Cross-Market Investor Sentiment in Commodity Exchange-Traded Funds

Hsiu-Lang Chen*
College of Business Administration
University of Illinois at Chicago

August 5, 2013

* I thank Chun Zhang, Jiayin Hou, and Saembyeol Park for research assistance. For their helpful comments, I also thank the participants of the Annual Meetings of Midwest Finance Association held in Chicago, March 14-16, 2013. Address correspondence to Hsiu-lang Chen, Department of Finance, 601 South Morgan Street, Chicago, IL 60607, or e-mail: hsiulang@uic.edu.

Cross-Market Investor Sentiment in Commodity Exchange-Traded Funds

Abstract

This study shows how the investor sentiment in the stock market affects prices of commodity exchange-traded funds (ETFs). We provide quantitative evidence that the tracking errors of commodity ETFs differ in the bull versus the bear stock market, and the aggregate tracking error of commodity ETFs is sensitive to the sentiment measure constructed by Baker and Wurgler (2007). We exploit a profitable trading strategy based on investor sentiment in the stock market and commodity market. We use commodity ETFs and Spider, an ETF tracking S&P 500 Index, in a long-short investment strategy according to prior sentiment signals. The sentiment-driven demand for commodity ETFs exists, and it is a short-term phenomenon. This unique evidence indicates investor sentiment affects asset valuation across markets.

I. Introduction

It is well known that liquid financial markets are not always as orderly as the efficient market advocates might suggest (see Grossman and Stiglitz, 1980). When some investors trade on a ‘‘noisy’’ signal that is unrelated to fundamentals, asset prices likely deviate from their intrinsic value. In a model of two types of investors, rational arbitrageurs who are sentiment-free and irrational traders prone to exogenous sentiment, DeLong, Shleifer, Summers, and Waldmann (1990) argue that rational arbitrageurs mainly face limits from short time horizons or from costs and risks of trading and short selling. Shleifer and Vishny (1997) show how agency problems between an arbitrageur and his or her source of capital can also hinder arbitrage. As a result, sentiment-based demands might drive prices away from their fundamental values. Empirically, examination of this issue is still contentious. The absence of precise valuation models for stocks makes measuring deviations from theoretical prices difficult. Similar problems arise from the difficulty in measuring investor sentiment. Our study of passively managed commodity exchange-traded funds (ETFs) might be able to subside these issues.[1]

Most public statements by institutional investors emphasize the primary advantage of commodity investments as diversification, providing a return that has little correlation with core equity and bond holdings. Since their introduction in 2004, commodity ETFs have grown from just over $1 billion to $109 billion by the end of 2011, with total net assets almost tripling in the last two years. Strong net issuance and surging gold and silver prices were the primary drivers behind the increase in assets during this time. By construction, a passively managed commodity

[1] An ETF is an investment company, typically an open-end investment company (open-end fund), whose shares are traded intraday on stock exchanges at market-determined prices. Investors may buy or sell ETF shares through a broker just as they would the shares of any publicly traded company. The first ETF—a broad-based domestic equity fund tracking the S&P 500 index—was introduced in 1993. Until 2008, the U.S. Securities and Exchange Commission’s (SEC) exemptive relief was granted only to ETFs that tracked designated indexes. According to the 2012 Investment Company Fact Book, by the end of 2011, the total number of index-based and actively managed ETFs had grown to 1,134, and total net assets were $1.05 trillion.

ETF tracks its underlying index, and thus the tracking errors of the ETF, gross returns on the ETF minus returns on the index, are free of any complex pattern of compensation for systematic risk. Unlike the discounts on closed-end mutual funds, the market price of an ETF is close to the value of its underlying index assets net of the expense ratio, because its portfolio composition is transparent, and authorized participants (institutional investors) are allowed to assemble a basket of underlying index assets in exchange for shares of the ETF. [2] Any correlation between the commodity and the stock market due to economic links will show in both commodity ETFs prices and their designated index value. Thus the *tracking errors* would be immunized from any interaction of economic fundamentals between the equity market and the commodity market. Even though the tracking errors of an ETF may not be zero all the time for reasons of imperfect replication of the underlying index, some unknown microstructure issues, or a mismatch in timing between receiving fund flows and acquiring assets, the average of tracking errors should be independent of a hype of a market other than the market in which the ETF assets underlie.

However, commodity ETFs are traded in a regular stock exchange. The aggregate sentiment in the stock market is likely to influence prices of commodity ETFs.[3] We label it as the cross-market sentiment effect. A pure psychology experiment conducted by Moreland and Beach (1992) supports this conjecture. With a control of no interaction, they show that students' mere exposure to the same classroom has strong effects on attraction and similarity to others. People also tend to conform to the judgments and behaviors of others. In a sequential decision model, Banerjee (1992) shows that people will do what others are doing rather than use their

[2] For example, United States Oil Fund, LP (USO), a commodity ETF, declared on its prospectus dated on April 10, 2006, that the price of USO's units on the American Stock Exchange would closely track the spot price of a barrel of WTI light, sweet crude oil, less USO's expenses. USO sought to achieve its investment objective by investing in a mix of oil futures contracts and other oil interests such as options on oil futures contracts, forward contracts for oil, and over-the-counter transactions based on the price of oil and other petroleum-based fuels.

[3] Sentiment in this paper simply represents the expectations of market participants relative to a norm—a bullish (bearish) investor expects returns to be above (below) average, whatever the average is.

own information. Investors investing in commodity ETFs likely enjoy the exciting celebration along with stock investors for a bullish stock market.

Several studies have documented the interaction between the sentiment and the broad stock market returns. For example, Brown and Cliff (2004) show that sentiment levels and changes are strongly correlated with contemporaneous market returns, but sentiment has little predictive power for near-term future stock returns. Using a direct survey measure of investor sentiment, Brown and Cliff (2005) provide evidence that optimism is associated with overvaluation and low returns over the subsequent one to three years as the valuation level returns to its intrinsic value. Ben-Rephael, Kandel, and Wohl (2012) also find that investor sentiment, proxied by net exchanges of equity funds, creates noise in aggregate market prices. Baker and Wurgler (2006 and 2007) argue that investor sentiment, the propensity to speculate, drives the relative demand for speculative investments, which are typically hard to value and tend to be difficult to arbitrage and therefore possibly cause cross-sectional effects on stock returns.[4] They document that speculative and hard-to-arbitrage stocks have lower (higher) future returns on average than bond-like stocks when sentiment is measured to be high (low). In this study, we investigate whether investor sentiment in the stock market affects daily tracking errors of commodity ETFs. Ultimately, we explore whether we can quantify the cross-market sentiment effects by exploiting profits from a long-short investment strategy. The findings of this study complement the existing literature that asserts investor sentiment affects the broad market returns and the cross-section of stock returns.

