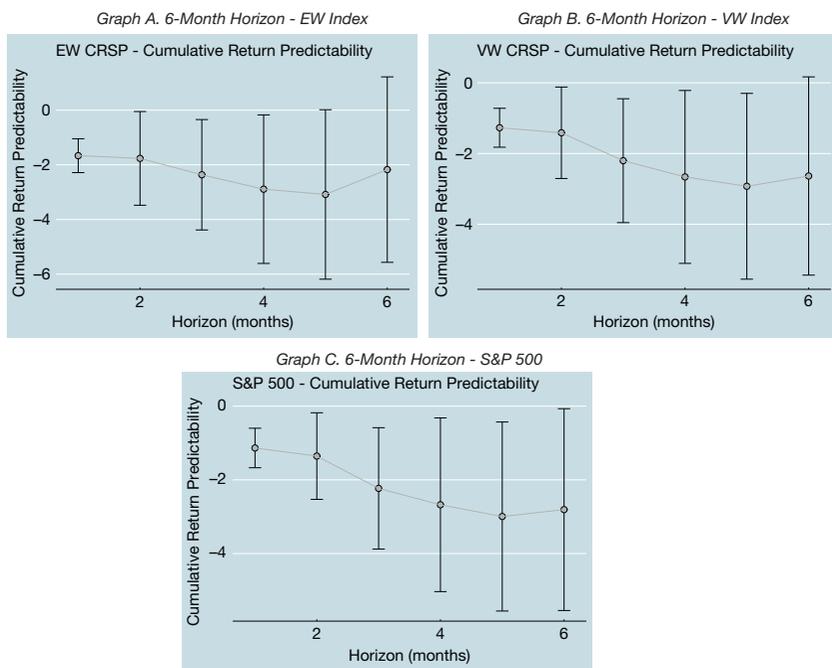
TABLE 6
Return Predictability with SSI at Long Horizons

Each column in Table 6 represents a regression in which the CRSP equal-weighted, CRSP value-weighted, or S&P 500 index cumulative future returns are regressed on the Speculation Sentiment Index value: $r_{t+i} = a + \beta SSI_t + \epsilon_{t+i}$ in which r_{t+i} is the index cumulative return over the next i months, SSI_t is the Speculation Sentiment Index value, β is the estimated coefficient on SSI_t , and ϵ_{t+i} is the error term. The cumulative return r_t is measured at one ($i=1$), two ($i=2$), three ($i=3$), four ($i=4$), five ($i=5$), and six ($i=6$) month horizons. Panel A reports the results for the CRSP equal-weighted index, Panel B reports the results for the CRSP value-weighted index, and Panel C reports the results for the S&P 500 index. Standard errors are based on Hodrick (1992), using code from Alexander Chincó's website, and t -statistics are reported, in parenthesis, below each estimated coefficient. The sample returns run from Nov. 2006 to July 2019 (the later dates in 2019 are excluded so that there are an equal number of observations for each return horizon). All variables, except for returns, are standardized. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	r_{t+1}	r_{t+2}	r_{t+3}	r_{t+4}	r_{t+5}	r_{t+6}
	1	2	3	4	5	6
<i>Panel A. EW CRSP</i>						
SSI_t	-1.67*** (-4.42)	-1.76* (-1.69)	-2.36* (-1.93)	-2.89* (-1.75)	-3.08 (-1.64)	-2.18 (-1.05)
Adj. R^2	0.11	0.05	0.05	0.06	0.05	0.01
N	153	153	153	153	153	153
<i>Panel B. VW CRSP</i>						
SSI_t	-1.27*** (-3.80)	-1.41* (-1.80)	-2.20** (-2.07)	-2.66* (-1.79)	-2.93* (-1.83)	-2.63 (-1.55)
Adj. R^2	0.08	0.04	0.07	0.08	0.07	0.04
N	153	153	153	153	153	153
<i>Panel C. S&P 500</i>						
SSI_t	-1.14*** (-3.48)	-1.36* (-1.90)	-2.23** (-2.23)	-2.68* (-1.87)	-2.99* (-1.92)	-2.81* (-1.69)
Adj. R^2	0.07	0.04	0.08	0.09	0.08	0.06
N	153	153	153	153	153	153

