Market Sentiment in Commodity Futures Returns

Lin Gao[a] and Stephan Süss[b]

June 28, 2015

Abstract

We identify a strong presence of sentiment exposure in commodity futures returns. Sentiment is able to provide additional explanatory power for comovement among commodity futures beyond the macro- and equity-related sources. Commodity futures with low open interest growth, high volatilities, low momentum, or low futures basis are more sensitive to change in sentiment. Similar to Baker and Wurgler (2006), we construct a market sentiment index by Partial Least Squares regressions (PLS) with non-return based stock market proxies, in particular higher moments of the option implied return distribution. Moreover, our sentiment index can be built on a daily basis.

JEL Classification: G12, G13

Keywords: Commodity Futures, Market Sentiment

---

[a] Corresponding author. University of Luxembourg, Luxembourg School of Finance, 4 rue Albert Borschette, L-1246 Luxembourg. Tel: +352 466 644 5482. Email: lin.gao@uni.lu

[b] University of St. Gallen, Swiss Institute of Banking and Finance, Rosenbergstrasse 52, CH-9000 St. Gallen, Switzerland. Tel: +41 43 244 7585. Email: stephan.suess@unisg.ch.

Economists have traditionally regarded commodity futures prices as fully informative about future economic activity and asset prices. The theory of backwardation implies that the risk premium depends only on fundamentals such as the net supply–demand imbalance among hedgers in the futures market (Keynes (1923) and Hicks (1939)). However, fundamentals could not explain the puzzle demonstrated in Pindyck and Rotemberg (1990). The authors argue that the co-movement among futures returns of different commodity markets should only be driven by common macroeconomic information. After controlling for these factors, commodity returns still significantly correlate with one another. Pindyck and Rotemberg (1990) interpret this phenomenon as a potential herding effect without further empirical tests.

In this paper, we find that sentiment contains information about commodity futures returns and explains the co-movement among the returns of various commodity futures. Through a sentiment index which is available both at a monthly and a daily level, we identify a significant role of investor sentiment in commodity futures returns, especially in the recent period. Its explanatory power remains robust after controlling for stock market returns, macroeconomic variables, and commodity related factors. Therefore, sentiment represents a distinct source of premia. When sorting the commodity futures based on their conditional characteristics into portfolios, we find that commodity futures with high open interest growth, high volatilities, low momentum, or high futures basis are likely to earn higher returns during bullish sentiment periods, whereas portfolios with low open interest growth, high volatilities, low momentum, or low futures basis are more sensible to change in sentiment.

As argued by Baker and Wurgler (2006), two conditions exist for sentiment to take effect in the price formation: the existence of speculative demand and arbitrage constraints in the market. The commodity futures market has features that suggests the existence of market sentiment. The share of financial investors in commodity futures has grown dramatically in the recent decade. As an illustration, Figure 1 shows the proportion of the different types of

1

traders with long positions in the wheat futures market to the total number of long traders. Financial investors, represented by swap dealers and money managers as defined by the Commodity Futures Trading Commission (CFTC), make up a total of around 80%. Similar situations prevail in other commodity futures markets. This development became known in the academic literature as the "financialization" of commodity markets.

[**Please insert Figure 1 around here**]

It is particularly worth mentioning that in Figure 1 swap dealers (index providers) account for around 60% of the total long traders in this market. Swap dealers are typically known as taking a long position in the futures market in order to hedge their index products, which are sold to their clients over the counter. As a result, the commodity futures markets do not just involve a large-scaled participation of financial investors, but the pressure, coming from the long side of the index providers, is rather persistent and substantial relative to the weights of the other types of traders.

A growing co-movement between commodities and stocks (Buyuksahin, Haigh, and Robe (2010)) and among different commodity groups (Tang and Xiong (2012)) has been observed since the beginning of the last decade[1]. Domanski and Heath (2007), Masters (2008), Buyuksahin, Haigh, Harris, Overdahl, and Robe (2008), Mayer (2009), Buyuksahin and Robe (2011), and Tang and Xiong (2012) attribute the effect to the increasing share of financial investors in the market for commodity futures (financialization) and to the hedging activities of commodity index providers. Mou (2011) disclose that index provider's rolling between contracts of different maturities has a significant effect on price levels. Buyuksahin and Robe (2012) argue that trading activities of hedge funds, particularly those that are active in both the commodity and the stock market, help explain the recent increase in correlation. Singleton (2011) conclude that the intermediate-term growth rates of index

---

[1]This strong comovement is in sharp contrast with the period before around 2003. For the earlier period, commodity futures returns are found to be very little correlated with equities (see, e.g. Dusak (1973), Erb and Harvey (2006) or Gorton and Rouwenhorst (2006).

provider positions and hedge fund positions had the largest impact on futures prices during the 2006–2010 sample period. Consequently, financial investors' behavior leads to changes in the risk premia of commodity futures (Hamilton and Wu (2014)). Henderson, Peason, and Wang (2015) provide evidence that investor flows into and out of commodity-linked notes have important impacts on commodity prices.

Financial traders help increase the liquidity of the market. However, due to the existence of short-sale constraints, prices could be extorted by the mood of those financial investors. Schenkman and Xiong (2003) suggest three reasons that prevent arbitrageurs from short selling. First, the price of borrowing a security can be expensive because the default risk of the potential price increase has been priced into the security. Second, arbitrageurs may be risk averse and are therefore deterred by the short-selling risk. Third, capital constraints of arbitrageurs in an extreme market situation may limit short selling. Their theory finds support in the commodity market evidence presented by Acharya, Lochstoer, and Ramadorai (2013), where increases in producers' hedging demand or speculators' capital constraints increase hedging costs via price-pressure on futures. Cheng, Kirilenko, and Xiong (2012) show that when the VIX index increases, the positions of commodity futures arbitrageurs will decrease. They argue that arbitrageurs' capital is more constrained during these periods. In commodity markets, due to delivery risk, most asset managers of pension funds are limited to participating directly or going short in the market. Therefore, they can only invest through a commodity index. Hedge fund managers may go short, but obtaining commodities physically is relatively difficult. This also adds to the fact that even arbitrageurs such as hedge funds intend to exploit the difference between the traded price and the fundamental value, they are betting against a substantial number of investors who are willing to take a long position in commodities. Hence, such arbitrage turns from riskless opportunities to costly speculative bets (See, e.g. Kondor (2009)).

Such a dominance of the index providers, who are usually long on the futures market, makes price corrections through short selling rarely effective. Thus, the growing participation

3

of financial investors, in particular index investors, combined with a market having limited short selling possibilities, creates the necessary conditions for sentiment to play a significant role.

What is market sentiment? Market sentiment is a concept that reflects investors' moods and beliefs. Most theoretical papers (e.g. DeLong, Shleifer, Summers, and Waldmann (1990)) on market sentiment distinguish between two types of traders: noise traders, who hold random beliefs, and arbitrageurs, who hold Bayesian beliefs. Concerning the effect of market sentiment on returns, the level of the noise traders' beliefs relative to the Bayesian beliefs is often referred to as market sentiment (see, e.g. Tetlock (2007)).

On the one hand, when a dominant number of noise traders have expectations of future prices that are beyond the expectations of rational arbitrageurs, market sentiment is bullish. In turn, a high sentiment spurs speculative behavior regarding both the number of investors and the average magnitude of the investments. As such, it is highly related to market liquidity.[2] In the face of existing arbitrage constraints, however, investor sentiment can drive asset prices away from their fundamental values. In Section 2, we provide a theoretical motivation in the framework of a noisy rational expectations equilibrium. Our model suggests that when arbitrage through short-selling is difficult, sentiment offers an additional premium to commodity futures returns which is associated with investors' risk tolerance γ_S.

On the other hand, sentiment also captures the uncertainty and risk that market participants perceive (Miller (1977)). Hong and Stein (2003) show that differences of opinion, combined with short-sale constraints, lead to mis-valuation, which is highly dependent on the risk that investors perceive. In this view, the investors' disagreements and fears affect prices in the markets with relatively weak arbitrage forces. Connolly, Stivers, and Sun (2005) find that stock–bond return correlation is low during high market uncertainty, as proxied by the VIX index. Etula (2013) models the broker-dealer risk appetite using information on

---

[2]There has been a set of recent articles investigating the dependence between investor mood and market liquidity. See, for example, Nagel (2011), Da, Liu, and Schaumburg (2012), and the references therein.

the aggregate assets and liability values of U.S. security broker-dealers and households. He shows that the risk bearing capacity proxy can predict in particular energy returns up to two quarters ahead.

Prior research on investor sentiment has examined stock returns (see, for example, Baker and Wurgler (2006)), equity option prices (Han (2008)), as well as on stock–bond return correlations (Connolly, Stivers, and Sun (2005)). To the best of our knowledge, this study is the first to provide evidence that investor sentiment drives commodity futures returns. In the face of the continuous changes regarding investor composition and high prices variations in the recent period, this study improves our understanding of this market.

We construct a sentiment index in the spirit of Baker and Wurgler (2006) (BW2006, henceforth). As sentiment proxies, we adopt CBOE's VIX and SKEW indexes, which measure the return volatility and skewness implied in S&P 500 index options. Both variables have been employed individually in previous studies as sentiment proxies in equity and bond markets (see, e.g., Coudert and Gex (2008) and Han (2008)). In addition, we compute a daily version of the closed-end fund discount and the dividend premium. So far, these proxies have only appeared at a monthly or even lower frequency in the literature, such as in BW2006. To reduce the dimension of these exogenous variables, we construct a composite sentiment index by the Partial Least Squares (PLS) method, an approach similar to the Principal Component Analysis (PCA) employed in BW2006.

We consciously adopt equity-related proxies instead of commodity-related proxies for our sentiment index. First, the equity market is still the most liquid market, hence proxies from this market can be representative of general market sentiment. Second, with equity-proxies, we are spared the additional effort needed to argue that the explanatory power of our sentiment proxies does not come from the systematic risks of the (same) market.

Compared to BW2006, our sentiment index has the following advantages: First, our index can be built on a daily basis, which is a substantial advantage regarding the timeliness

of the sentiment measure. Second, using higher moments of the option implied return distribution, the measurement errors and other confounding influences inherent in nonreturn proxy variables can be reduced. Moreover, the extreme investor's mood in the tails of the asset returns can be captured.

In addition, this study contributes to the literature on commodity futures pricing. So far, only a few factors have been found to have significant explanatory or even predictive power. Some of them are macro variables, such as the short rate (Frankel and Wurgler (2010)), exchange rates (Chen, Rogoff, and Rossi (2010)), and industrial production growth (Borensztein and Reinhart (1994)). Others are commodity-related. Hong and Yogo and (2010, 2012) provide empirical evidence that open interest has predictive power for commodity futures returns. Fama and French (1987), Bailey and Chan (1993) and Gorton, Hayashi, and Rouwenhorst (2013) find that the difference between spot and futures prices (basis) is substantially associated with commodity futures returns. Fuertes, Miffre, and Rallis (2010) and Moskowitz, Ooi, and Pedersen (2012) suggest the existence of a momentum effect. We show that our sentiment index remains highly significant after controlling for these factors and provides a significant incremental explanatory power.