This study also contributes to the ETFs literature. Elton, Gruber, Comer, and Li (2002) identify that two causes of the underperformance of Standard and Poor's Depository Receipts

[4] D'Avolio (2002) documents that stocks that are young, small, unprofitable, or experiencing extreme growth tend to be more costly to buy and to sell short. Wurgler and Zhuravskaya (2002) also find such stocks have a high degree of idiosyncratic variation in their returns, which makes betting on them riskier.

(SPDR, commonly referred as Spider) relative to the S&P 500 Index, which Spider tracks, are mainly due to the management fee of 18.45 basis points (bps) and the loss of return from dividend reinvestment of 9.95 bps. After correcting some of the measurement errors in net asset value, Engle and Sarkar (2006) show the average premium of exchange-traded equity index funds was less than 5 bps and the standard deviation was less than 20 bps. Decloure and Zhong (2007) and Levy and Lieberman (2013) analyze the “stale pricing” problem of securities traded in foreign country markets in generating a premium of country ETFs. Levy and Lieberman (2013) further find that whereas country ETF prices are mostly driven by their NAV returns during synchronized trading hours, the S&P 500 index has a dominant effect during non-synchronized trading hours. The “stale pricing” problem is not the major issue for commodity ETFs, except for precious-metal ETFs, which typically track the spot prices of precious metals in London. Because this study aims to determine whether investor sentiment in one market affects asset prices in another market, incorporating all possible channels through which investor sentiment might have influence is important. If the replication of an index a commodity ETF tracks is imperfect, investor sentiment could amplify the tracking errors. Therefore, we confine our study to the ETF’s price relative to its underlying index, not relative to its NAV. In addition, we use the underlying index to gauge the prospect of the commodity market relative to that of the stock market and to construct the index-adjusted performance measure for an investment strategy exploiting the sentiment effect. We study whether the average of tracking errors of commodity ETFs differs depending on investor sentiment in the stock market. The result of this study suggests that in addition to the stale pricing problems, behavioral factors may account for some mispricing in ETFs.

The rest of the paper proceeds as follows. Section II describes the data. Section III presents statistics on tracking errors and tracking-error volatility for commodity ETFs. Section IV tests whether investor sentiment in the stock market affects prices of commodity ETFs. Section V exploits a profitable trading strategy based on investor sentiment in the stock market and commodity market. Section VI further performs a Fama-French risk-factor model for a robustness check, and section VII concludes.

II. Data

The CRSP daily return files and SEC's EDGAR database constitute our main data sources. We retrieve returns of all commodity ETFs as well as Spider (the ticker symbol: SPY) from CRSP.[5] We hand-collect all historical expense ratios for all commodity ETFs, and Spider from SEC's EDGAR database. We construct the daily gross returns by adding the expense ratios to the reported net-of-expense returns for these ETFs. We only consider ETFs that have at least one-year daily returns, and have specified in the prospectus the underlying indexes that ETFs will track.[6] We retrieve the data of indexes tracked by ETFs from the web sites of the companies that publish the indexes, as well as the data service providers. Our sample includes 33 commodity ETFs from November 18, 2004, to December 31, 2011.

III. Tracking Errors and Tracking-Error Volatility

We first present the quartile distribution of tracking errors of commodity ETFs over the entire sample period. An ETF's tracking error is the ETF's gross return minus its benchmark index return. Our sample contains nine commodity ETFs designed to provide double returns, inverse

[5] An ETF is a security with a common share code of 73 in CRSP.
[6] We exclude four ETFs (ticker symbols: BNO, UGA, WITE, and GLTR). We cannot have the complete data for the indexes the first two ETFs track. The last two ETFs track an index comprising a customized deposit of bullion metals.

returns, or double inverse returns of the indexes they track. For these leveraged/inverse ETFs, we define their tracking errors accordingly.[7] Panel A of Table 1 shows that commodity ETFs entail non-trivial tracking errors, whereas Spider tracks the S&P 500 Index nearly perfectly. The median commodity ETF trails its index by 0.7 basis points. To present the stability of an ETF in tracking its index, we calculate the volatility of the ETF's daily tracking errors over the sample period. Panel B of Table 1 shows the median tracking-error volatility (TEV) among these 33 commodity ETFs is about 2.05%. For a reference comparison, we also construct the quartile distribution of TEV for Spider in such a way that the volatility of Spider's tracking errors is calculated each time for a period in which TEV of a commodity ETF is calculated. Spider only entails a median tracking-error volatility of 0.16%. The standard deviation of the cross-sectional TEVs of commodity ETFs is about 1.73%. The big deviation might indicate the large variety of ways in which commodity ETFs implement tracking strategies—some investing in underlying assets directly whereas some using financial instruments to gain exposure to the underlying assets. One must take the large variation in tracking implementation into consideration when we empirically test whether the tracking errors of commodity ETFs can possibly indicate investor sentiment in the stock market.

Table 2 shows the quartile distribution of tracking errors over contrast periods, bull versus bear. We define a bull stock market versus a bear stock market according to daily returns of RMRF, one of the Fama-French three factors. The bull (bear) stock markets include days that RMRF is positive (negative). Spider trails the S&P 500 Index in a bull stock market, whereas

[7] For example, the tracking error of ProShares Ultra Gold ETF (UGL) is its gross returns minus double returns on the daily performance of gold bullion as measured by the U.S. Dollar p.m. fixing price for delivery in London, the benchmark index UGL tracks. Similarly, the tracking error of ProShares UltraShort Gold ETF (GLL) is its gross returns minus inverse double returns on the benchmark index. According the description on page 21 of the initial prospectus on November 21, 2008, these two funds will not invest in bullion, but rather will use financial instruments to gain exposure to these precious metals. Not investing directly in bullion may introduce additional tracking errors.

overshoots the S&P 500 Index in a bear market in an almost identical magnitude of about 2 bps. However, commodity ETFs exhibit an opposite effect. The median commodity ETF overshoots its index by 8.5 bps in a bull stock market but trails its index by 11.3 bps in a bear stock market. Investor sentiment in the stock market might affect traded securities regardless of whether the underlying assets of the securities are traded in the same market. In Panel B, we report the volatility of the ETF's daily tracking errors over these two contrast periods. Commodity ETFs seem to have a slightly higher TEV in the bull stock market, whereas Spider has a marginally higher TEV in the bear market.