FIGURE 3
Return Predictability Horizons

Each figure depicts the regression coefficients β from the regressions $r_{t+i} = a + \beta \text{SSI}_t + \epsilon_{t+i}$ in which r_{t+i} is either the CRSP equal-weighted index cumulative return, the CRSP value-weighted index cumulative return, or the S&P 500 index cumulative return through month $t+i$, SSI_t is the Speculation Sentiment Index value in month t , and ϵ_{t+i} is the error term. The cumulative return r_{t+i} is measured at one ($i=1$), two ($i=2$), three ($i=3$), four ($i=4$), five ($i=5$), and six ($i=6$) month horizons. 90% confidence intervals are depicted as the error bars. Standard errors are based on Hodrick (1992), using code from Alexander Chincó's website. In each graph, the vertical axis represents the coefficient β from the univariate regressions and the horizontal axis represents the number of months over which the cumulative return is calculated. The sample returns run from Nov. 2006 to July 2019 (the dates later in 2019 are excluded so that there are an equal number of observations for each return horizon).



Thus, while I focus on monthly return predictability, there is evidence that speculative demand shocks may take several months to fully resolve themselves.

The results in Table 6 are also depicted in Figure 3. Each panel in Figure 3 depicts the results for either the CRSP equal-weighted index, the CRSP value-weighted index, or the S&P 500 index. The vertical axis represents the coefficient β from the univariate regressions and the horizontal axis represents the number of months the cumulative return is calculated over. 90% confidence intervals are depicted with each data point using Hodrick (1992) standard errors. In all three panels, predicted returns are evident in the first month and continue to grow until approximately month four.

C. Return Predictability with Controls

A reasonable concern is that SSI is highly correlated with a known predictor of aggregate returns or that it encapsulates an established sentiment proxy. To this end, I control for other known predictors of returns and sentiment proxies by performing the predictive bivariate regression

$$(7) \quad r_{t+1} = a + \beta \text{SSI}_t + \gamma \text{CONT}_t + \epsilon_{t+1},$$

in which CONT_t is a control variable and γ is the coefficient on CONT_t . The results for the regressions are reported in Table 7. The additional control variables with the regression number included in parenthesis are cyclically adjusted earnings-to-price (CAEP) (1), term spread (TERM) (2), dividend-to-price (DP) (3), short-rate (RATE) (4), variance risk premium (VRP) (5), intermediary capital risk factor (INTC) (6), innovation to aggregate liquidity (ΔLIQ) (7), short interest (SHORT) (8), VIX (VIX) (9), Baker–Wurgler investor sentiment level (SENT) (10), aligned investor sentiment level (HJTZ) (11), closed-end fund discount (CEFD) (12), consumer confidence level (CONF) (13), change in consumer confidence level (ΔCONF) (14), and investor lottery demand (FMAX) (15).³⁰ Section IA.9 of the Supplementary Material studies the correlations between SSI and the control variables.

Beginning with the CRSP equal-weighted index results in Panel A of Table 7, the coefficients on SSI are statistically significant at a 5% p -value threshold or better in each regression, except for regression 5, which corresponds to the variance risk premium. In that regression, the coefficient on SSI is statistically significant at a 10% p -value threshold. Moreover, the coefficient values range from -1.86 to -1.14 . Turning to the CRSP value-weighted index results in Panel B, the coefficients on SSI are statistically significant at the 5% p -value threshold or better for all regressions except for 5. The coefficients in CRSP value-weighted index regressions range from -1.33 to -0.85 . Finally, for the S&P 500 index results in Panel C, the coefficients on SSI are statistically significant at the 5% p -value threshold or better for all regressions except for regression 5. The coefficients in the S&P 500 index regressions range from -1.18 to -0.75 . Together, the bivariate regression results in Table 7 demonstrate an economically meaningful and statistically significant relation between SSI and future market returns, even after controlling for other known predictors of returns and sentiment proxies.