Our study is organized as follows: Section 2 presents a theoretical motivation for the sentiment exposure in commodity futures returns. Section 3 describes our sentiment proxies and outlines the mechanism of the construction of the index. The details about our dataset can be found in Section 4. Our empirical results are outlined in Section 5. Finally, Section 6 concludes.

2   A Theoretical Motivation

This section gives a brief outline of the structure of the model. It characterizes the prices of commodity futures in a simple two-period noisy rational expectations equilibrium. The setup has roots in Chen, Hong, and Stein (2002). It shows that the prices of commodity

futures are not only driven by supply and demand for the underlying good, but are also driven by investors' impact. As financial investment in commodity markets is typically conducted by commodity index products, it is difficult for a large part of the investors to short commodity goods. As such, we show that futures prices can be extorted away from their fundamental values. In addition, as increased output may drive commodity prices down, increased financial investor engagement will lead to rising contango levels.

2.1 The Model Economy without Short-Selling Restrictions

Our model treats the pricing of a single homogeneous good in a noisy rational expectations equilibrium model. Consider an artificial economy with two discrete time periods t and T ($T > t$) in which the commodity is traded. The final time T price is determined by the aggregate consumption demand at T. For simplicity, we assume that the complete commodity supply Q_t, which is produced at t, is consumed at T at the stochastic price

$$\tilde{p}_T = \tilde{a} - Q_t, \tag{1}$$

where the uncertainty about future consumption demand is modeled by a normally distributed shock $\tilde{a}$. There exist two main classes of traders in the economy. First, there is a group of commodity producers P. Without loss of generality, let their mass be normalized to $(1 - \omega)$, where $0 \leq \omega < 1$. We assume that they have an identical constant absolute-risk-aversion (CARA) utility, with risk tolerance γ_P. In period t, they observe the price of a futures contract $F_t(T)$ maturing at T. Let the interest rate r be zero. This avoids the need for a specific futures pricing model. We assume that producers have time-t expectations of $E_t^P[\tilde{p}_T|F_t(T)] = \mu$ and $Var_t^P[\tilde{p}_T|F_t(T)] = \sigma_\varepsilon^2$. Their production decision is determined by some convex cost function $C(Q_t)$, increasing in the output level Q_t. For simplicity, we assume that $C(Q_t) = \frac{1}{2}cQ_t^2$. Furthermore, producers can hedge their price exposure by selling z_t futures contracts. This means that they effectively make "purchases of futures" equal to

7

$x_t = Q_t - z_t$. We maximize the time-T wealth of a representative producer i

$$\tilde{W}_{i,T}^P = (F_t(T)Q_{i,t} - \frac{1}{2}cQ_{i,t}^2) + [\tilde{p}_T - F_t(T)]\, x_{i,t} + W_{i,t}^P, \tag{2}$$

where $W_{i,t}^P$ denotes the corresponding initial wealth. This leads to a production decision of $Q_{i,t} = \frac{F_t(T)}{c}$, which aggregates to a total supply of $Q_t = (1-\omega)\frac{F_t(T)}{c}$. It is thus dependent on the futures pricing prevailing in time t. Forward prices therefore have a quantity and price impact on commodities. The representative producer also makes "futures purchases" of $x_{i,t}^P = \gamma_P \frac{\mu - F_t(T)}{\sigma_\varepsilon^2}$.

The second class of traders consists of financial investors, such as institutional traders who try to hedge their commodity exposures, professional commodity index traders, or retail investors. To model heterogeneity among them, we assume that each member either belongs to a group of "optimists" U or "pessimists" D, with individual valuations of $E_t^U[\tilde{p}_T] = V^U = \mu + H$ and $E_t^D[\tilde{p}_T] = V^D = \mu - H$, respectively, where $H > 0$. For the sake of simplicity, assume that the group shares of the optimists and the pessimists are both equal to 0.5, so that the average belief of financial investors is unbiased. Assuming a CARA utility function with risk tolerance γ_S for all financial speculators, individual investor demand for futures is therefore $x_{i,t}^S = \gamma_S \frac{V_i^{U/D} - F_t(T)}{\sigma_\varepsilon^2}$. The futures market clearing condition in t is $x^S(t) = \sum_i x_i^S(t) = z_t$. As $z_t = Q_t - x_t^P$, we have $x_t^P + x_t^S = Q_t$, i.e.,

$$\frac{(1-\omega)\gamma_P}{\sigma_\varepsilon^2}(\mu - F_t(T)) + \frac{\omega\gamma_S}{2\sigma_\varepsilon^2}(\mu - H - F_t(T)) + \frac{\omega\gamma_S}{2\sigma_\varepsilon^2}(\mu + H - F_t(T)) = \frac{(1-\omega)F_t(T)}{c}. \tag{3}$$

Solving for $F_t(T)$ leads to

$$F_t(T) = \frac{\mu}{1 + \frac{(1-\omega)\sigma_\varepsilon^2}{c\bar{\gamma}}}, \tag{4}$$

where $\bar{\gamma} = \gamma_P + \omega(\gamma_S - \gamma_P)$ is the weighted average risk tolerance of the market participants. For simplicity of notation, let $\phi = 1 + \frac{(1-\omega)\sigma_\varepsilon^2}{c\bar{\gamma}}$ be the expected gross return of $F_t(T)$, so that $F_t(T) = \frac{\mu}{\phi}$. For expectations of the producers to be rational, we must have $\mu = E_t^P[\tilde{p}_T] = E[\tilde{a}] - Q_t = E[\tilde{a}] - \frac{(1-\omega)\mu}{c\phi}$ and $Var(\tilde{p}_T) = Var(\tilde{a}) = \sigma_\varepsilon^2$. Solving for μ, we can write $F_t(T)$

8

and $\tilde{p}_T$ in primitives by

$$
\begin{aligned}
F_t(T) &= \frac{E[\tilde{a}]}{\theta}, \\
\tilde{p}_T &= \frac{\phi E[\tilde{a}]}{\theta} + (\tilde{a} - E[\tilde{a}]), \quad \text{where} \\
\phi &= 1 + \frac{(1-\omega)Var(\tilde{a})}{c\bar{\gamma}}, \quad \theta = \phi + \frac{1-\omega}{c}, \text{ and } \bar{\gamma} = \gamma_P + \omega(\gamma_S - \gamma_P).
\end{aligned}
\tag{5}
$$

Note that $E[\tilde{p}_T] = \phi F_t(T)$, which allows an investigation of the return premia in more detail.

This result has several appealing features: First, the expected return premium is increasing in $Var(\tilde{a})$. Producers sell futures contracts to decrease their exposures to the consumption risk. With a high ω there is only little insurance need for producers, i.e., $(1-\omega)Var(\tilde{a})$ is small. We observe a futures price close to the expected fundamental commodity value μ. Most of the risk is diversified by the bulk of financial investors. Furthermore, we observe that the difference between the future price and the fundamental value decreases with high production costs c. A high $F_t(T)$ induces producers to expand their production level. If production costs are high, the elasticity of the supply is very small, i.e., $\partial Q_t / \partial F_t(T) = \frac{1}{c}$. A higher share of financial investors has two distinct effects: first, they act as providers of insurance to producers against consumption risk. The more financial investors are in the market, the more the risk is diversified and the higher the production level. Second, they influence the average risk tolerance in the market $\bar{\gamma}$. However, ex ante, the sign of this effect depends on whether their risk tolerance is higher or lower than that of the producers. In the case of $\gamma_P < \gamma_S$, it makes sense to increase the production level and diversify a large part of the consumption risk over the financial investors. However, if $\gamma_P > \gamma_S$, a larger fraction of the consumption risk remains with the producers. Note that the future price does not depend on the dispersion of the investors' valuations H. If a specific investor i has a valuation higher than μ, he takes a short position and vice versa. The deviation of $E_t[\tilde{a}]$ from $E[\tilde{a}]$ entirely derives from the effect of expanding production levels and not from speculation.

2.2 Imposing Short-Selling Restrictions

This section extends our basic model by imposing restrictions on short-selling by financial investors. As before, group members have individual valuations of $E_t^U[\tilde{p}_T] = V^U = \mu + H$ and $E_t^D[\tilde{p}_T] = V^D = \mu - H$, respectively. For concreteness, one may think of the long-only investors as institutional financial firms replicating a commodity index. Alternatively, this group may consist of retail investors, who are restricted from trading directly on the derivatives market and take long positions in some commodity index tracking product. If their valuations exceed the prevailing market price, they stay out of the market instead of taking short positions. We think that this is an appropriate characterization of the commodity investment market. For simplicity, we assume that the traders of both subgroups have the same CARA utility, with a constant risk tolerance of γ_S. Note, that both the investor groups still make up an unbiased valuation on average.

Futures market clearing implies

$$
\begin{aligned}
\frac{F_t(T)}{c} = {} & \frac{(1-\omega)\gamma_P}{\sigma_\varepsilon^2}(\mu - F_t(T)) + \\
& \frac{\omega\gamma_S}{2\sigma_\varepsilon^2}\mathbf{1}_{\{F_t(T)\leq\mu-H\}}(\mu - H - F_t(T)) + \frac{\omega\gamma_S}{2\sigma_\varepsilon^2}\mathbf{1}_{\{F_t(T)\leq\mu+H\}}(\mu + H - F_t(T)).
\end{aligned}
\tag{6}
$$

where $\mathbf{1}$ denotes the indicator function. Solving for $F_t(T)$ now results in

$$
F_t(T) = \frac{\mu}{\phi} + \frac{\omega\gamma_S H}{2\bar{\gamma}'\phi},
\tag{7}
$$

where $\bar{\gamma}' = \gamma_P + \mathbf{1}_{\{F_t(T)\leq\mu+H\}}\omega((\frac{1}{2} + \frac{1}{2}\mathbf{1}_{\{F_t(T)\leq\mu-H\}})\gamma_S - \gamma_P)$ is again the average risk tolerance of market participants. Rational expectations require that $\mu = \frac{E[\tilde{a}]}{1+\frac{1-\omega}{c\phi}} - \frac{\omega\gamma_S H}{2c\bar{\gamma}\phi}$, such that

$$
F_t(T) = \frac{E[\tilde{a}]}{\theta} + \frac{\omega\gamma_S H}{2c\bar{\gamma}\phi}\frac{\phi c - \omega}{\theta}
\tag{8}
$$

10

and

$$\tilde{p}_T = \frac{E[\tilde{a}]}{1 + \frac{1-\omega}{c\phi}} - \frac{\omega\gamma_S H}{2c\bar{\gamma}\theta} + (\tilde{a} - E[\tilde{a}]). \tag{9}$$

As such, we observe a premium on futures prices, as pessimists stay out of the market. This premium increases with the risk tolerance of the speculators γ_S. As is well known in the equity pricing literature, preventing part of the investors from taking short selling positions leads to an aggravated price level. This is expressed here by the higher futures price.

3 The Sentiment Index

In this section, we explain the construction of our sentiment index. Our approach combines the moments of the option implied S&P 500 return distribution, which have been verified in the literature as sentiment proxies, with the BW2006 variables. Specifically, we employ option implied volatility, option implied skewness, together with the BW2006 proxies— closed-end fund discounts (CEFD), first-day returns of initial public offerings (IPOs), NYSE turnover, and dividend premium[3].

3.1 Sentiment Proxies in Baker and Wurgler (2006)

The four proxies adopted by our study, CEFD, IPO, Turnover, and Dividend Premium, have been given in detail in BW2006 as sentiment proxies. Therefore, we summarize the most important features of these proxies and explain how they are expected to be related to commodity markets.