Brown and Cliff (2004) document that sentiment levels and changes are strongly correlated with contemporaneous market returns. Investors are likely more hyper in a bullish stock market. These investors in turn are more likely excited about commodity ETFs that are also traded in the same market. The evidence shown in Table 2 that the daily tracking errors for commodity ETFs tend to be positive (negative) when the stock market is bull (bear) supports this conjecture. A similar behavior has been documented in a psychology experiment conducted by Moreland and Beach (1992) that shows students' mere exposure to the same classroom has strong effects on attraction and similarity to others. Banerjee (1992) also shows that people will do what others do rather than use their own information.

IV. Do commodity ETFs' tracking errors reflect investor sentiment in the stock market?

The literature has well documented that investor sentiment affects the broad market and the cross-section of stock returns. However, whether aggregate investor sentiment in one market affects asset prices in another market is still unknown. In other words, can investor sentiment in the stock market affect all securities traded in the market, including securities whose underlying assets are commodities? We simply view investor sentiment as optimism or pessimism about stocks in general and thus assume investor sentiment is positive (negative) in a bull (bear) stock market. We define bull and bear stock market according to daily returns of RMRF. Similarly, a day for an ETF is classified as a bull (bear) market in a commodity market if the daily return on an index tracked by the ETF minus daily one-month T-bill rate is positive (negative). For each ETF since its inception, we classify trading days into four periods: (1) bull stock market and bull commodity market (BullSBullC), (2) bull stock market and bear commodity market (BullSBearC), (3) bear stock market and bull commodity market (BearSBullC), and (4) bear stock market and bear commodity market (BearSBearC). We examine whether investor sentiment in the stock market affects the commodity ETFs' tracking errors after controlling for sentiment in the commodity.

Because of the concern that commodity ETFs engage in quite different index tracking strategies in addition to commodity types varying from oil to bullion, we conduct a test on the basis of individual commodity ETFs in Table 3. We first test for whether the mean of daily tracking errors (TEs) of an ETF is identical in two contrast periods defined by the stock market, BullSBullC versus BearSBullC as well as BullSBearC versus BearSBearC. For a reference comparison, we also test for whether the mean of daily TEs of Spider is the same in two contrast periods defined by the underlying index of each commodity ETF, BullSBullC versus BullSBearC as well as BearSBullC versus BearSBearC. Panel A shows that the tracking errors of 29 out of 33 commodity ETFs differ significantly in the bull stock market versus the bear stock market, after controlling for bullish sentiment in the commodity market. Similarly, Panel B shows that investor sentiment in the stock market affect the tracking errors of 30 commodity ETFs after controlling for bearish sentiment in the commodity market. After controlling for investor sentiment in the stock market, by contrast, Spider shows little evidence that sentiment in the commodity market significantly affects its tracking errors.

If the tracking errors of commodity ETFs really reflect investor sentiment in the stock market in which the ETFs are traded, the aggregate tracking errors are anticipated to be more sensitive to the sentiment, in the sense that the aggregate tracking errors of commodity ETFs will have higher sentiment betas after controlling for other factors that might affect the tracking errors. The other factors we consider are liquidity and the stock market condition. We use two liquidity measures, Amihud illiquidity and turnover, for this study. Amihud (2002) illiquidity is the absolute return divided by the dollar volume.[8] It captures the daily price response associated with one dollar of trading volume. The turnover is the ratio of trading volume to the number of shares outstanding. Amihud and Mendelson (1986) document that turnover is negatively related to illiquidity costs. Investor sentiment measure (SENT) is constructed by Baker and Wurgler (2007) and directly retrieved from Wurgler's website. Given that the sentiment measure is only available annually and monthly up to December 2010, daily tracking errors and liquidity measures are converted to monthly data first. Instead of retrieving monthly returns and liquidity variables directly, we average the daily data for each month for each ETF in the hopes that the averaging can reduce noises in the tracking-error calculation so the aggregate tracking errors of commodity ETFs can represent investor sentiment in the stock market. We calculate the cross-sectional average of monthly data in an equal weight for the portfolio of commodity ETFs. ETFs that use financial instruments to gain leverage exposure to their underlying assets may introduce additional tracking errors that are not necessarily related to investor sentiment. As a result, we

[8] The daily dollar volume is the trade volume times the daily closing price. Daily closing price, trade volumes, and number of shares outstanding of commodity ETFs are retrieved from CRSP.

exclude leveraged and inverse ETFs from the analysis. We have 24 non-leveraged/non-inverse commodity ETFs in total and 74 months between November 2004 and December 2010.

We regress monthly tracking errors of the commodity ETF portfolio on the liquidity measures and sentiment measures. To control the conditions of the general economy and stock market, we add RMRF to the independent variables. Table 4 shows the aggregate tracking errors of commodity ETFs load positively and significantly on the sentiment regardless of whether we consider liquidity. Because the tracking error has adjusted for any fundamental influence associated with the economy by deducting returns on an index tracked by an ETF, we expect the aggregate tracking errors no longer strongly co-move with the stock market excess returns. When negative liquidity shocks hit, sentiment-driven demand, perhaps due to limits to arbitrage, will amplify the aggregate tracking errors as shown in the coefficients of two liquidity measures. As a reference, we apply the same test to the tracking errors of Spider. As expected, the tracking errors of Spider are not sensitive to investor sentiment at all because sentiment, if any is present, affects both Spider and S&P 500 company stocks simultaneously and similarly. Sentiment-driven mispricing might also wane much sooner on Spider relative to commodity ETFs, or arbitrage forces to correct the mispricing on Spider might be more effective.