It is worthwhile to focus on one particular control variable, the variance risk premium which is the control in regression 5; in Table 7, the regressions which use the variance risk premium as a control consistently show smaller coefficients on SSI and less statistical significance. Therefore, special attention is required in exploring the relation between SSI and VRP. VRP is the spread between implied and realized variance and several studies argue that it serves as a proxy for aggregate market risk aversion (Rosenberg and Engle (2002), Bakshi and Madan (2006), and Bollerslev et al. (2009)). Consistent with serving as a proxy for aggregate market risk aversion, large values of VRP are associated with higher subsequent returns. However, the correlation coefficient for VRP and SSI is -0.45 , highlighting a common component of both variables. As SSI measures realized differences in demand between unsophisticated speculators and sophisticated institutions, the strong negative correlation between VRP and SSI suggests that the variance risk premium, in addition to being correlated with aggregate risk aversion, may also contain

³⁰Throughout the article, I use average monthly VIX index values. End-of-month values are highly correlated with monthly average values (0.96 correlation coefficient during the paper's sample period of Nov. 2006–Dec. 2019).

TABLE 7
Return Predictability with SSI and Controls

The regressions regress the future CRSP equal-weighted, CRSP value-weighted, or S&P 500 index monthly returns on the Speculation Sentiment Index value and a control variable: $r_{t+1} = a + \beta SSI_t + \gamma CONT_t + \epsilon_{t+1}$ in which r_{t+1} is the future index monthly return, SSI_t is the Speculation Sentiment Index value, β is the estimated coefficient on SSI_t , $CONT_t$ is a control variable, γ is the estimated coefficient on $CONT_t$, and ϵ_{t+1} is the error term. The control variables are cyclically adjusted earnings-to-price (CAEP), term spread (TERM), dividend-to-price (DP), short spread (SHORT), variance risk premium (VRP), intermediary capital risk factor (INTC), innovation to aggregate liquidity (ΔLIQ), short interest (SHORT), VIX (VIX), Baker–Wurgler sentiment level (SENT), aligned investor sentiment level (HJTZ), closed-end fund discount (CEFD), consumer confidence level (CONF), change in consumer confidence ($\Delta CONF$), and investor lottery demand (FMAX). Panel A reports the results for the CRSP equal-weighted index, Panel B reports the results for the CRSP value-weighted index, and Panel C reports the results for the S&P 500 index. White standard errors are used to account for heteroscedasticity and t -statistics are reported, in parenthesis, below each estimated coefficient. The sample returns run from Nov. 2006 to Dec. 2019 (if the control variable is available through 2019). All variables, except for returns, are standardized. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	CAEP	TERM	DP	RATE	VRP	INTC	ΔLIQ	SHORT	VIX	SENT	HJTZ	CEFD	CONF	$\Delta CONF$	FMAX
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
<i>Panel A. EW CRSP</i>															
SSI_t	-1.68*** (-3.01)	-1.68*** (-3.08)	-1.72*** (-3.05)	-1.68*** (-3.09)	-1.14* (-1.83)	-1.63** (-2.40)	-1.51*** (-3.05)	-1.56*** (-2.91)	-1.86*** (-3.28)	-1.58*** (-2.85)	-1.59*** (-2.96)	-1.74*** (-3.14)	-1.69*** (-3.08)	-1.66*** (-3.01)	-1.49** (-2.47)
$CONT_t$	0.94* (1.66)	0.21 (0.63)	1.04 (1.59)	-0.53* (-1.80)	1.18** (2.28)	0.10 (0.16)	0.64 (1.16)	-0.87** (-2.10)	0.69 (1.15)	-0.98** (-2.53)	-0.26 (-0.49)	0.94** (2.38)	-0.24 (-0.46)	0.12 (0.29)	0.37 (0.68)
Adj. R^2	0.14	0.10	0.15	0.11	0.15	0.11	0.11	0.13	0.12	0.14	0.11	0.14	0.10	0.10	0.11
N	158	158	158	158	158	146	158	158	158	147	147	147	158	158	158
<i>Panel B. VW CRSP</i>															
SSI_t	-1.28*** (-2.64)	-1.27*** (-2.66)	-1.29*** (-2.67)	-1.28*** (-2.66)	-0.85 (-1.49)	-1.33** (-2.12)	-1.10*** (-2.71)	-1.21** (-2.56)	-1.32*** (-2.67)	-1.22** (-2.48)	-1.12** (-2.43)	-1.30*** (-2.68)	-1.26*** (-2.63)	-1.28*** (-2.63)	-1.31** (-2.35)
$CONT_t$	0.33 (0.70)	-0.05 (-0.17)	0.36 (0.66)	-0.29 (-1.04)	0.95* (1.91)	-0.13 (-0.24)	0.72 (1.35)	-0.49 (-1.32)	0.19 (0.38)	-0.48 (-1.41)	-0.50 (-1.06)	0.46 (1.35)	0.12 (0.28)	-0.09 (-0.24)	-0.08 (-0.16)
Adj. R^2	0.08	0.07	0.08	0.08	0.11	0.08	0.10	0.09	0.08	0.09	0.09	0.09	0.08	0.08	0.08
N	158	158	158	158	158	146	158	158	158	147	147	147	158	158	158
<i>Panel C. S&P 500</i>															
SSI_t	-1.14** (-2.55)	-1.13** (-2.57)	-1.15** (-2.58)	-1.14** (-2.57)	-0.75 (-1.40)	-1.18** (-1.99)	-0.96** (-2.53)	-1.08** (-2.46)	-1.15** (-2.49)	-1.09** (-2.38)	-0.97** (-2.26)	-1.16** (-2.59)	-1.12** (-2.52)	-1.15** (-2.55)	-1.19** (-2.24)
$CONT_t$	0.21 (0.46)	-0.07 (-0.23)	0.24 (0.45)	-0.31 (-1.14)	0.87* (1.76)	-0.09 (-0.19)	0.75 (1.44)	-0.46 (-1.27)	0.05 (0.10)	-0.43 (-1.31)	-0.54 (-1.18)	0.44 (1.35)	0.19 (0.46)	-0.11 (-0.31)	-0.11 (-0.22)
Adj. R^2	0.06	0.06	0.07	0.07	0.10	0.07	0.09	0.07	0.06	0.07	0.08	0.07	0.06	0.06	0.06
N	158	158	158	158	158	146	158	158	158	147	147	147	158	158	158