Closed-End Fund Discounts (CEFD). Closed-end funds are exchange-traded companies that hold a portfolio of securities. They are "closed-end" because they issue a fixed number of outstanding shares, which are not redeemable until the fund liquidates. In violation of the

---

[3]In our previous experiment, we also tried the option implied volatility premium, implied skewness premium, the number of IPOs, and the share of equity issues in new issues. Due to repetitive or insignificant results, we dropped these proxies to keep our factor structure slim.

law of one price, closed-end funds typically trade at a lower market price than the net asset values (NAV) of the fund's actual security holdings. This difference is called the closed-end fund discount. Empirical research relates the explanation of CEFD to investor sentiment: Lee, Shleifer, and Thaler (1990, 1991) argue that the CEFD is inversely related to sentiment. Rising discount values reflect periods of bearish investor expectations.[4]

First-Day Returns of IPOs. New equity issues typically exhibit positive first-day returns on average, and are frequently considered as "underpriced". Baker and Wurgler (2007) argue that first-day IPO returns reflect investors' enthusiasm. They show that extreme peaks and troughs of first-day IPO returns are highly correlated with IPO volume and other sentiment proxies. Derrien (2005) find that the share of retail investors is positively related to IPO returns even after controlling for the participation of institutional investors.

Trading Volume. In markets with short-sale restrictions, increasing divergence of opinions about the fundamental asset price leads to rising trading volumes, allocating the shares to the most optimistic market participants. As such, in periods of high market sentiment, (overly) optimistic investors add liquidity to the market. However, in pessimistic periods, these participants are less likely to trade, because short selling is restricted. The trading volume thus decreases.[5] Hence, in equilibrium, bubbles are typically accompanied by large trading volumes and volatile prices. Jones (2001) finds a negative relation between turnover and future returns.

Dividend Premium. Baker and Wurgler (2006) define this as the difference between the average market-to-book ratios of dividend-paying stocks and non-dividend-paying stocks. Dividend-paying stocks resemble bonds, because their predictable dividend streams provide safety, which is highly valued in market declines. Baker and Wurgler (2004b, 2004a) argue that managers cater to investor sentiment for safety with dividend payments. Hence,

---

[4]The most important studies on closed-end fund discounts include Zweig (1973), Lee, Shleifer, and Thaler (1991), Chopra, Lee, and Thaler (1993), Pontiff (1996), Neal and Wheatley (1998), and Baker and Wurgler (2004b, 2007).

[5]A theoretical model for this can be found in Schenkman and Xiong (2003) and Hong and Stein (2003).

dividend premia should be inversely related to sentiment.

3.2 Option Implied Factors as Sentiment Proxies

In the following, we list the option implied factors from the S&P 500 index and explain their relation to sentiment. Using equity-related instead of commodity-related proxies helps us avoid the potential concern that the latter may reflect some systematic risks in the commodity market.

Option Implied Volatility. The option implied volatility index, VIX, which is defined as the square root of the variance swap rate implied in S&P 500 options, is frequently referred to as an "investor fear gauge" by practitioners (Whaley (2000)). Apart from the fact that the VIX reflects stock market uncertainty, it also contains information about risk aversion. Carr and Wu (2009) decompose the squared VIX into an expected conditional variance part and a variance risk premium part. Academic studies relate the expected variance to economic uncertainty and link the variance risk premia to risk aversion (See, for instance, Bekaert and Engstrom (2009), Bollerslev, Tauchen, and Zhou (2009), and Drechsler and Yaron (2011)). We find that the VIX is highly correlated with the variance risk premia and its explanatory power for commodity returns comes mainly from the variance risk premia part. Therefore, to remain simple in the factor estimation, we continue to use the VIX index as our sentiment proxy. Whaley (2000) shows that the index has its highest levels during periods of financial market turmoil and investor fear, as investors hedge their stock portfolios with long positions in S&P 100 put options[6]. We therefore expect that this factor is negatively correlated with commodity returns.

Option Implied Skewness. The use of implied skewness as a possible proxy for measuring investor sentiment has been motivated by extensive empirical work suggesting limits to

---

[6]Before its change in the calculation method in September 2003, the VIX was originally defined as the Black–Scholes implied volatility of a hypothetical 30 calendar day at-the-money option written on the S&P 100. This definition is now disseminated under the ticker symbol VXO.

arbitrage in the options market, something which permits sentiment to affect options prices[7]. A large body of literature has focused on the impact of investor sentiment on the steepness of the volatility smile or, equivalently, on the skewness of the implied return distribution. Han (2008) reveals that the index volatility smile of S&P 500 options is steeper when market sentiment becomes more bearish. Accordingly, the risk neutral skewness of monthly index returns is more negative during these periods[8]. It is well documented that the empirical distribution of monthly equity index returns is approximately symmetric, whereas its risk neutral counterpart is determined by the slope of the pricing kernel: investors' desire to avoid value loss in their portfolios shifts the probability mass to the negative tail of the implied equity index return distribution relative to its physical counterpart (Bakshi and Madan (2006)). This can be observed as a negative implied skewness. As such, the implied skewness can be interpreted as a direct market assessment of downside risk relative to upside risk.

The option implied sentiment proxies have several desirable properties. First, the required data are available at high frequency and liquidity levels, which promises a timely measure of sentiment with low noise levels. Furthermore, they are derived from observable market data and so reveal actual investor behavior and avoid the reactivity biases of, e.g., surveys. Convenient side effects are that the values can be determined with comparatively low effort and a high degree of reliability. Furthermore, options are typically traded by sophisticated investors and institutional traders. Our proxies should therefore reflect the opinion of investors with a comparatively high market impact.

---

[7]See Figlewski (1989), Bollen and Whaley (2004), Ofek, Richardson, and Whitelaw (2004), and the references therein.

[8]For similar evidence, see Bakshi, Kapadia, and Madan (2003), Rehman and Vilkov (2012), and the references therein.

14

3.3 The Construction of the Sentiment Index

We construct our index based on six proxies introduced above: changes in implied volatility (ΔVIX_t) and skewness $(\Delta SKEW_t)$, determined by first differences in CBOE's VIX and SKEW indexes, changes in closed-end fund discounts $(\Delta cefd_t)$, first-day returns of IPOs (r_t^{IPO}), changes in trading volume $(\Delta turnover_t)$, and changes in the dividend premium $(\Delta pdnd_t)$.

Our next task is to reduce the dimension of the estimation problem. Thereby, our aim is to extract as much information as possible from the sentiment proxies. A very popular method is Principal Component Analysis (PCA), which has been implemented in several well-known studies (see, e.g., Stock and Watson (2002), Baker and Wurgler (2006), and Ludvigson and Ng (2009)). However, its factor loadings are not necessarily related to the informational content of the sentiment proxies for commodity returns. In other words, the principal component constructed to capture the maximum common variations in the proxies may not necessarily contain the information that is relevant for the market investigated. Given the idiosyncratic property of the commodity futures markets compared to equities (Erb and Harvey (2006)), investors in commodity markets are expected to react differently to sentiment measure changes than would those in other markets. As such, an index that simply captures the maximum variations in the proxies may not take into account the special market characteristics of commodities.

To extract information from our proxies more efficiently, we use a Partial Least Squares regression (PLS). PLS is similar to PCA in the sense that it produces uncorrelated components by combining the factors linearly. Its advantage compared to the latter is that the factor loadings (which are constant in our case) are determined so as to maximize the covariance of the proxies with the endogenous variable.[9] A discussion on the advantages of PLS can also be found in Kelly and Pruitt (2013) and Huang, Jiang, Tu, and Zhou (2014).

---

[9]Empirical work by Groen and Pesenti (2010) suggests that PLS performs particularly well, especially when the factor structure is weak.

The PLS algorithm is based on Hastie, Tibshirani, and Friedman (2009). We take the first PLS component as the sentiment index for the whole commodity market. As a robustness check, we also employ PCA and multivariate regressions in the empirical part.

Our sentiment index $SENT$ is computed as the first PLS component of the six selected variables. We present below the loadings of our composite sentiment index for the complete commodity market.

$$SENT_t = -3.802 \cdot \Delta VIX_t + 0.730 \cdot \Delta SKEW_t$$

$$-1.513 \cdot \Delta turnover_t + 0.798 \cdot r_t^{IPO} - 1.003 \cdot \Delta cefd_t - 1.188 \cdot \Delta pdnd_t.$$

Each sentiment proxy enters the equation with the expected sign. For illustration purpose, we recalculate the index with the levels instead of taking differences in the time series of the six proxy variables. Figure 2 exhibits the composite index for the complete commodity market. We observe that the sentiment index drops dramatically during the Russian crisis at the end of 1998 and peak during the internet bubble at the end of 2000. After remaining at a high level between 2004 and 2007, the index reaches its bottom at the end of 2008 in the wake of the subprime crisis. It dips deeply again in the recent sovereign debt crisis.

4 Data Description

We obtain monthly and daily settlement prices of commodity futures from Thomson Datastream. All selected contracts are traded on U.S. exchanges. As in Gorton and Rouwenhorst (2006), we construct rolling commodity futures excess returns by selecting the nearest to maturity contract that will not mature during the next month. Returns are defined as the first difference between the log prices at t and $t-1$. Commodity group returns are defined as the equally-weighted average of each commodity return. Our sample period spans from February 1996 to December 2013.

In total, we have eight commodity groups under investigation. Following Bodie and Rosansky (1980) and Gorton and Rouwenhorst (2006), seven of them are formed based on their natural characteristics: animal products (animal), energy, grains, industrial materials (industrials), industrial metals (metals), precious metals (precious), and softs. A detailed description of the group composition can be found in Appendix A. We then take the average return of the seven groups to represent the commodity market in general. All returns time series are of equal length. The group formation diversifies the idiosyncratic risk of individual commodities to some extent.

We follow Fama and French (2001) and Baker and Wurgler (2004b) in computing the daily and monthly dividend premium and closed-end fund discounts. The details of the construction are provided in Appendix B. Furthermore, we collect the average first-day returns of IPOs from Jay Ritter's webpage. The trading volume is proxied by NYSE Turnover and is taken from *NYSE Fact Book*.

5   Empirical Evidence

This section presents our empirical results. In Section 5.1, we provide univariate analyses to illustrate the individual effects of our sentiment proxies. Then, we investigate the explanatory power of our index for commodity futures returns. Section 5.2 implements the cross-sectional analysis and explores the sentiment exposure of commodities with different characteristics. In Section 5.3, we ascertain the co-movement of commodity returns with sentiment and other related sources by taking into account the time-varying changes. Section 5.4 conducts robustness checks and provides results on the incremental explanatory power of sentiment compared to the commodity-related factors. In Section 5.5, we show the predictive ability of sentiment for future commodity returns on a daily frequency level.

17

5.1 Sentiment in Commodity Returns

5.1.1 Univariate Analyses

This section provides univariate analyses by regressing commodity futures returns on the six sentiment proxies: the VIX and the SKEW indexes, NYSE turnover, IPO returns, closed-end fund discounts, and the dividend premium. We examine both the current and the lagged proxies. The results are outlined in Table 1.