If the tracking errors of commodity ETFs indeed reflect investor sentiment in the stock market, the daily tracking errors of commodity ETFs can serve as an indicator of daily sentiment for the stock market. Thus, the daily tracking errors of commodity ETFs complement indicators of annual and monthly sentiment measures provided by Baker and Wurgler (2006 and 2007).

V. Using Sentiment to Predict Returns

So far we have shown investor sentiment in the stock market will likely swing the ETFs' tracking errors given that commodity ETFs are traded in the stock market. In other words, the positive sentiment in the stock market drives the prices of commodity ETFs above the intrinsic values of their underlying commodities. We follow the suggestion by Baker and Wurgler (2007) that the strongest tests of the effects of sentiment involve return predictability.

To possibly quantify the impact of sentiment in the stock market on commodity ETFs, we perform a long-short investment strategy involving a commodity ETF and Spider, depending on whether investor sentiment is simply positive or negative. A zero-cost investment in an efficient market should generate zero return after adjusting for risks. We investigate whether current investor-sentiment levels predict future returns on the long-short strategy as sentiment wanes differently on commodity ETFs and Spider, or as arbitrage forces accumulate to correct mispricing of these two securities at different paces. Additionally, using a commodity ETF and Spider, instead of a commodity ETF and its underlying assets, in the long-short investment strategy offers a practical advantage. Unlike in stock ETFs, an authorized participant in the creation (redemption) of a commodity ETF may have to deliver (receive) a combination of cash and physical assets underlying the commodity ETF (e.g., SPDR Gold ETF [GLD]), or a combination of cash and Treasuries (e.g., United States Oil ETF [USO]). Types of assets for delivery in exchange for commodity ETF shares vary accordingly and may be illiquid as well as not available for short.

In the single-market strategy, we explore an investment opportunity based on investor sentiment in the stock market. At the beginning of each day, we long an ETF and short SPY if the stock market was bull the prior day, whereas we short an ETF and long SPY if the stock market was bear the prior day. In the cross-market strategy, we explore an investment opportunity based on investor sentiment in both stock and commodity markets. At the beginning of each day, we long an ETF and short SPY if the stock market was bull and the ETF market was bear the prior day (BullSBearC). We short an ETF and long SPY if the stock market was bear and the ETF market was bull the prior day (BearSBullC). Using ETFs' reported net-of-expense returns, we calculate daily performance for each strategy for each ETF.

Table 5 shows that the proposed investment strategy is profitable. For example, the single-market strategy on the basis of individual ETFs generates 14.1 bps per day on average, whereas the cross-market strategy results in about 18.3 bps. Both are significant at the level of 1%. As a reference, we perform a plain strategy of long ETF and short SPY constantly. Without relying on sentiment signals, the plain strategy results in zero performance. This result indicates that both commodity ETFs and Spider are exposed to systematic risk factors similarly during the sample period, and thus the profit from the long-short strategy is not just compensation for bearing the systematic risk. When we pool all performance of strategies across commodity ETFs, the average performance of sentiment strategies is significant and positive.[9] The strong evidence also appears in a singular strategy under either single-market or cross-market sentiment consideration. For example, a sentiment strategy based on positive sentiment in the stock market generates 15.2 bps per day on average, whereas a strategy based on positive sentiment in the stock market and negative sentiment in the commodity market results in 22.3 bps. Both are significant at the 1% level.

To test the sentiment effects further, we split the time series into extreme bull (bear) days depending on whether the daily market excess returns are ranked at the top (bottom) third over the entire bull (bear) market period. If the sentiment effects indeed exist, we anticipate that investors are more hyper in the extreme bull market while more pessimistic in the extreme bear

[9] We checked the extreme profit/loss for a potential data error but have not yet found one. On October 10, 2008, DBS had returns of -19.98% (its underlying index DBSLIX had return of 0.87%), while SPY had returns of -2.42%. As a result, the lowest performance of the strategy is -17.56% (-19.98-(-2.42)). On September 17, 2008, DBS had returns of 16.75% (its underlying index DBSLIX had a return of -5.56%) while SPY had returns of -4.50%. As a result, the highest performance of the strategy is 21.25% (16.75-(-4.50)).

market. As a result, a sentiment strategy based on these extreme signals is expected to generate higher returns, and Panel B of Table 5 confirms this expectation. For example, a sentiment strategy based on extreme positive sentiment in the stock market generates 23.3 bps per day on average, whereas a strategy based on extreme positive sentiment in the stock market and extreme negative sentiment in the commodity market results in 44.9 bps. Again, both are significant at the 1% level.

Although our proposed long-short strategy involves two liquid securities, a commodity ETF and Spider, performance of the strategy might still be exposed to unknown economic risk factors related to the fundamental difference between the commodity market and the stock market. To take this possibility into consideration, we regress performance of long-short sentiment strategies on the return difference between the S&P 500 Index and the commodity index which the ETF tracks. According to the position of an ETF's trading signal in the timeline, we further classify each strategy into three mutually exclusive groups depending on whether a sentiment signal is fresh new, in the middle of a consecutive signal sequence, or at the tail of a consecutive signal sequence. We pool daily performance of each strategy across all ETFs in each group and perform a regression in Table 6. The strategy following the fresh positive sentiment in the stock market generates positive raw returns of 43 bps, and we can have such a strategy for 1,427 fund-days over the sample period. Although the fundamental difference between the stock and commodity markets explains about half of the returns, the strategy still delivers a significant alpha of 23.7 bps.

Passively managed ETFs are neither difficult to value nor hard to arbitrage. As sentiment wanes (perhaps spurred by fundamental news or an absence thereof) or as arbitrage forces eventually accumulate to correct mispricing, we anticipate that the sentiment-driven demand for commodity ETFs is a short-term phenomenon and is likely to be corrected quickly. Table 6 confirms this conjecture. Only strategies following fresh positive sentiment in the stock market or following fresh positive sentiment in the stock market and fresh negative sentiment in the commodity market can generate significantly positive alphas.