information about speculation sentiment. For example, it is possible that part of the spread between implied and realized volatility is driven by nonfundamental speculative demand shocks (e.g., bearish speculation sentiment bidding up put option prices; and implied volatility accordingly).

To further examine the relation between VRP and SSI, in unreported analysis, univariate regressions using VRP as a predictor of the CRSP equal-weighted index return, the CRSP value-weighted index return, and the S&P 500 index return are performed. The regressions use sample returns from Nov. 2006 through Dec. 2019 and yield coefficient estimates, with t -statistics in parenthesis, of 1.69 (4.10), 1.33 (3.24), and 1.21 (2.97), respectively. Moreover, adjusted R^2 s from the univariate VRP regressions are 0.11, 0.09, and 0.08 respectively. Note, from Table 3, the comparable SSI univariate predictive regression coefficients, with t -statistics in parenthesis, are -1.67 (-3.08), -1.27 (-2.67), and -1.14 (-2.57) and the adjusted R^2 s are 0.11, 0.08, and 0.07. Thus, both SSI and VRP are strong univariate predictors of aggregate returns and both have substantial adjusted R^2 s. When combined, the coefficients on both SSI and VRP are attenuated and so are the coefficients' t -statistics, as can be seen in Table 7. While the coefficients are weaker, the bivariate regressions display stronger adjusted R^2 s: for the CRSP equal-weighted index regression the value is 0.15, for the CRSP value-weighted index regression the value is 0.11, and for the S&P 500 index regression the value is 0.10. Therefore, the incremental explanatory power by combining SSI with VRP, and vice versa, is substantial. As such, while there is a common component shared by the two variables, both contain important and distinct information in explaining the variation of future returns.