[**Please insert Table 1 around here**]

Consistent with our expectations, changes in the VIX enter with a negative sign in all commodity groups. This suggests that investor uncertainty is priced negatively in commodity returns. Its R^2 values range from 0.006 (animal) to 0.181 (commodity).

In accordance with the intuition outlined previously, implied skewness contains predictive power for commodity returns. Its lagged values exert a positive impact on the majority of the eight commodity groups. An increase in skewness by one standard deviation leads to an increase in the return of 0.34 percentage point for the whole commodity market and 0.57 percentage point for metals.[10] Changes in the dividend premium and closed-end fund discounts are priced negatively for all commodity groups, except animal. These findings are in line with the results of Baker and Wurgler (2006). NYSE turnover growth is negatively related to commodity returns. This result supports the conclusion of Jones (2001), in which high turnover is typically associated with low market returns.

---

[10]We also examine implied kurtosis. Due to its high correlation with skewness (-0.89), we do not find significant additional predictive power for implied skewness.

5.1.2 The Dependence between Investor Sentiment and Commodity Futures Returns

In the sequel, we investigate the explanatory power of our investment indexes ($SENT$) for commodity futures returns. We estimate the regression

$$r_{i,t} = \hat{\alpha}_i + \hat{\beta}_i \cdot SENT_{i,t} + \tilde{\varepsilon}_{i,t} \tag{10}$$

for each commodity group i. The results are shown in Panel A of Table 2. Our estimates convey that the composite sentiment factor is statistically highly significant at the 1% level for each of the commodity groups.

[Please insert Table 2 around here]

A rise in the sentiment index by one standard deviation leads to a rise in the expected commodity return by 31 to 46 basis points per month. As such, the effect is also economically significant. Sentiment exhibits a strong explanatory power with R^2 values of up to 17.9%.

To investigate the robustness of our findings to the index construction method, we report the results from the Principal Component Analyses and the F-test and R^2 from the multivariate regressions in Panels B and C of Table 2, respectively. Naturally, the highest R^2 values are observed for multivariate regressions. It is worth mentioning, however, that with merely one factor in the PLS regression, the amount of variation captured by the sentiment index is of comparable magnitude. The sentiment index constructed by PCA is also statistically highly significant for all groups except animal, albeit with lower explanatory power. In general, the sentiment effect is robust to the methods discussed above.

5.1.3 Controlling for Macroeconomic Influences

In a further step, we control for possible macroeconomic effects on our indexes that might potentially increase their correlation with commodity futures returns. We regress each of

the sentiment proxies on the macro-factors employed in Baker and Wurgler (2006) (data sources are given in parentheses): growth in the industrial production index (Federal Reserve Statistical Release), growth in consumer durables, nondurables, and services (BEA National Income Accounts Table), and a dummy variable for NBER recessions. Furthermore, we consider five additional control variables that are of potential importance for commodity markets: the trade-weighted US Dollar index to major currencies (Federal Reserve), the inflation rate (OECD), the US unemployment rate (OECD), term spread (Federal Reserve), and the default spread (Federal Reserve). We take the residuals from these regressions and use them as "clean" proxies for the construction of our investor sentiment index. Table 3 shows our results for the clean composite index ($SENT_t^{\perp}$).

[**Please insert Table 3 around here**]

Again, we provide the corresponding results for multivariate regressions and principal component analyses. The impact of our index remains highly significant. Comparing the results with those in Table 2, we find that despite a slight reduction in the magnitude, the factor loadings remain approximately on the same scale. Examination of the other two index construction methods yield similar results. Thus, the composite sentiment index is robust to controlling for macro effects. This suggests that the explanatory power of our sentiment measure does not mainly derive from correlations with macroeconomic factors.

5.2 The Asymmetric Sentiment Effect and the Cross-Section of Commodity Futures Returns

In this section, we examine whether some commodity groups, after sorting by different characteristics, are more subject to the sentiment effect than the others. To carry out these tests, we first examine the positive and negative sentiment impact before sorting commodity futures portfolios based on their characteristics.

5.2.1 The Asymmetric Sentiment Effect on Commodity Futures Returns

Previous studies in the academic literature have shown that asset returns tend to be more sensitive to downward than upward sentiment changes (see, e.g., Lee, Jiang, and Indro (2002)). This section investigates whether this asymmetric effect is also present in commodity futures returns. We split our return time series into subsamples of non-negative ($SENT_t \geq 0$) and negative sentiment changes ($SENT_t < 0$). Then, we regress the commodity returns on our sentiment index separately for both subsamples. The results are given in Panels A.1 and A.2 of Table 4.

[Please insert Table 4 around here]

All commodity groups exhibit a stronger sentiment exposure to the negative shifts, both in terms of statistical magnitude and economic significance. For instance, a decrease in sentiment by one standard deviation is associated with a 2.56% loss in the expected commodity return. The same extent of increase is associated with a mere additional return of 1.66%. This finding is robust after controlling for the macro-factors. The results are given in Panels B.1 and B.2.

5.2.2 Commodity Portfolios Sorted by Different Characteristics

Having shown the presence and variation of sentiment premia in commodity returns across time and in bear–bull situations, we are now interested in looking at commodity portfolios sorted by different characteristics, such as futures basis, volatility, momentum, and open interest. The portfolio sorting approach has been adopted in recent studies on commodity markets (e.g., Gorton, Hayashi, and Rouwenhorst (2013) and Szymanowska, de Roon, Nijman, and van den Goorbergh (forthcoming)). Specifically, we sort the 22 commodities into four portfolios based on their characteristics listed above from $t - 12$ to $t - 1$ at the end of every month $t - 1$. Then we calculate their portfolio return by distinguishing whether

21

the change of the clean sentiment index at t is the positive or negative. For each sort, we calculate the annualized mean returns of the portfolios conditioned on the positive and negative sentiment and report them in the first and second row in each panel of Table 5. In the third row we show the return difference between the two cases. The return differences show how sensitive the portfolio is to changes in sentiment. In addition, we also present the difference between returns of the portfolio in the highest quartile and the one in the lowest quartile.

Table 5 presents the results. It displays distinct patterns of portfolio returns resulting both from the sorts and from the sign of the market sentiment. When market sentiment is positive, low basis portfolios earn lower annualized mean returns (8.95%) than high basis portfolios (11.68%). However, when sentiment is negative, the portfolio with the lowest basis bears a yearly loss of -24.91%, whereas portfolios from the second until fourth basis bin obtain returns between -4.09% and -6.92%. The difference between the portfolios during optimistic periods is not significant, while low basis portfolio earn significantly lower returns than high basis portfolio in the bearish periods. The return difference between high and low sentiment is the highest for the low basis portfolio, showing that this portfolio is extremely sensitive to change of sentiment. This evidence corresponds to our theoretical model inference that in situations of backwardation, the sentiment impact is the highest.

High volatility commodity futures yield high return when investors are bullish and low returns when investors are bearish. The portfolio with the highest volatility earns an annualized mean return of 16.94% when sentiment is positive, 2.77% higher than the one with the lowest volatility. Conversely, during negative sentiment periods, the high volatility portfolio loses 6.77%, 11.42 pencentage points more than the low volatility portfolio. High volatility portfolios are hence more sensitive to the sign of sentiment than low volatility portfolios. This evidence supports the conclusion of Barberis and Xiong (2012) that investors tend to invest in high volatility assets for a greater chance of realizing gains.

The high momentum portfolio earns higher returns when sentiment is negative than when sentiment is positive. Compared to high momentum porfolios, low momentum portfoilios are more sensitive to change in sentiment. Commodity futures with lower open interest growth are more exposed to sentiment than those with higher open interest. The portfolio with the lowest open interest yield a return of 15.43%, 2.65 percentage points less than the portfolio with the highest open interest during bullish periods. This difference becomes more dramatic during the bearish periods (9.12 percentage points). If open interest is comparative to "size" in the equity market, this finding is similar to the inference of Baker and Wurgler (2006) that small stocks are more prone to sentiment than large stocks.

5.3 The Time-Series Analysis

5.3.1 Common Driving Forces for the Co-Movement of Commodity Markets

As argued in Pindyck and Rotemberg (1990), the correlations between commodity returns should be driven by common systematic risk factors, such as macroeconomic influences. Testing this hypothesis, their results showed that macroeconomic factors could not explain all common variations among the diverse commodity future returns. The authors attribute this evidence to a potential sentiment impact. Hence, we test the hypothesis whether investor sentiment is the missing common factor to explain the high co-movement among seemingly unrelated commodity returns. Doing this, we closely follow the methodology of Pindyck and Rotemberg (1990). We regress commodity group returns on industrial production levels (IP, OECD), inflation rates (CPI, OECD), the trade-weighted US Dollar index to major currencies ($Dollar$, Federal Reserve), monetary basis M3 ($M3$, OECD), 3-month treasury rates (SR, Federal Reserve). In addition, as recent studies (e.g., Bakshi, Panayotov, and Skoulakis (2011)) have shed light on the importance of the Baltic Dry Index in explaining commodity returns[11], we include the BDI index (BDI, Bloomberg) in our

---

[11]Bakshi, Panayotov, and Skoulakis (2011) documents the predictive ability of the BDI index in pricing stock and commodity returns. However, we find that the contemporaneous relation between the index and

regression specification, resulting in

$$r_{i,t} = \hat{\alpha}_i + \hat{\beta}_{i,1} \cdot IP_t + \hat{\beta}_{i,2} \cdot CPI_t + \hat{\beta}_{i,3} \cdot Dollar_t + \hat{\beta}_{i,4} \cdot SR_t + \hat{\beta}_{i,5} \cdot M3_t +$$
$$\hat{\beta}_{i,8} \cdot BDI_t + \hat{\beta}_{i,9} \cdot SENT_t + \tilde{\varepsilon}_{i,t}. \tag{11}$$

[Please insert Table 6 around here]

The results are given in the upper panel of Table 6. Comparable to our previous regression specifications, the sentiment coefficient is statistically significant at the 5% level for all commodities except for energy. After controlling the eight groups of commodity returns by the factors mentioned above, the residuals should have no significant correlation, i.e., their correlation matrix should resemble the identity matrix. Following (Pindyck and Rotemberg 1990), we implement a likelihood ratio test on the correlation matrix of the errors from the regressions. Specifically, we test the null hypothesis whether the correlation matrix of the return errors is equal to an identity matrix. According to Morrison (1967), the likelihood ratio of the restricted and unrestricted functions is

$$\lambda = |Corr|^{T/2},$$

where $|Corr|$ is the determinant of the correlation matrix. Our test statistic is hence $-2log\lambda$, which has a χ^2 distribution with $(1/2)N(N-1)$ degrees of freedom, where N is the number of commodities.

We report the Likelihood ratio test statistics in the bottom panel of Table 6. The sentiment factor alone reduces the χ^2-statistic from 209 (unrestricted) to 178 (restricted), which is substantial and comparable with macro factors (180). With macro factors together, the likelihood ratio is reduced to 15, though the correlation matrix is still significantly different from an identity matrix. Hence, sentiment contributes to the explanation of commodty return comovement.

---

the commodity returns is stronger than a lagged relation.