VI. Robustness Check

We have performed sentiment strategies based on fund-days individually. We further examine the effects of sentiment in predicting future returns on the basis of portfolios in Table 7. In each long-short sentiment strategy described in the previous section, at the beginning of each day, we form three portfolios by including ETFs that have consecutive trading signals over the prior 1-, 2-, and 3-day intervals. Each portfolio is composed of pairs of a commodity ETF and Spider in a long-short position according to the prior consecutive sentiment signals. Using ETFs' reported net-of-expense returns, we calculate daily performance for each ETF for up to five days following the trading signal. To increase the power of our tests, we construct overlapping portfolios by following the methodology used in Jegadeesh and Titman (1993).[10] Each portfolio is equally weighted and held for up to five days following the portfolio formation. For each holding period, we regress the portfolio's daily performance on the 4-factor (Fama-French 3 factors plus a momentum) portfolios retrieved directly from the French website and report the intercept, which is in a percentage format.

[10]In any given day *t*, the strategies hold a series of portfolios that are selected in the current day as well as in the previous *K*-1 days, where *K* is the holding period. Specifically, a bull strategy that takes a long position on a commodity ETF and a short position on SPY on the basis of consecutive bullish signals over the past *J* days and holds them for *K* days. For instance, for a five-day holding strategy for positive sentiment in the stock market over the prior day, a Friday portfolio comprises commodity ETFs with the prior bullish stock signal on Thursday, Wednesday, and so on up to the previous Friday. Each day cohort is assigned an equal weight in this portfolio.

On average, there are about 14 commodity ETFs in a portfolio formed by an investor sentiment signal in the stock market. Over the sample period, 54% of the days are preceded by one-day bullish sentiment in the stock market and 45% of the days are preceded by one-day bearish sentiment. If a sentiment strategy cannot be performed on a regular basis, results from an overlapping portfolio strategy are less meaningful. In this regard, we focus only on the single-market and cross-market strategies based on sentiment in the prior day. On average, the portfolio in the single-market sentiment strategy based on the prior-day signal generates the 4-facor alpha of 17.2 bps, and its performance decays quickly to 4.4 bps in five days. A similar result is shown in the cross-market sentiment strategy.

VII. Conclusion

This study explores how investor sentiment in the stock market affects prices of commodity ETFs. We provide quantitative evidence that the tracking errors of commodity ETFs differ in the bull versus the bear stock market, and thus the aggregate tracking error of commodity ETFs is sensitive to the sentiment measure constructed by Baker and Wurgler (1997). We further exploit a profitable trading strategy based on investor sentiment in the stock market and commodity markets. We use commodity ETFs and Spider in a long-short strategy according to the prior sentiment signals. The sentiment-driven demand for commodity ETFs exists and is a short-term phenomenon. Only strategies following fresh positive sentiment in the stock market or following fresh positive sentiment in the stock market and fresh negative sentiment in the commodity market can generate significantly positive index-adjusted alphas about 23.7 bps to 8.6 bps per day on average, respectively. Following the methodology used in Jegadeesh and Titman (1993), we document that the portfolio in the single-market sentiment strategy based on the prior-day signal generates the 4-factor alpha of 17.2 bps, and its performance decays speedily to 4.4 bps in five days.

References

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

Amihud, Yakov, and Haim Mendelson, 1986, Asset pricing and the bid–ask spread, *Journal of Financial Economics 17*, 223–249.

Anderson, Richard Loree, and Theodore Alfonso Bancroft, 1952, *Statistical Theory in Research*, McGraw-Hill Book Company, New York.

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

Baker, Malcolm, and Jeffrey Wurgler, 2007, Investor sentiment in the stock market, *Journal of Economic Perspectives 21*, 129-151.

Banerjee, Abhijit V., 1992, A simple model of herd behavior, *Quarterly Journal of Economics 107*, 797-817.

Ben-Rephael, Azi, Shmuel Kandel, and Avi Wohl, 2012, Measuring investor sentiment with mutual fund flows, *Journal of Financial Economics 104,* 363-382.

Brown, Gregory W., and Michael T. Cliff, 2004, Investor sentiment and the near-term stock market, *Journal of Empirical Finance*, 1-27.

Brown, Gregory W., and Michael T. Cliff, 2005, Investor sentiment and asset valuation, *Journal of Business 78*, 405–440.

D'Avolio, Gene, 2002, The market for borrowing stock, *Journal of Financial Economics 66*, 271-306.

Delcoure, Natalya, and Maosen Zhong, 2007, On the premiums of iShares, *Journal of Empirical Finance 14*, 168-195.

DeLong, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann, 1990, Noise trader risk in financial markets, *Journal of Political Economy 98*, 703-738.

Elton, Edwin J., Martin J. Gruber, George Comer, Kai Li, 2002, Spiders: Where are the bugs? *Journal of Business 75*, 453–472.

Engle, Robert, Debojyoti Sarkar, 2006, Premiums-discounts and exchange traded funds, *Journal of Derivatives 13*, 27–45.

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

Levy, Ariel, and Offer Lieberman, 2013, Overreaction of country ETFs to US market returns: Intraday vs. daily horizons and the role of synchronized trading, *Journal of Banking and Finance 37*, 1412-1421.

Moreland, Richard L. and Scott R. Beach, 1992, Exposure effects in the classroom: The development of affinity among students, *Journal of Experimental Social Psychology 28*, 255-276.

Shleifer, Andrei, and Robert W. Vishny, 1997, The limits of arbitrage, *Journal of Finance 52*, 35-55.

Wurgler, Jeffrey, and Ekaterina Zhuravskaya, 2002, Does arbitrage flatten demand curves for stocks? *Journal of Business 75*, 583-609.

Table 1 Summary Statistics

A tracking error of an ETF is its gross return minus its benchmark index return. To calculate gross returns, we retrieve historical expense ratios of an ETF from SEC's EDGAR database and add the ratios to the ETF-reported net-of-expense returns on a daily basis. The benchmark index an ETF tracks is identified based on released prospectuses. We calculate daily tracking errors for an ETF since its inception and record the quartile distribution of tracking errors (TEs) for each commodity ETF. Our sample period starts on November 18, 2004, the first inception date in commodity ETFs, and ends by December 31, 2011. In Panel A, we report the cross-sectional average of quartile distributions of TEs. We also report the average (AVG) and standard deviation (STD) of the means of TEs across these 33 ETFs. For a reference comparison, we report the same statistics for Spider (symbol: SPY), an ETF tracking S&P 500 Index, over the same sample period. In Panel B, we calculate tracking-error volatility (TEV) for each ETF over its lifetime, and report the quartile distribution based on TEVs of all commodity ETFs. A tracking-error volatility of an ETF is the standard deviation of its tracking errors. We also report the average and standard deviation of the TEVs across these ETFs. To construct the quartile distribution of TEVs for SPY, we first calculate the TEV of SPY over a period of each ETF's lifetime and then report the statistics. All TEs and TEVs are in a percentage format.