In the Supplementary Material, I provide several alternative specifications of SSI as robustness tests: Section IA.7.1 provides a measure of SSI which is corrected for autocorrelation, Section IA.7.2 provides a measure of SSI which is orthogonal to aggregate ETF flows, Section IA.7.3 provides a measure of SSI which controls for the cost of arbitrage capital and other macro-economic variables, Section IA.7.4 provides a measure of SSI based on dollar share change rather than percent share change, Section IA.7.5 provides a measure of SSI that includes new leveraged ETFs as they come to market, Section IA.7.6 provides a measure of SSI that is formed from only the three leveraged-long ETFs and also a measure of SSI that is formed from only the three leveraged-short ETFs, and Section IA.7.7 considers each long-short index pair separately (e.g., SDS and SSO). The results using these alternative specifications of SSI are qualitatively the same in economic magnitude and statistical significance to those presented in this section. In addition to alternative specifications of SSI, I also consider an additional test in the Supplementary Material: Section IA.7.8 shows that institutional ownership changes in leveraged ETFs *positively* predict aggregate returns (i.e., while institutions rarely trade leveraged ETFs, when they do, their trades appear informed).

D. Out-of-Sample R^2 Analysis

To this point, the return predictability results are all from in-sample tests. As such, it is important to also examine the out-of-sample predictive performance of SSI. As highlighted in Welch and Goyal (2008), out-of-sample analysis is both

a critical diagnostic for the in-sample results and also interesting in itself for investors who seek to use SSI for market-timing purposes. With this motivation, I analyze the out-of-sample R^2 s (R_{OS}^2 s) for SSI in the spirit of Campbell and Thompson (2008) and Welch and Goyal (2008). R_{OS}^2 measures whether or not a variable is a better predictor of returns than the historical average return. If the value is positive, the economic interpretation is that the predictor outperforms the historical average. $R_{OS,\tau}^2$ is computed as

$$(8) \quad R_{OS,\tau}^2 = 1 - \frac{\sum_{t=\tau}^T (r_t - \hat{r}_t)^2}{\sum_{t=\tau}^T (r_t - \bar{r}_t)^2},$$

in which \hat{r}_t is the fitted value of the monthly CRSP equal-weighted index return, the CRSP value-weighted index return, or the S&P 500 index return using coefficients from out-of-sample univariate predictive regressions, \bar{r}_t is the historical average monthly return, r_t is the realized monthly index return, and τ is the start date.

Table 8 reports $R_{OS,\tau}^2$ results for the three benchmark indices' returns using eight different start dates: Jan. 2010, Jan. 2011, Jan. 2012, Jan. 2013, Jan. 2014, Jan. 2015, Jan. 2016, and Jan. 2017. For each start date and each benchmark index, $R_{OS,\tau}^2$ is calculated through Dec. 2019. All estimated coefficients for the fitted values of index returns use return time series that begins in Nov. 2006. The historical average return for the benchmark returns is computed using a time series that begins in Feb. 1926. As in Campbell and Thompson (2008), calculations of $R_{OS,\tau}^2$ use estimated coefficients from regressions through month $t - 1$ to obtain the fitted index return and the historical average return is also computed through month $t - 1$. The column labeled "EW" reports the results for the CRSP equal-weighted index, the column labeled "VW" reports the results for the CRSP value-weighted index, and the column labeled "SP500" reports the results for the S&P 500 index. To assess statistical significance for each value of $R_{OS,\tau}^2$, I utilize the parametric bootstrap outlined in Section IA.4 of the Supplementary Material, using 10,000 simulated histories, and compute a one-tailed p -value.

Table 8 reports positive $R_{OS,\tau}^2$ s for each start date with the CRSP equal-weighted index and all but two are statistically greater than zero at a 10% p -value

TABLE 8
Out-of-Sample Return Predictability

Table 8 reports the out-of-sample $R_{OS,\tau}^2$ for SSI based on the calculation from Campbell and Thompson (2008). Fitted values of returns in month t are obtained by using coefficients estimated on the Nov. 2006 through month $t - 1$ sample. The long-term average returns are obtained by taking the average monthly return from Feb. 1926 through month $t - 1$. Statistical significance is reported as a one-tailed p -value using a parametric bootstrap. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Horizon	Out-of-Sample R^2		
	EW	VW	SP500
01/10–12/19	0.05**	0.01*	0.00
01/11–12/19	0.06**	0.02**	0.01*
01/12–12/19	0.01	-0.02	-0.02
01/13–12/19	0.05*	-0.00	-0.01
01/14–12/19	0.11***	0.05**	0.04**
01/15–12/19	0.10**	0.05**	0.05**
01/16–12/19	0.07*	0.03*	0.03*
01/17–12/19	0.04	0.00	0.01