5.3.2 Has the Sentiment Effect Increased during the Past Decade?

As outlined in the introduction, several researchers have observed a dramatically growing participation of financial investors in commodity markets during the past decade. As such, market sentiment is expected to play an increasing role in commodity futures returns. The following part aims to shed light on this hypothesis. Specifically, we split our sample into three subperiods: many studies have documented the growing financialization in commodity market since the second half of 2003. Therefore, we define our first break point at July 2003.[12] To take into account the financial crisis in 2008, we take September 2008 as a second break point. Our choices result in three subperiods: 02/1996 to 06/2003 (pre), 07/2003 to 08/2008 (middle), and 09/2008 to 12/2013 (post). For each one, we compute the sentiment index of the composite commodity market and factor loadings with PLS regressions. A likelihood ratio test is then implemented to examine changes in the sample sensitivity of the subperiods:

$$-2[LLF_{full} - (LLF_{pre} + LLF_{middle} + LLF_{post})] \sim \chi^2(df),$$

where $LLF_{full}, LLF_{pre}, LLF_{middle}, LLF_{post}$ denote the log-likelihood values of the full sample, as well as the three subperiods, respectively. The number of degrees of freedom, df, is 16. Panel A of Table 7 shows the likelihood ratio test results with the composite index as explanatory variable. Panels B shows the corresponding results for our clean index.

[**Please insert Table 7 around here**]

The chi-square statistics are highly significant at the 1% level, even after controlling for systematic risk exposure. A rise of one standard deviation in the clean sentiment index increases expected commodity returns by 1.29% and 1.60% in the 07/2003–09/2008 and the 09/2008–12/2013 period respectively — an increase of 0.55% and 0.86% compared to the 02/1996–06/2003 period. These facts confirm the hypothesis that the role of sentiment in

---

[12]This choice is also confirmed by a stochastic break point test.

commodity futures returns has been more accentuated after 2003.

5.4 Further Robustness Checks

5.4.1 A Comparison with the Sentiment Index of Baker and Wurgler (2007) and the Consumer Confidence Index

We explore the relation of our index to indexes that are well established in the academic literature: the Baker and Wurgler (2007) index (BW) and the US consumer confidence index (CCI).[13] To compare the indexes on a common basis, we employ the first principal component of our sentiment proxies, as BW is constructed by PCA. Figure 3 compares our commodity sentiment index with the other two indexes. Their co-movements are quite high, with slight lead–lag effects in various subperiods.

[**Please insert Figure 3 around here**]

To examine the relation more precisely, we conduct an autoregressive analysis by regressing our sentiment index on two leads, two lags, and the contemporaneous values of the other two indexes, i.e.,

$$SENT_t = \hat{\alpha} + \hat{\beta}_{lag2} \cdot BW_{t-2} + \hat{\beta}_{lag1} \cdot BW_{t-1} + \hat{\beta}_0 \cdot BW_t + \hat{\beta}_{lead1} \cdot BW_{t+1} + \hat{\beta}_{lead2} \cdot BW_{t+2} + \tilde{\varepsilon}_t.$$

Similarly, we apply the same equation to the consumer confidence index.

$$SENT_t = \hat{\alpha} + \hat{\beta}_{lag2} \cdot CCI_{t-2} + \hat{\beta}_{lag1} \cdot CCI_{t-1} + \hat{\beta}_0 \cdot CCI_t + \hat{\beta}_{lead1} \cdot CCI_{t+1} + \hat{\beta}_{lead2} \cdot CCI_{t+2} + \tilde{\varepsilon}_t.$$

where BW and CCI denote the values of the Baker and Wurgler (2007) index and the Consumer Confidence index respectively.

---

[13]The Baker and Wurgler (2007) index is available at Jeffrey Wurgler's website. The US consumer confidence index is retrieved from Bloomberg.

Our index is both positively correlated with the BW index and the CCI index and the correlation coefficients are statistically significant. The contemporaneous relation between our index and BW dominates the lead–lag effects ($\hat{\beta}_0$=0.216, t-stat=4.019). Interestingly, our index leads the CCI index by approximately one lag ($\hat{\beta}_{lead1}$=0.272, t-stat=2.743). Summarizing, our index is closely related with the two indexes that are regarded as sentiment measures in the literature, but it also contains independent sources of sentiment information.

5.4.2 Controlling for Stock Returns

As we employ in our study the moments of the implied equity index return distribution as sentiment proxies, we analyze whether the explanatory power of our index stems from the correlation between commodity futures and stock market returns.

To get a first impression, we regress the commodity futures returns on S&P 500 index returns to display the direct correlation between the two markets.[14] Panel A of Table 8 shows the resulting coefficients and the t-statistics. They are statistically highly significant for five of eight commodity groups. The coefficients range between 0.13 (animal) and 1.762 (metals). This result reveals a quite substantial link between commodity futures and equity returns.

[**Please insert Table 8 around here**]

To uncover whether the sentiment index has explanatory power beyond this relation, we orthogonalize the clean sentiment index ($SENT_t^{\perp}$) to S&P 500 stock index returns. Then, we regress the commodity futures returns on the resulting index values ($SENT_t^{\perp\sharp}$), i.e.,

$$r_{i,t} = \hat{\alpha}_i + \hat{\beta}_i SENT_{i,t}^{\perp\sharp} + \tilde{\varepsilon}_t. \tag{12}$$

The results are shown in Panel B of Table 8. Compared to our findings in Table 2, the

---

[14]Our results are robust to using the trailing 1 -year stock returns, which is often examined in the literature (Duffie, Saita, and Wang (2007)).

sentiment coefficient remains statistically significant at the 5% level despite of a slight reduction in the magnitudes and statistical significances. In general, sentiment remains a robust explanatory variable even after controlling for equity returns. This evidence suggests that sentiment provides an additional source of premia.

5.4.3 A Horse Race Test with Commodity-Related Factors

This section investigates whether investor sentiment has incremental explanatory power beyond commodity-related factors. The commodity-related factors we consider are (i) open interest growth of commodity futures contracts, (ii) momentum, and (iii) the futures basis. As already outlined, these factors are well established in the financial literature.

We compute the monthly open interest as the product of the number of outstanding contracts, the contract size and the futures settlement price at the end of each month. Subsequently, we take the log difference of open interest in our analysis. Our momentum factor at t is defined as the cumulative raw return for the period from 9 $(t-9)$ through 1 $(t-1)$ months prior to the observation return.[15] Following, e.g., Fama and French (1987), we calculate the futures basis as

$$\frac{(P_t^{T_2} - P_t^{T_1})}{P_t^{T_1}} \cdot \frac{360}{(T_2 - T_1)}, \tag{13}$$

where P_t^T indicates time-t futures prices with maturity T. T_1 denotes the time to maturity of the nearest-to-maturity contract. T_2 indicates the time to maturity of the subsequent contract, i.e., $T_2 > T_1$.

Our list of commodity-related variables may not be exhaustive. However, we view these variables as representative of a bullish or bearish mood on the commodities market besides the macroeconomic risks that are incorporated in these factors. To investigate the impact

---

[15]We compared the explanatory power for cumulated returns of 1–12 months lagged 1 month, $\sum_{k=1}^{K} r_{t-k}, K = 1, ..., 12$, individually for the commodity returns. As a result, the 9-month return lagged by 1 month was adopted, due to its having the strongest explanatory power among the commodities.

of our factors, we pool the commodities in each group and regress the commodity futures returns on the basis, open interest, and momentum, i.e.,

$$r_{i,t} = \hat{\alpha}_i + \hat{\beta}_i^{open} \cdot Open_{i,t} + \hat{\beta}_i^{bas} \cdot Basis_{i,t-1} + \hat{\beta}_i^{mom} \cdot Momentum_{i,t} + \tilde{\varepsilon}_{i,t}, \qquad (14)$$

where $Open_{i,t}$ $Basis_{i,t}$, and $Momentum_{i,t}$ denote open interest growth, futures basis, and momentum of group i, respectively. The results are given in Panel A of Table 9.

[**Please insert Table 9 around here**]

Open interest is highly significant for all commodity groups. Similar results are found for the futures basis. In the univariate regressions (unreported), momentum contains highly significant predictive power for commodity returns. After adding the lagged basis, however, the significance disappears. To investigate the incremental explanatory power of investor sentiment, we add our index to the regression specification. The results are outlined in Panel B. They display a dramatic increase in the explanatory power. Moreover, the sentiment factor is significant at a level of 1% for all the examined commodity groups. Summarizing, the evidence above indicates a substantial explanatory power of the sentiment index for commodity returns, even after controlling for macroeconomic effects, equity index returns, and established commodity-related factors.

5.5 Predictive Regressions on a Daily Frequency Level

An investor's mood can change very quickly and varies substantially from day to day. As most of our proxies can be measured with a very high frequency, this advantage allows for a timely measure of investor sentiment. This section examines whether investor sentiment has predictive power for commodity futures returns. We extend the CEFD and dividend premium to a daily level, as described in Appendix B. We investigate the relation between our sentiment index and commodity futures returns at daily frequencies. Interestingly, we

29

find a salient lagged impact of the daily sentiment proxies on the commodity futures returns. Specifically, we estimate the following regression

$$r_{i,t} = \hat{\alpha}_i + \hat{\beta}_i \cdot SENT_{i,t-1} + \hat{\zeta}_i \cdot r_{i,t-1} + \tilde{\varepsilon}_{i,t}. \tag{15}$$

where $SENT_{i,t-1}$ is determined by our sentiment proxies, except the first-day returns of IPOs and NYSE turnover, which are both only available at monthly frequencies. The results of these regressions are given in Table 10.

[**Please insert Table 10 around here**]

Our results reveal a statistically highly significant predictive relation between the daily sentiment index and commodity futures returns. A rise in the sentiment index by one standard deviation leads to a rise in the expected commodity return by 2 to 3 basis points per day, which is substantial for daily return predictions.

6   Conclusion

This study highlights the existence of sentiment exposure in commodity futures returns. We provide a noisy rational expectations equilibrium model to show that financial investors provide an additional risk premium in the existence of short-sale constraints.

Through a composite sentiment index, characterized by higher moments of the implied equity index return distribution, as well as sentiment proxies already established in the academic literature, we discover a strong impact of investor sentiment. This result remains significant after controlling for macro-economic influences, equity market returns, as well as commodity related factors, such as open interest, futures basis, and momentum. Moreover, our results are robust to different methods of constructing the sentiment index. Following the trend of increasing financial investment in commodity futures, the role of sentiment in

30

commodity returns became more pronounced during the recent decade. Consistent with our model inference, that financial investors offer an additional risk premium, our empirical results demonstrate that commodity futures with low open interest, high volatility, low momentum, or low basis are more likely to be exposed to sentiment. In addition, we find a predictive power of investor sentiment for commodity futures returns at a daily frequency level.