	# of ETFs	Quartile Distribution					AVG	STD
		Lowest	25%	50%	75%	Highest		
Panel A. Tracking Errors								
Commodity	33	-15.785	-1.233	-0.007	1.193	16.954	-0.019	0.052
SPY		-1.621	-0.073	-0.001	0.070	2.939	0.000	0.217
Panel B. Tracking-Error Volatility								
Commodity	33	0.596	1.345	2.053	3.425	6.950	2.621	1.733
SPY		0.063	0.079	0.164	0.241	0.250	0.165	0.078

Table 2 Statistics of Daily Tracking Errors in Contrast Periods

The calculation of daily tracking errors for an ETF is described in Table 1. We classify the entire sample period into two contrast periods, bull versus bear, according to daily excess returns of the stock market. The bull (bear) stock markets include days that RMRF is positive (negative), where RMRF is one of the Fama-French three factors. For each commodity ETF in each period, we calculate the quartile distribution of tracking errors (TEs) and report the cross-sectional average for quartile distributions of TEs in Panel A. The average and standard deviation of the means of TEs across these ETFs are also reported in the last two columns. For a reference comparison, we report the same statistics for Spider. In Panel B, we calculate tracking-error volatility (TEV) for each ETF in each period, and report the quartile distribution based on TEVs of all commodity ETFs. The average and standard deviation of the TEVs across these ETFs are also reported in the last two columns. To construct the quartile distribution of TEVs for SPY, we first calculate the TEV of SPY each time for a period in which the TEV of a commodity ETF is calculated, and then report the statistics. All TEs and TEVs are in a percentage format.

	# of ETFs	Quartile Distribution					AVG	STD
		Lowest	25%	50%	75%	Highest		
Panel A. Tracking Errors								
Bull								
Commodity	33	-14.320	-1.088	0.085	1.249	15.459	0.107	0.590
SPY		-1.399	-0.093	-0.020	0.043	2.939	-0.025	0.206
Bear								
Commodity	33	-12.391	-1.405	-0.113	1.065	12.257	-0.172	0.733
SPY		-1.621	-0.042	0.021	0.102	1.556	0.031	0.228
Panel B. Tracking-Error Volatility								
Bull								
Commodity	33	0.562	1.310	2.011	3.489	6.457	2.589	1.716
SPY		0.054	0.071	0.157	0.228	0.235	0.155	0.075
Bear								
Commodity	33	0.581	1.426	1.992	2.873	6.988	2.548	1.672
SPY		0.070	0.083	0.169	0.252	0.263	0.172	0.082

Table 3 Cross-Market Tests on Individual Tracking Errors

We test for whether the mean of daily tracking errors (TEs) of an ETF differs in two markets, the stock market and the commodity market. In the stock market, a day is classified as a bull (bear) market if the daily RMRF, one of the Fama-French three factors, is positive (negative). In a commodity market, a day for an ETF is classified as a bull (bear) market if the daily return on an index tracked by the ETF minus daily one-month T-bill rate is positive (negative). For each ETF since its inception, the cross-market classification results in four periods: (1) bull stock market and bull commodity market (BullSBullC), (2) bull stock market and bear commodity market (BullSBearC), (3) bear stock market and bull commodity market (BearSBullC), and (4) bear stock market and bear commodity market (BearSBearC). For each ETF, we test for whether the mean of the TEs is identical in two contrast periods: BullSBullC versus BearSBullC as well as BullSBearC versus BearSBearC. For Spider (SPY) as a reference, we calculate the means and variances of its TEs in a period defined by each ETF and test for whether the mean of SPY's TEs is identical in two contrast periods: BullSBullC versus BullSBearC as well as BearSBullC versus BearSBearC. We list ETFs that have the significant difference and classify them according to the p-value (1%, 5%, or 10%) of the tests. For SPY, we list ETFs for which the period is defined and SPY has the significant difference in the tests. To consider that the population variances may not be equal in two periods, we use the modified t-test according to Satterthwaite's procedure described by Anderson and Bancroft (1952, p. 83).

p-value	#	Ticker Symbols of Commodity ETFs
Panel A:	H_0: The mean of TEs of an ETF is identical in both **BullSBullC** and **BearSBullC**.	
$p \leq 1\%$	25	USCI,GLD,IAU,DBC,SLV,USO,GSG,DBE,DBP,DBS,DBA,DBO,DGL, DBB,UCD,CMD,UCO,SCO,GLL,ZSL,AGQ,SIVR,DNO,PPLT,PALL
$p \leq 5\%$	2	USL,UGL
$p \leq 10\%$	2	SGOL,UNL
Panel B:	H_0: The mean of TEs of an ETF is identical in both **BullSBearC** and **BearSBearC**.	
$p \leq 1\%$	21	USCI,GLD,IAU,DBC,SLV,USO,GSG,DBE,DBO,DBB,UNG,GCC,UCD, CMD,UCO,SCO,ZSL,AGQ,SIVR,DNO,PALL
$p \leq 5\%$	6	DBS,DBA,UGL,GLL,SGOL,UNL
$p \leq 10\%$	3	DBP,USL,CORN
Panel C:	H_0: The mean of SPY's TEs is identical in both **BullSBullC** and **BullSBearC**.	
$p \leq 1\%$	0	
$p \leq 5\%$	1	SIVR
$p \leq 10\%$	0	
Panel D:	H_0: The mean of SPY's TEs is identical in both **BearSBullC** and **BearSBearC**.	
$p \leq 1\%$	2	UGL,GLL
$p \leq 5\%$	3	ZSL,AGQ,CORN
$p \leq 10\%$	4	SLV,USO,DBB,SIVR