threshold or better. The values for the CRSP value-weighted index are similar but slightly weaker: nearly all values are positive valued and the majority are statistically greater than zero. The results for the S&P 500 index are the weakest, but still report positive valued $R_{OS,t}^2$ s across all but two of the start dates and half of the values are statistically greater than zero. As a point of comparison, the in-sample predictive regressions in Table 3 report R^2 s for the post-2009 sample of 0.04, 0.03, and 0.02 for the CRSP equal-weighted index, the CRSP value-weighted index, and the S&P 500. Thus, the results in Table 8, which are calculated over similar dates, are not materially different.

Collectively, the results in Table 8 complement earlier results by showing that the return predictability of SSI holds for both in-sample and out-of-sample tests. Moreover, the out-of-sample tests also provide insights for investors that condition on SSI for market-timing purposes. Specifically, Campbell and Thompson (2008) show that an investor with mean–variance utility may increase their expected return proportional to

$$(9) \quad \left(\frac{R_{OS}^2}{1 - R_{OS}^2} \right) \left(\frac{1 + S^2}{S^2} \right),$$

in which S^2 is the squared Sharpe Ratio of the investor's portfolio absent the conditioning information. In other words, positive values of $R_{OS,t}^2$ may be interpreted as SSI providing information that improves the investor's Sharpe Ratio. In the Supplementary Material, Section IA.8 studies in greater detail the performance of portfolios that condition on realized values of SSI. In that analysis, I consider two basic managed portfolios: one is a long-only portfolio that only purchases a market index when SSI is negative and one is a long-short portfolio that buys the market index when SSI is negative and shorts the market index when SSI is positive.³¹ Both the long-only portfolios and the long-short portfolios generate statistically significant and economically meaningful alpha.

VI. Conclusion

This article provides a direct and clean measure of investor sentiment using a novel market setting: the leveraged ETFs primary market. Leveraged ETFs are special because i) a distinct investor clientele trades the ETF shares (“dumb” money, short-horizon traders) and another distinct investor clientele trades the shares' underlying assets (relatively smarter institutions) and ii) mispricing between the ETF shares and the underlying assets is corrected via *observable* arbitrage trades. Thus, observed arbitrage trades proxy for market-wide latent demand shocks that gave rise to the initial mispricing. In other words, these arbitrage trades signal aggregate disagreement between “dumb” and “smart” money. With exception to a few papers, the leveraged ETF primary market has gone largely unnoticed despite its incredibly rich information on investor sentiment.

The paper's sentiment measure (SSI) proxies for nonfundamental demand and it is a powerful predictor of market returns. Furthermore, traders may improve their

³¹For the managed portfolio analysis, I do not standardize SSI to prevent any lookahead bias.

portfolios' risk–return characteristics by conditioning on realized values of SSI. Given the strong empirical evidence and economic motivation, it is likely that the SSI will serve as an important sentiment proxy in future asset pricing and corporate finance studies.

Supplementary Material

Supplementary Material for this article is available at <https://doi.org/10.1017/S0022109022000291>.