Table 1: Univariate Regression Results

This table gives the results of the univariate regression analyses for commodity futures returns on the individual sentiment candidate proxies, i.e.,

$$r_t = \hat{\alpha} + \hat{\beta} \cdot proxy_t + \tilde{\varepsilon}_t,$$

where $proxy_t$ denotes the time-t values of our sentiment proxies. They are changes in the VIX index (ΔVIX_t) and changes in CBOE's SKEW index ($\Delta SKEW_t$). In addition, we examine changes in the NYSE turnover ($\Delta turnover_t$), the average of first-day IPO returns (r_t^{IPO}), the closed-end fund discount ($\Delta cefd_t$), and the dividend premium ($\Delta pdnd_t$). Our analysis treats seven commodity groups (animal, energy, grains, industrials, metals, precious, and softs), as well as an equally-weighted index representing the commodity market (commodity). The sample contains monthly data and stretches from February 1996 to December 2013. The t-statistics calculated with Newey–West standard errors are shown in square brackets.

	commodity	animal	energy	grains	indus	metals	precious	softs
ΔVIX_t	-0.0177	-0.0036	-0.0165	-0.0202	-0.0207	-0.0291	-0.0177	-0.0161
	[-4.283]	[-1.005]	[-2.159]	[-3.959]	[-4.547]	[-3.907]	[-3.659]	[-4.540]
R^2	0.181	0.006	0.033	0.089	0.093	0.128	0.083	0.082
$\Delta SKEW_t$	0.0034	-0.0038	0.0023	0.0058	0.0034	0.0057	0.0071	0.0033
	[0.919]	[-1.227]	[0.362]	[1.303]	[0.770]	[0.819]	[1.540]	[0.842]
R^2	0.007	0.007	0.001	0.007	0.002	0.005	0.013	0.003
$\Delta turnover_t$	-0.0070	0.0020	-0.0102	-0.0096	-0.0131	-0.0018	-0.0069	-0.0094
	[-2.142]	[0.652]	[-1.854]	[-1.669]	[-2.652]	[-0.329]	[-1.362]	[-2.531]
R^2	0.029	0.002	0.013	0.020	0.038	0.001	0.013	0.028
r_t^{IPO}	0.0037	0.0059	0.0145	0.0008	0.0008	0.0037	0.0078	-0.0075
	[1.207]	[2.248]	[2.176]	[0.223]	[0.236]	[0.633]	[1.801]	[-2.034]
R^2	0.008	0.018	0.026	0.000	0.000	0.002	0.016	0.018
$\Delta cefd_t$	-0.0047	0.0007	-0.0041	-0.0043	-0.0036	-0.0103	-0.0040	-0.0071
	[-1.597]	[0.263]	[-0.814]	[-0.829]	[-0.774]	[-2.205]	[-0.900]	[-2.232]
R^2	0.013	0.000	0.002	0.004	0.003	0.016	0.004	0.016
$\Delta pdnd_t$	-0.0055	-0.0003	-0.0133	-0.0021	-0.0054	-0.0027	-0.0102	-0.0047
	[-2.571]	[-0.105]	[-3.009]	[-0.634]	[-1.336]	[-0.523]	[-3.184]	[-1.060]
R^2	0.018	0.000	0.022	0.001	0.006	0.001	0.028	0.007

Table 2: The Dependence between Investor Sentiment and Commodity Returns

This table illustrates the dependence between investor sentiment and commodity returns. For each commodity group i, Panel A shows the univariate regression results of monthly commodity returns ($r_{i,t}$) on the composite sentiment factor ($SENT_{i,t}$) constructed by partial least squares regressions (PLS), i.e.,

$$r_{i,t} = \hat{\alpha}_i + \hat{\beta}_i \cdot SENT_{i,t} + \tilde{\varepsilon}_{i,t}.$$

As a robustness check, we follow Baker and Wurgler (2007) and conduct a principal component analysis (PCA) on our seven sentiment proxies. Panel B.1. shows the regression results of commodity futures returns on the first principal component. Panel B.2 shows the results in the multivariate regressions, in which we regress commodity futures returns on the six sentiment proxies. We report the F-test statistics on the null hypothesis that the coefficients are equal to zero and the R^2. The t-statistics calculated with Newey–West standard errors are in square brackets. *, **, and *** denote significance at 10%, 5%, and 1% level, respectively.

	commodity	animal	energy	grains	indus	metals	precious	softs
Panel A: PLS Analysis								
$\hat{\beta}$	0.0033	0.0046	0.0031	0.0034	0.0032	0.0039	0.0032	0.0035
	[6.804]	[2.695]	[3.833]	[4.552]	[4.843]	[5.558]	[5.041]	[5.311]
R^2	0.179	0.033	0.065	0.089	0.099	0.127	0.107	0.117
Panel B: Robustness Check: Alternative Methods								
Panel B.1: PCA Analysis								
$\hat{\beta}$	0.0128	0.0022	0.0183	0.0128	0.0148	0.0159	0.0156	0.0099
	[4.021]	[0.793]	[4.086]	[2.728]	[3.589]	[2.668]	[4.426]	[2.912]
R^2	0.143	0.004	0.062	0.054	0.071	0.057	0.097	0.047
Panel B.2: Multivariate Regression Tests								
F-test(6, 217)	8.207***	1.457	2.422**	3.699***	3.972***	6.299***	4.259***	4.892***
R^2	0.191	0.040	0.065	0.096	0.103	0.154	0.109	0.124

Table 3: Controlling for Macroeconomic Factors

This table shows the results of our robustness checks with regard to macroeconomic variables. In the spirit of Baker and Wurgler (2006), we orthogonalize each of the sentiment proxies to eleven macro variables. As before, we estimate the first component of the "clean" $(SENT_t^{\perp})$ by partial least squares regressions (PLS). Panel A shows the regression results of monthly commodity futures returns on the first PLS component, i.e.,

$$r_{i,t} = \hat{\alpha}_i + \hat{\beta}_i \cdot SENT_{i,t}^{\perp} + \tilde{\varepsilon}_{i,t}.$$

Panel B.1 shows the regression results of commodity returns on their first principal component, determined by ordinary PCAs. Panel B.2 shows the results of the multivariate regressions, in which we regress commodity returns on the six "clean" sentiment proxies. We report the F-test statistics on the null hypothesis that all the coefficients are equal to zero and their R^2. The t-statistics calculated by Newey–West standard errors are in square brackets.

	commodity	animal	energy	grains	indus	metals	precious	softs
Panel A: PLS Analysis								
$\hat{\beta}$	0.0032	0.0048	0.0032	0.0034	0.0033	0.0040	0.0031	0.0038
	[5.656]	[2.567]	[3.662]	[4.214]	[4.377]	[4.164]	[4.029]	[4.680]
R^2	0.131	0.030	0.059	0.077	0.083	0.075	0.071	0.093
Panel B: Robustness Check: Alternative Methods								
Panel B.1: PCA Analysis								
$\hat{\beta}$	0.0102	0.0023	0.0159	0.0104	0.0124	0.0109	0.0123	0.0074
	[3.612]	[0.873]	[3.851]	[2.440]	[3.083]	[2.012]	[3.620]	[2.201]
R^2	0.094	0.004	0.048	0.037	0.052	0.028	0.062	0.027
Panel B.2: Multivariate Regression Tests								
F-test(6, 217)	5.608***	1.324	2.294**	3.091***	3.319***	3.606***	2.698**	3.888***
R^2	0.139	0.037	0.062	0.082	0.087	0.094	0.072	0.101

Table 4: Asymmetric Sentiment Effect on Commodity Futures Returns

This table shows the asymmetric effect of sentiment on commodity returns. First, we split our sample into subsamples with positive (≥ 0) and negative (< 0) sentiment index changes. We report the coefficients $\hat{\beta}$, t-statistics, and R^2 values for the regressions

$$r_{i,t} = \hat{\alpha}_{i,1} + \hat{\beta}_{i,1}^+ \cdot SENT_t + \tilde{\varepsilon}_{i,t}, \qquad SENT_t \geq 0,$$

and

$$r_{i,t} = \hat{\alpha}_{i,2} + \hat{\beta}_{i,2}^- \cdot SENT_t + \tilde{\varepsilon}_{i,t}, \qquad SENT_t < 0.$$

Panel A.1 and A.2 show the estimations with the raw sentiment index as explanatory variable, for $SENT_t \geq 0$ and $SENT_t < 0$, respectively. The corresponding values for the time series orthogonalized to macro effects ($SENT^\perp$) are given in Panel B.1 and B.2. The t-statistics calculated with Newey–West standard errors are in square brackets.

	commodity	animal	energy	grains	indus	metals	precious	softs
Panel A.1: Positive Sentiment								
SENT$^+$	0.0166	0.0122	0.0158	0.0372	0.0285	0.0110	0.0142	0.0210
	[2.547]	[2.600]	[1.366]	[5.486]	[3.089]	[0.656]	[1.965]	[2.493]
R^2	0.076	0.063	0.014	0.128	0.060	0.008	0.022	0.058
Panel A.2: Negative Sentiment								
SENT$^-$	0.0256	0.0161	0.0344	0.0194	0.0301	0.0411	0.0268	0.0182
	[3.216]	[1.685]	[3.165]	[1.597]	[4.446]	[3.092]	[2.709]	[2.603]
R^2	0.200	0.037	0.080	0.044	0.122	0.162	0.104	0.056
Panel B.1: Positive "Clean" Sentiment								
SENT$^{\perp+}$	0.0158	0.0073	0.0227	0.0433	0.0196	0.0112	0.0082	0.0193
	[2.157]	[1.459]	[1.822]	[5.527]	[2.296]	[0.724]	[0.959]	[2.094]
R^2	0.058	0.018	0.026	0.153	0.030	0.007	0.008	0.042
Panel B.2: Negative "Clean" Sentiment								
SENT$^{\perp-}$	0.0221	0.0020	0.0288	0.0164	0.0259	0.0431	0.0202	0.0173
	[2.876]	[0.199]	[2.437]	[1.395]	[2.905]	[3.210]	[2.128]	[2.772]
R^2	0.145	0.001	0.051	0.029	0.080	0.159	0.052	0.049

Table 5: The Sentiment Exposure of Portfolios Sorted by Different Characteristics

This table shows the annualized mean returns of the commodity portfolios sorted by their basis, volatility, momentum, and open interest growth. For every month t, we sort portfolios based on the characteristics above from $t-12$ to $t-1$. "Pos" and "Neg" denote the cases when the change of the clean sentiment index at t is positive or negative. "Pos−Neg" stands for the difference between the returns in positive and negative sentiment. Returns are expressed in percentage points. *, **, and *** denote significance of the mean return difference at 10%, 5%, and 1% level, respectively.