Table 4 The Aggregate Tracking Errors and Investor Sentiment

The calculation of daily tracking errors for an ETF is described in Table 1. We calculate two daily liquidity measures, Amihud illiquidity and turnover, for each ETF. Amihud (2002) illiquidity is the absolute return divided by the dollar volume. The turnover is the ratio of trading volume to the number of shares outstanding. The investor sentiment measure (SENT) is constructed by Baker and Wurgler (2007) and directly retrieved from Wurgler's website. The sentiment measure is only available annually and monthly up to December 2010. Thus daily tracking errors and liquidity measures are converted to monthly data. We average the daily data for each month for each ETF and then calculate the cross-sectional average of monthly data for the entire commodity ETF. To avoiding the confounding, we exclude leveraged and inversed ETFs from the analysis for this table. We have 24 non-leveraged/non-inversed commodity ETFs in total and 74 months between November 2004 and December 2010. We regress monthly tracking errors of the commodity ETF portfolio on the liquidity measures and sentiment measures. To control the conditions of the general economy and the stock market, RMRF, we add one of the Fama-French three factors to the independent variables. Both tracking errors and RMRF are in a percentage format. The t-value associated with a coefficient estimate is in parentheses. As a reference, we also regress the tracking errors (TEs) of Spider on the same variables. TEs and liquidity measures for Spider are constructed in the same way to obtain monthly data. For brevity, we only report results of two models for SPY in the last two columns. Note that Amihud illiquidity is defined as 10^6 |r| / \$Vol for commodity ETFs and as 10^8 |r| / \$Vol for Spider.

Independent Variable	Dependent Variable					
	Tracking Errors of Commodity ETFs				TEs of SPY	
	Model 1	Model 2	Model 3	Model 4	Ref 1	Ref 2
Constant	-0.015	0.008	-0.026	-0.004	0.0002	0.002
	(-1.98)	(0.59)	(-2.66)	(-0.26)	(0.10)	(0.32)
RMRF	0.001	0.0003	0.002	0.002	-0.0003	-0.0004
	(0.61)	(0.21)	(1.23)	(0.93)	(-1.03)	(-1.10)
Turnover		-0.317		-0.348		-0.008
		(-2.12)		(-2.37)		(-0.65)
Amihud Illiquidity			0.689	0.781		5.994
			(1.76)	(2.05)		(0.07)
SENT	0.063	0.057	0.063	0.057	-0.0009	-0.0008
	(2.49)	(2.28)	(2.53)	(2.32)	(-0.17)	(-0.14)
Adjusted R^2	5.45	9.91	8.15	13.83	0.00	0.00

Table 5 Investment Strategies Based on Cross-Market Investor Sentiments

In the stock market, a day is classified as a bull (bear) market if the daily RMRF, one of the Fama-French three factors, is positive (negative). In a commodity market, a day for an ETF is classified as a bull (bear) market if the daily return on an index tracked by the ETF minus daily one-month T-bill rate is positive (negative). We further classify the extreme bull (bear) market for which the daily excess returns on the market are ranked at the top (bottom) third over the entire bull (bear) market period. For each ETF, we perform two long-short investment strategies depending on signals in a single market or cross-markets. In the single-market strategy, at the beginning of each day, we long an ETF and short SPY if the stock market was bull the prior day, whereas we short an ETF and long SPY if the stock market was bear the prior day. In the cross-market strategy, at the beginning of each day, we long an ETF and short SPY if the stock market was bull and the ETF market was bear the prior day (BullSBearC). We short an ETF and long SPY if the stock market was bear and the ETF market was bull the prior day (BearSBullC). Using ETFs' reported net-of-expense returns, we calculate daily performance for each strategy for each ETF. In Panel A, we report the distribution of average performance of each strategy on the basis of individual ETFs. In Panel B, we pool daily performance of each strategy across all ETFs and report the performance distribution based on all fund-days. Numbers in performance are in a percentage format. We report the t-value associated with the test if the average performance is zero. As a reference, we also report the performance of a plain strategy, simply longing an ETF and shorting SPY daily, without relying on any investment signal. To avoid the confounding, we exclude leveraged and inverse ETFs from the analysis for this table. The sample period from November 18, 2004, to December 31, 2011, contains 24 non-leveraged/non-inverse commodity ETFs.

Strategy	**#Obs**	**Quartile Distribution**					**AVG**	**STD**	**t**
		Min	25%	50%	75%	Max			
Panel A. On the basis of individual ETFs									
Without Signals	24	-0.202	-0.031	0.013	0.061	0.091	0.007	0.074	0.48
With Signals									
Single-Market	24	-0.102	0.109	0.150	0.182	0.305	0.141	0.097	7.10
Cross-Market	24	-0.194	0.136	0.160	0.262	0.440	0.183	0.138	6.48
Panel B. On the basis of pooling observations across all ETFs									
Without Signals	25690	-17.56	-1.023	0.020	1.095	21.25	0.010	2.102	0.80
With Signals									
Single-market	25525	-14.05	-0.981	0.069	1.142	21.25	0.156	2.097	11.91
Bull	13861	-12.32	-0.888	0.076	1.089	21.25	0.152	1.950	9.18
Extreme Bull	5337	-12.32	-0.909	0.123	1.256	21.25	0.233	2.212	7.69
Bear	11664	-14.05	-1.101	0.061	1.206	17.56	0.161	2.260	7.71
Extreme Bear	4669	-12.11	-1.027	0.269	1.533	17.56	0.391	2.541	10.50
Cross-Market	11149	-12.11	-1.024	0.108	1.205	17.56	0.180	2.120	8.97
BullSBearC	5879	-10.60	-0.908	0.139	1.245	14.57	0.223	2.039	8.37
Extreme BullSBearC	697	-8.01	-1.244	0.301	1.947	12.64	0.449	2.758	4.30
BearSBullC	5270	-12.11	-1.126	0.051	1.162	17.56	0.133	2.206	4.36
Extreme BearSBullC	702	-8.38	-1.270	0.079	1.638	17.56	0.340	2.844	3.17