References

- Antoniewicz, R. S., and J. Heinrichs. "Understanding Exchange-Traded Funds: How ETFs Work." Working Paper, available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2523540 (2014).
- Baker, M.; M. F. Hoeyer; and J. Wurgler. "The Risk Anomaly Tradeoff of Leverage." NBER Working Paper No. 22116 (2016).
- Baker, M., and J. Wurgler. "Investor Sentiment and the Cross-Section of Stock Returns." *Journal of Finance*, 61 (2006), 1645–1680.
- Baker, M., and J. Wurgler. "Investor Sentiment in the Stock Market." *Journal of Economic Perspectives*, 21 (2007), 129–151.
- Bakshi, G., and D. Madan. "A Theory of Volatility Spreads." *Management Science*, 52 (2006), 1945–1956.
- Bali, T. G.; S. J. Brown; S. Murray; and Y. Tang. "A Lottery-Demand-Based Explanation of the Beta Anomaly." *Journal of Financial and Quantitative Analysis*, 52 (2017), 2369–2397.
- Bali, T. G.; N. Cakici; and R. F. Whitelaw. "Maxing Out: Stocks as Lotteries and the Cross-Section of Expected Returns." *Journal of Financial Economics*, 99 (2011), 427–446.
- Barberis, N., and R. Thaler. "A Survey of Behavioral Finance." In *Handbook of the Economics of Finance*, Vol. 1. Cambridge, MA: NBER (2003), 1053–1128.
- Ben-David, I.; F. Franzoni; and R. Moussawi. "Exchange-Traded Funds." *Annual Review of Financial Economics*, 9 (2017), 169–189.
- Ben-David, I.; F. Franzoni; and R. Moussawi. "Do ETFs Increase Volatility?" *Journal of Finance*, 73 (2018), 2471–2535.
- Ben-Rephael, A.; S. Kandel; and A. Wohl. "Measuring Investor Sentiment with Mutual Fund Flows." *Journal of Financial Economics*, 104 (2012), 363–382.
- Bessembinder, H. "Predictable ETF Order Flows and Market Quality." *Journal of Trading*, 10 (2015), 17–23.
- Black, F.; M. C. Jensen; and M. Scholes. "The Capital Asset Pricing Model: SOME EMPIRICAL TESTS." *Studies in the Theory of Capital Markets*, 81 (1972), 79–121.
- Bollerslev, T.; G. Tauchen; and H. Zhou. "Expected Stock Returns and Variance Risk Premia." *Review of Financial Studies*, 22 (2009), 4463–4492.
- Brown, D. C.; S. W. Davies; and M. Ringgenberg. "ETF Arbitrage, Non-fundamental Demand, and Return Predictability." *Review of Finance*, 25 (2021), 937–972.
- Campbell, J. Y., and S. B. Thompson. "Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?" *Review of Financial Studies*, 21 (2008), 1509–1531.
- Cheng, M., and A. Madhavan. "The Dynamics of Leveraged and Inverse-Exchange Traded Funds." *Journal of Investment Management*, 7 (2009), 43–62.
- Coval, J., and E. Stafford. "Asset Fire Sales (and Purchases) in Equity Markets." *Journal of Financial Economics*, 86 (2007), 479–512.
- Da, Z.; J. Engelberg; and P. Gao. "The Sum of All Fears Investor Sentiment and Asset Prices." *Review of Financial Studies*, 28 (2014), 1–32.
- DellaVigna, S., and J. M. Pollet. "Demographics and Industry Returns." *American Economic Review*, 97 (2007), 1667–1702.
- DellaVigna, S., and J. M. Pollet. "Investor Inattention and Friday Earnings Announcements." *Journal of Finance*, 64 (2009), 709–749.
- DeVault, L.; R. Sias; and L. Starks. "Sentiment Metrics and Investor Demand." *Journal of Finance*, 74 (2019), 985–1024.
- Edmans, A.; D. Garcia; and Ø. Norli. "Sports Sentiment and Stock Returns." *Journal of Finance*, 62 (2007), 1967–1998.