		Low	(2)	(3)	High	High − Low
Basis	Pos	8.948	13.829	10.361	11.678	2.730
	Neg	-24.914	-5.703	-4.092	-6.920	17.994***
	Pos − Neg	33.862***	19.532***	14.453***	18.598***	-15.264***
Volatility	Pos	14.163	16.971	17.815	16.938	2.774*
	Neg	4.656	-6.932	-6.848	-6.770	-11.425***
	Pos − Neg	9.508***	23.903***	24.663***	23.707***	14.200***
Momentum	Pos	14.405	12.079	10.063	9.689	-4.716***
	Neg	-9.185	-2.976	-4.351	-6.422	2.762
	Pos − Neg	23.590***	15.055***	14.414***	16.111***	-7.478***
Open Interest	Pos	15.431	17.812	16.399	18.078	2.647*
	Neg	-10.824	-2.290	-0.802	-1.700	9.124***
	Pos − Neg	26.255***	20.102***	17.202***	19.778***	-6.477***

Table 6: Common Driving Forces for the Co-Movement of Commodity Markets

This table shows the correlations among commodity group returns before and after controlling for the sentiment effect and macroeconomic factors. Our choice of macroeconomic factors closely follows Pindyck and Rotemberg (1990). The regression equation estimated is

$$r_{i,t} = \hat{c}_i + \hat{\beta}_{i,1} \cdot IP_t + \hat{\beta}_{i,2} \cdot CPI_t + \hat{\beta}_{i,3} \cdot Dollar_t + \hat{\beta}_{i,4} \cdot SR_t + \hat{\beta}_{i,5} \cdot TS_t$$
$$+ \hat{\beta}_{i,6} \cdot DS_t + \hat{\beta}_{i,7} \cdot r_t^{emg} + \hat{\beta}_{i,8} \cdot BDI_t + \hat{\beta}_{i,9} \cdot SENT_t^{\perp} + \tilde{\varepsilon}_{i,t},$$

where IP, $Dollar$, and SR denote the log difference of industrial production, the average Dollar exchange rate against commodity currencies, and a short rate, respectively. TS and DS are term spread and default spread. BDI denotes the log changes of the Baltic Dry Index. The likelihood ratio test statistics of the correlations among the commodity returns with restricted and unrestricted model are reported in the last column. The t-statistics calculated with Newey–West standard errors are in square brackets.

	commodity	animal	energy	grains	indus	metals	precious	softs
$\hat{c}$	-0.0018	0.0079	0.0002	0.0294	-0.0109	-0.0027	-0.0159	-0.0198
	[-0.121]	[0.503]	[0.006]	[1.115]	[-0.417]	[-0.080]	[-0.736]	[-0.980]
IP	1.5001	1.1262	0.5279	0.7356	2.5141	2.1565	2.5784	0.8844
	[2.050]	[1.189]	[0.268]	[0.609]	[1.661]	[1.086]	[1.453]	[0.863]
CPI	3.2513	3.2140	12.1297	-1.3429	-1.6892	9.8884	2.3361	-1.5863
	[2.100]	[1.581]	[3.303]	[-0.498]	[-0.849]	[2.407]	[0.991]	[-0.667]
Dollar	-1.6163	-0.2624	-2.6499	-1.6090	-0.8867	-2.9991	-2.0941	-0.8642
	[-4.089]	[-0.645]	[-3.534]	[-2.500]	[-1.400]	[-3.692]	[-3.236]	[-1.382]
SR	-0.0006	-0.0019	0.0005	-0.0042	-0.0009	-0.0017	0.0014	0.0024
	[-0.308]	[-0.894]	[0.154]	[-1.075]	[-0.242]	[-0.433]	[0.489]	[0.866]
TS	-0.0002	-0.0003	0.0052	-0.0117	0.0029	-0.0011	-0.0021	0.0055
	[-0.063]	[-0.073]	[0.893]	[-1.611]	[0.562]	[-0.159]	[-0.422]	[1.177]
DS	0.0032	-0.0025	-0.0136	0.0039	0.0077	0.0035	0.0175	0.0056
	[0.347]	[-0.356]	[-0.660]	[0.359]	[0.589]	[0.152]	[1.747]	[0.517]
BDI	0.0424	0.0174	0.0616	0.0495	0.1109	0.0499	0.0308	-0.0412
	[0.925]	[0.490]	[0.957]	[0.906]	[2.426]	[0.575]	[0.457]	[-0.731]
$SENT^{\perp}$	0.0145	0.0076	0.0212	0.0184	0.0182	0.0217	0.0160	0.0176
	[6.000]	[2.548]	[4.483]	[5.165]	[4.079]	[3.635]	[4.713]	[5.343]
R^2	0.276	0.063	0.199	0.128	0.137	0.226	0.160	0.115

Likelihood Ratio Test of Correlation

Unconditional Correlation	209.33
Control for $SENT^{\perp}$	178.28
Control for MACRO	179.90
Control for $SENT^{\perp}$+MACRO	149.93

37

Table 7: The Increasing Role of Investor Sentiment during the Past Decade

This table tests the hypothesis that commodity futures returns have been increasingly exposed to sentiment changes. For this purpose, we split our sample into three subperiods: the first sample spans the period between 02/1996 and 06/2003 (pre); the second sample stretches from 07/2003 to 08/2008 (middle); the third sample covers the period after the financial crisis from 09/2008 to 12/2013 (post). We provide the results of the Partial Least Squares regressions (PLS) for the commodity indexes, as well as likelihood ratio test statistics. The estimation equation for the complete sample is the same as in Table 2. The corresponding equations for the three subsamples are

$$r_{t<07/2003} = \hat{\alpha}^{pre} + \hat{\beta}^{pre} \cdot SENT_{t<07/2003} + \tilde{\varepsilon}_{t<07/2003}.$$

$$r_{07/2003 \leq t \leq 08/2008} = \hat{\alpha}^{middle} + \hat{\beta}^{middle} \cdot SENT_{07/2003 \leq t \leq 08/2008} + \tilde{\varepsilon}_{07/2003 \leq t \leq 08/2008}.$$

$$r_{t>08/2008} = \hat{\alpha}^{post} + \hat{\beta}^{post} \cdot SENT_{t>08/2008} + \tilde{\varepsilon}_{t>08/2008}.$$

We compute the likelihood ratio statistic, comparing the sum of the logarithmic likelihood values of the three subsamples (LLF_{pre}, LLF_{middle}, and LLF_{post}) with the corresponding value for the full sample (LLF_{full}):

$$-2[LLF_{full} - (LLF_{pre} + LLF_{middle} + LLF_{post})] \sim \chi^2(df),$$

where the number of degrees of freedom, df, is 16. The results are given in **Panel A**. **Panel B** shows the corresponding results after controlling for macroeconomic effects and S&P 500 returns ($SENT^{\perp}$). Both sentiment indexes are standardized to have zero mean and unit variance. The t-statistics calculated by Newey–West standard errors are in square brackets. *, **, and *** denote significance at 10%, 5%, and 1% level respectively.

	02/1996 — 06/2003	07/2003 — 08/2008	08/2008 — 12/2013	Complete Sample
Panel A: Likelihood Ratio Test with $SENT$				
$\hat{\beta}$	0.0074	0.0155	0.0088	0.0033
	3.358	2.202	6.292	6.804
R^2	0.115	0.075	0.390	0.179
Likelihood Ratio	41.770***			
Panel B: Likelihood Ratio Test with $SENT^{\perp}$				
$\hat{\beta}$	0.0074	0.0129	0.0160	0.0032
	2.315	2.991	2.138	5.656
R^2	0.058	0.130	0.069	0.131
Likelihood Ratio	50.445***			

Table 8: Controlling for Equity Market Returns

This table shows the results of our robustness check of the sentiment effect to S&P 500 returns. First, we regress the monthly commodity futures returns of each group i on S&P 500 equity index returns r^{sp}, i.e.,

$$r_{i,t} = \hat{\alpha}_i + \hat{\varsigma}_i \cdot r^{sp}_{i,t} + \tilde{\varepsilon}_{i,t}.$$

The results are given in Panel A. Next, we orthogonalize the clean sentiment indexes, i.e., controlled for the eleven macro variables as shown in Table 3, $SENT^{\perp}_{i,t}$, to the S&P 500 returns. The resulting time series are denoted by $SENT^{\perp\sharp}_{i,t}$. In Panel B, we regress the commodity futures returns on $SENT^{\perp\sharp}_{i,t}$, i.e.,

$$r_{i,t} = \hat{\alpha}_i + \hat{\beta}_i \cdot SENT^{\perp\sharp}_{i,t} + \tilde{\varepsilon}_{i,t}.$$

The t-statistics calculated by Newey–West standard errors are in square brackets.

	commodity	animal	energy	grains	indus	metals	precious	softs
Panel A: Regression on SP500 Return								
$\hat{\varsigma}$	0.8764	0.1317	1.0010	0.8479	1.0540	1.7626	0.7611	0.5762
	[3.702]	[0.761]	[2.018]	[2.954]	[4.377]	[4.345]	[2.970]	[2.727]
R^2	0.177	0.003	0.049	0.063	0.096	0.186	0.061	0.042
Panel B: Regression on $SENT^{\perp\sharp}$								
$\hat{\beta}$	0.0037	0.0050	0.0044	0.0037	0.0037	0.0046	0.0039	0.0040
	[2.201]	[2.559]	[2.795]	[2.499]	[2.151]	[2.407]	[2.122]	[3.640]
R^2	0.022	0.030	0.035	0.028	0.021	0.026	0.021	0.059

Table 9: Pooled OLS Regressions with Commodity Related Factors and the Sentiment Index

This table investigates the incremental explanatory power of our sentiment index beyond established commodity-related factors: open interest ($Open_t$), futures basis ($Basis_{t-1}$), and momentum ($Momentum_t$). For each commodity group i, we run a pooled OLS regression of commodity returns on the factors mentioned above. Momentum is computed as the average of the preceding nine months (three quarters) return. $Open_t$ is calculated as the log difference of the open interest. $Basis_t$ is computed as $(P_t^{T_2} - P_t^{T_1})/P_t^{T_1} \cdot (360/(T_2 - T_1))$, where $P_t^{T_1}$ denotes the price level of the nearest-to-maturity contract. Similarly, $P_t^{T_2}$ denotes the corresponding value of the subsequent contract, i.e., $T_2 > T_1$. Panel A shows the results of the pooled OLS regressions with the commodity-related factors as regressors, i.e.,

$$r_{i,t} = \hat{\alpha}_i + \hat{\beta}_i^{open} \cdot Open_{i,t} + \hat{\beta}_i^{bas} \cdot Basis_{i,t-1} + \hat{\beta}_i^{mom} \cdot Momentum_{i,t} + \tilde{\varepsilon}_{i,t}.$$

In Panel B, we add our clean sentiment index ($SENT^{\perp}$) to the specification in Panel A.

$$r_{i,t} = \hat{\alpha}_i + \hat{\beta}_i^{open} \cdot Open_{i,t} + \hat{\beta}_i^{bas} \cdot Basis_{i,t-1} + \hat{\beta}_i^{mom} \cdot Momentum_{i,t} + \hat{\beta}_i^{sent} \cdot SENT_{i,t}^{\perp} + \tilde{\varepsilon}_{i,t}.$$

The t-statistics are in square brackets, calculated with Newey–West standard errors of a pre-specified lag length of 10.

	commodity	animal	energy	grains	indus	metals	precious	softs
Panel A: with the Commodity-Related Factors Only								
$\hat{\beta}_i^{open}$	0.3691	0.2501	0.6045	0.3955	0.2765	0.3625	0.3279	0.4480
	[14.315]	[15.825]	[13.014]	[11.409]	[7.877]	[8.509]	[8.938]	[20.391]
$\hat{\beta}_i^{bas}$	0.0413	0.0635	-0.0048	0.0313	0.0567	-0.0811	-0.0770	0.0539
	[7.299]	[6.871]	[-0.305]	[2.795]	[3.737]	[-2.515]	[-1.263]	[2.754]
$\hat{\beta}_i^{mom}$	0.1222	-0.1302	0.1340	0.0183	-0.0405	0.0600	0.1897	0.2428
	[2.444]	[-0.837]	[1.148]	[0.233]	[-0.405]	[0.467]	[1.997]	[4.374]
R^2	0.479	0.516	0.687	0.495	0.384	0.505	0.465	0.522
Panel B: With SENT$^{\perp}$ and the Commodity-Related Factors								
$\hat{\beta}_i^{open}$	0.3635	0.2488	0.5946	0.3883	0.2673	0.3494	0.3215	0.4414
	[15.767]	[16.276]	[12.510]	[13.074]	[8.440]	[8.712]	[9.059]	[20.991]
$\hat{\beta}_i^{bas}$	0.0421	0.0621	-0.0014	0.0320	0.0596	-0.0836	-0.0549	0.0544
	[7.985]	[6.753]	[-0.093]	[2.786]	[4.164]	[-2.513]	[-1.026]	[2.792]
$\hat{\beta}_i^{mom}$	0.1368	-0.1441	0.1404	0.0286	-0.0070	0.0941	0.2233	0.2317
	[3.211]	[-0.931]	[1.274]	[0.398]	[-0.073]	[0.741]	[2.449]	[3.860]
$\hat{\beta}_i^{SENT^{\perp}}$	0.0021	0.0019	0.0007	0.0026	0.0023	0.0031	0.0023	0.0025
	[6.333]	[2.519]	[1.705]	[5.614]	[4.645]	[4.145]	[5.200]	[4.569]
R^2	0.493	0.518	0.689	0.525	0.408	0.550	0.488	0.537