Table 6 Alphas of Investment Strategies

Table 5 defines investment strategies. For each ETF, at the beginning of each day, we perform long-short investment strategies depending on the prior-day bull/bear signals of the stock and commodity markets. According to the position of an ETF's trading signal in the timeline, we further classify each strategy into three mutually exclusive groups. Strategies following signals that are fresh new are labeled as the "1st Signal" group. Strategies following the signals that are 2nd or 3rd in the sequence of consecutive same signals are labeled as the group of "2≤ #Consecutive Signals ≤3." Strategies following the signals preceded by at least three consecutive same signals are labeled as the group of " #Consecutive Signals >3." Using ETFs' reported net-of-expense returns, we calculate daily performance for each strategy for each ETF. We pool daily performance of each strategy across all ETFs in each group and perform a regression. We regress strategy performance on the return difference between the S&P 500 index and the underlying index the ETF tracks. Note that when we long SPY and short an ETF in the Y variable, the X variable will be the S&P 500 index returns minus returns on the index that the ETF tracks. When we long an ETF and short SPY in the Y variable, the X variable will be returns on the index the ETF tracks minus the S&P 500 index returns. We report the regression results. In the brackets, we report the average of raw returns in strategies and numbers of observations for each group. Alphas, returns, and Adjusted R^2 ($AdjR^2$) are all in a percentage format. We only analyze non-leveraged/non-inverse commodity ETFs over the sample period from November 18, 2004, to December 31, 2011, for this table. The significance level of returns/alphas equaling to 0 or betas equaling to 1 is indicated by *** (1%), ** (5%), and * (10%).

$$\tilde{Y} = \alpha + \beta\tilde{X} + \tilde{\varepsilon}$$

$$Where\ \tilde{X} = \tilde{R}_{Index\ tracked\ by\ the\ ETF} - \tilde{R}_{S\&P500} \quad if \quad \tilde{Y} = \tilde{R}_{ETF} - \tilde{R}_{SPY}$$

$$\tilde{X} = \tilde{R}_{S\&P500} - \tilde{R}_{Index\ tracked\ by\ the\ ETF} \quad if \quad \tilde{Y} = \tilde{R}_{SPY} - \tilde{R}_{ETF}$$

Strategy	1st Signal			2≤ #Consecutive Signals ≤3			#Consecutive Signals >3		
	α	β	$AdjR^2$	α	β	$AdjR^2$	α	β	$AdjR^2$
Single-market	0.092***	0.509***	37.23	0.020	0.458***	31.65	-0.045	0.578***	41.90
	[0.237***	/ 3428]		[0.146***	/19397]		[0.130***	/ 2699]	
Bull	0.237***	0.557***	45.06	0.001	0.468***	30.69	-0.039	0.445***	25.40
	[0.430***	/ 1427]		[0.127***	/10688]		[0.083**	/ 1745]	
Bear	-0.008	0.446***	28.74	0.042**	0.450***	32.55	-0.015	0.646***	51.56
	[0.099**	/ 2001]		[0.169***	/ 8709]		[0.218**	/ 954]	
Cross-Market	0.051**	0.522***	36.81	0.002	0.370***	24.87	0.212	0.653***	43.41
	[0.196***	/ 5437]		[0.157***	/ 5638]		[0.732***	/ 74]	
BullSBearC	0.086***	0.519***	37.80	0.022	0.381***	23.98	0.150	0.573***	38.39
	[0.257***	/ 2731]		[0.187***	/ 3099]		[0.583***	/ 49]	
BearSBullC	0.016	0.525***	35.61	-0.025	0.362***	25.60	0.435	0.673**	43.15
	[0.136***	/ 2706]		[0.120**	/ 2539]		[1.025	/ 25]	

Table 7 The 4-Factor Alphas of Investment Strategies

Table 5 defines investment strategies. In each strategy, at the beginning of each day, we form three portfolios by including ETFs that have consecutive trading signals over the prior 1-, 2-, and 3-day intervals. Using ETFs' reported net-of-expense returns, we calculate daily performance for each ETF for up to five days following the trading signal. To increase the power of our tests, we construct overlapping portfolios by following the methodology used in Jegadeesh and Titman (1993). Each portfolio is equally weighted and held for up to five days following the portfolio formation. For each holding period, we regress the portfolio's daily performance on the 4-factor (Fama-French 3 factors plus a momentum) portfolios retrieved directly from the French website and report the intercept, which is in a percentage format. We only analyze non-leveraged/non-inverse commodity ETFs for this table. The significance level of alphas equaling to zero is indicated by *** (1%), ** (5%), and * (10%). We also report the average number of ETFs in each portfolio and the percentage of days for which we have trading signals in each strategy over the entire sample period from November 18, 2004, to December 31, 2011.

Strategy	# prior consecutive signals	# ETFs	% days covered	Holding Days				
				+1	+2	+3	+4	+5
Single-market	1	14	99	0.172***	0.116***	0.072***	0.048**	0.044**
	2	15	47	0.260***	0.257***	0.220***	0.189***	0.172***
	3	15	22	0.316***	0.309***	0.309***	0.237***	0.212***
Bull	1	14	54	0.141***	0.140***	0.140***	0.141***	0.140***
	2	14	28	0.158***	0.156***	0.156***	0.156***	0.155***
	3	15	14	0.083	0.081	0.080	0.081	0.081
Bear	1	14	45	0.110**	0.110**	0.110**	0.110**	0.110**
	2	15	19	0.218***	0.218***	0.218***	0.218***	0.218***
	3	15	8	0.305**	0.305**	0.305**	0.305**	0.305**
Cross-Market	1	7	87	0.165***	0.123***	0.080***	0.044*	0.048**
	2	4	29	0.319***	0.300***	0.223***	0.200***	0.186***
	3	3	8	0.314	0.326	0.330*	0.311	0.237
BullSBearC	1	7	47	0.153***	0.142***	0.135***	0.120***	0.128***
	2	4	17	0.227**	0.193**	0.181*	0.178*	0.174*
	3	3	5	-0.139	-0.160	-0.161	-0.161	-0.161
BearSBullC	1	7	40	0.083	0.093*	0.114**	0.113**	0.108**
	2	4	13	0.237*	0.231*	0.251**	0.282**	0.282**
	3	3	3	0.430	0.461	0.471	0.471	0.359