- Egan, M. L.; A. MacKay; and H. Yang. "Recovering Investor Expectations from Demand for Index Funds." NBER Working Paper No. 26608 (2020).
- Engelberg, J.; D. R. Mclean; J. Pontiff; and M. Ringgenberg. "Do Cross-Sectional Predictors Contain Systematic Information?" *Journal of Financial and Quantitative Analysis*, forthcoming (2022).
- Fama, E. F., and K. R. French. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics*, 33 (1993), 3–56.
- Frazzini, A., and L. H. Pedersen. "Embedded Leverage." *Review of Asset Pricing Studies*, 12 (2022), 1–52.
- Frazzini, A., and L. H. Pedersen. "Betting Against Beta." *Journal of Financial Economics*, 111 (2014), 1–25.
- Han, B., and A. Kumar. "Speculative Retail Trading and Asset Prices." *Journal of Financial and Quantitative Analysis*, 48 (2013), 377–404.
- He, Z.; B. Kelly; and A. Manela. "Intermediary Asset Pricing: New Evidence from Many Asset Classes." *Journal of Financial Economics*, 126 (2017), 1–35.
- Hirshleifer, D. "Investor Psychology and Asset Pricing." *Journal of Finance*, 56 (2001), 1533–1597.
- Hirshleifer, D.; S. S. Lim; and S. H. Teoh. "Driven to Distraction: Extraneous Events and Underreaction to Earnings News." *Journal of Finance*, 64 (2009), 2289–2325.
- Hodrick, R. J. "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement." *Review of Financial Studies*, 5 (1992), 357–386.
- Hong, H.; W. Torous; and R. Valkanov. "Do Industries Lead Stock Markets?" *Journal of Financial Economics*, 83 (2007), 367–396.
- Huang, D.; F. Jiang; J. Tu; and G. Zhou. "Investor Sentiment Aligned: A Powerful Predictor of Stock Returns." *Review of Financial Studies*, 28 (2015), 791–837.
- Huberman, G., and T. Regev. "Contagious Speculation and A Cure for Cancer: A Nonevent that Made Stock Prices Soar." *Journal of Finance*, 56 (2001), 387–396.
- Ivanov, I. T., and S. L. Lenkey. "Are Concerns About Leveraged ETFs Overblown?" Federal Reserve Board: Finance and Economics Discussion Series (2014).
- Jiang, F.; J. Lee; X. Martin; and G. Zhou. "Manager Sentiment and Stock Returns." *Journal of Financial Economics*, 132 (2018), 126–149.
- Jiang, W.; Y. Xiao; and H. Yan. "Financial Innovation and Investor Behavior: Evidence from the ETF Market." Working Paper (2020).
- Kumar, A. "Who Gambles in the Stock Market?" *Journal of Finance*, 64 (2009), 1889–1933.
- Lee, C.; A. Shleifer; and R. H. Thaler. "Investor Sentiment and the Closed-End Fund Puzzle." *Journal of Finance*, 46 (1991), 75–109.
- Lou, D. "A Flow-Based Explanation for Return Predictability." *Review of Financial Studies*, 25 (2012), 3457–3489.
- Madhavan, A. *Exchange-Traded Funds and the New Dynamics of Investing*. Oxford: Oxford University Press (2016).
- Neal, R., and S. M. Wheatley. "Do Measures of Investor Sentiment Predict Returns?" *Journal of Financial and Quantitative Analysis*, 33 (1998), 523–547.
- Pástor, L., and R. F. Stambaugh. "Liquidity Risk and Expected Stock Returns." *Journal of Political Economy*, 111 (2003), 642–685.
- Rapach, D. E.; M. C. Ringgenberg; and G. Zhou. "Short Interest and Aggregate Stock Returns." *Journal of Financial Economics*, 121 (2016), 46–65.
- Rosenberg, J. V., and R. F. Engle. "Empirical Pricing Kernels." *Journal of Financial Economics*, 64 (2002), 341–372.
- Saunders, E. M. "Stock Prices and Wall Street Weather." *American Economic Review*, 83 (1993), 1337–1345.
- Shleifer, A. "Do Demand Curves Discuss Slope Down?" *Journal of Finance*, 41 (1986), 579–590.
- Shum, P.; W. Hejazi; E. Haryanto; and A. Rodier. "Intraday Share Price Volatility and Leveraged ETF Rebalancing." *Review of Finance*, 20 (2015), 2379–2409.
- Staer, A. "Fund Flows and Underlying Returns: The Case of ETFs." *International Journal of Business*, 22 (2017), 30.
- Stambaugh, R. F. "Predictive Regressions." *Journal of Financial Economics*, 54 (1999), 375–421.
- Tuzun, T. "Are Leveraged and Inverse ETFs the New Portfolio Insurers?" Federal Reserve Working Paper, available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2340616 (2013).
- Welch, I., and A. Goyal. "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction." *Review of Financial Studies*, 21 (2008), 1455–1508.
- Zweig, M. E. "An Investor Expectations Stock Price Predictive Model Using Closed-End Fund Premiums." *Journal of Finance*, 28 (1973), 67–78.