Table 10: Lagged Sentiment Effect with a Daily Frequency Level

This table shows the lagged effect of sentiment on commodity returns. We show the results of the regression

$$r_{i,t} = \hat{\alpha}_i + \hat{\beta}_i \cdot SENT_{i,t-1} + \hat{\zeta}_i \cdot r_{i,t-1} + \tilde{\varepsilon}_{i,t}.$$

The t-statistics adjusted by Newey-West standard errors are reported in square brackets.

	commodity	animal	energy	grains	indus	metals	precious	softs
$SENT_{t-1}$	0.0002	0.0002	0.0002	0.0003	0.0002	0.0002	0.0002	0.0002
	[4.184]	[2.141]	[2.520]	[2.529]	[2.370]	[4.289]	[5.715]	[3.673]
r_{t-1}	0.0126	0.0176	-0.0265	-0.0073	0.0307	-0.0822	0.0299	0.0181
	[0.677]	[1.061]	[-1.579]	[-0.465]	[1.917]	[-4.097]	[1.611]	[1.188]
R^2	0.009	0.001	0.002	0.002	0.003	0.011	0.015	0.004

Figure 1: Proportion of the Different Types of Traders With Long Positions in the Wheat Futures Market

This figure plots the proportion of the different types of traders on the long side to the total long side traders in the wheat futures market. The trader categories are producers, swap dealers, money managers, and others, as defined by the CFTC. The sample period is from June 2006 to December 2013.
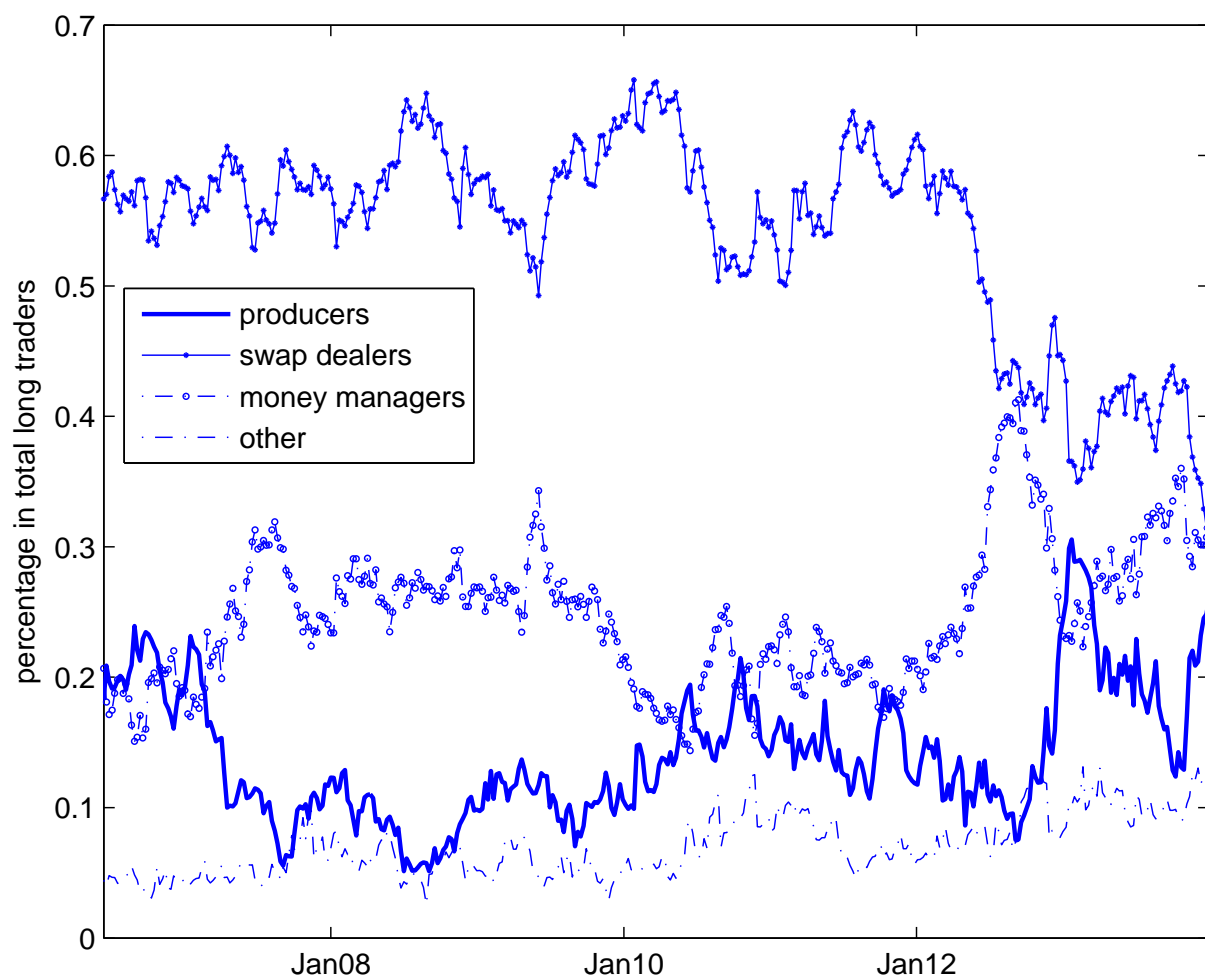
Figure 2: A Plot of Sentiment Index Levels

This figure illustrates the first PLS component of levels and changes in the six sentiment proxies: the VIX and SKEW indexes, NYSE turnover, first-day returns on IPOs, closed-end fund discounts, and dividend yields. The upper panel exhibits the sentiment index based on changes in the proxies. The lower panel displays the index constructed by levels of the sentiment proxies. Both indexes are standardized to have zero mean and unit variance over the sample period from 02/1996 to 12/2013.
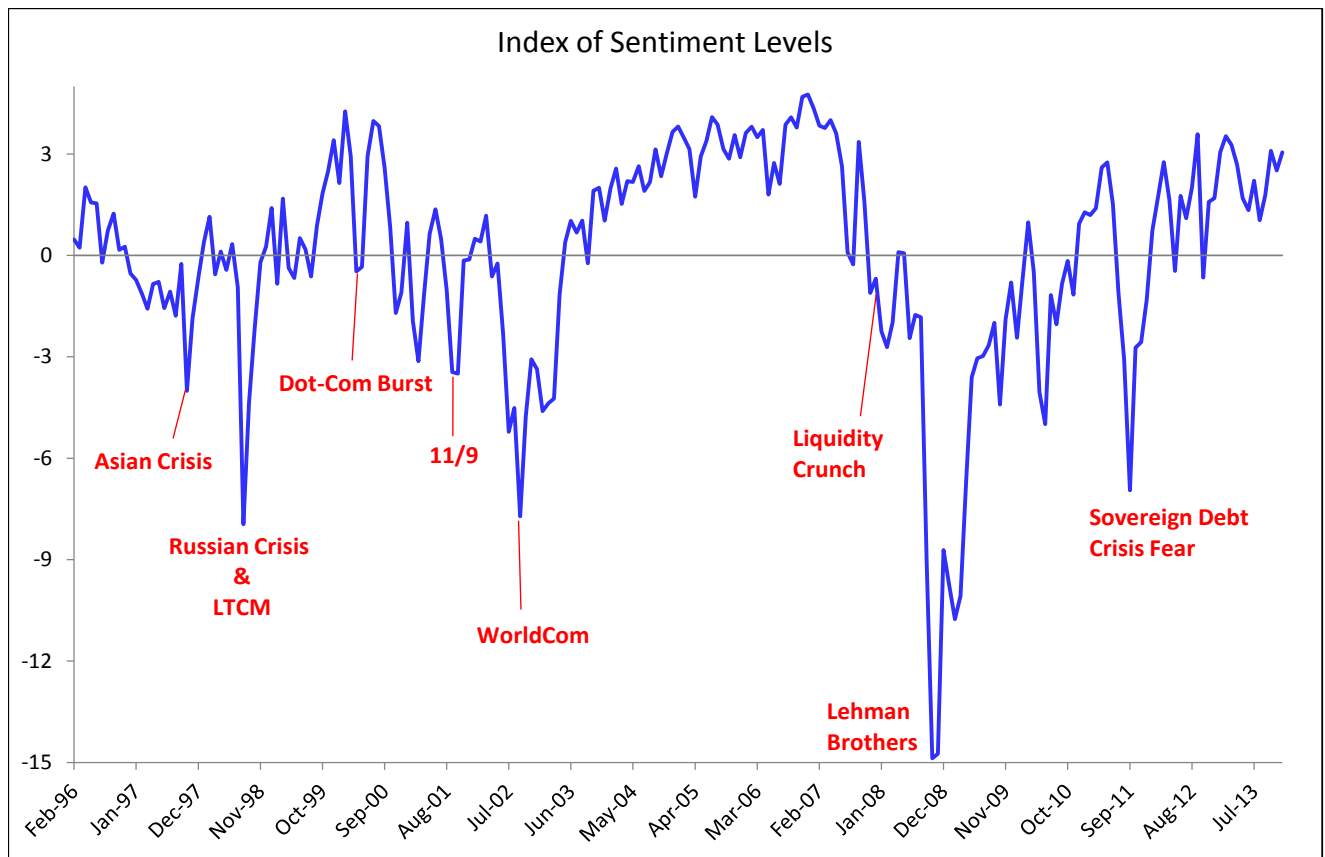

43

Figure 3: A Comparison of the Commodity Sentiment Index with the Baker and Wurgler (2007) Index and the Consumer Confidence Index

This figure illustrates the relation between our commodity sentiment index and two other sentiment indexes—the Baker and Wurgler (2007) sentiment index and the Consumer Confidence Index. For comparison on a common basis, our commodity sentiment index is defined as the first principal component of changes in the six measures of sentiment: changes in option implied variance and skewness, changes in NYSE turnover, average first-day returns of IPOs, and changes in closed-end fund discounts and in the dividend premium. The upper panel compares our index with the Baker and Wurgler (2007) sentiment index. The lower panel contains the corresponding time series for the Consumer Confidence Index. All indexes are standardized to have zero mean and unit variance over the sample period.

Our commodity sentiment index is displayed in a plain solid line. The other two indexes are displayed in dashed lines.

44
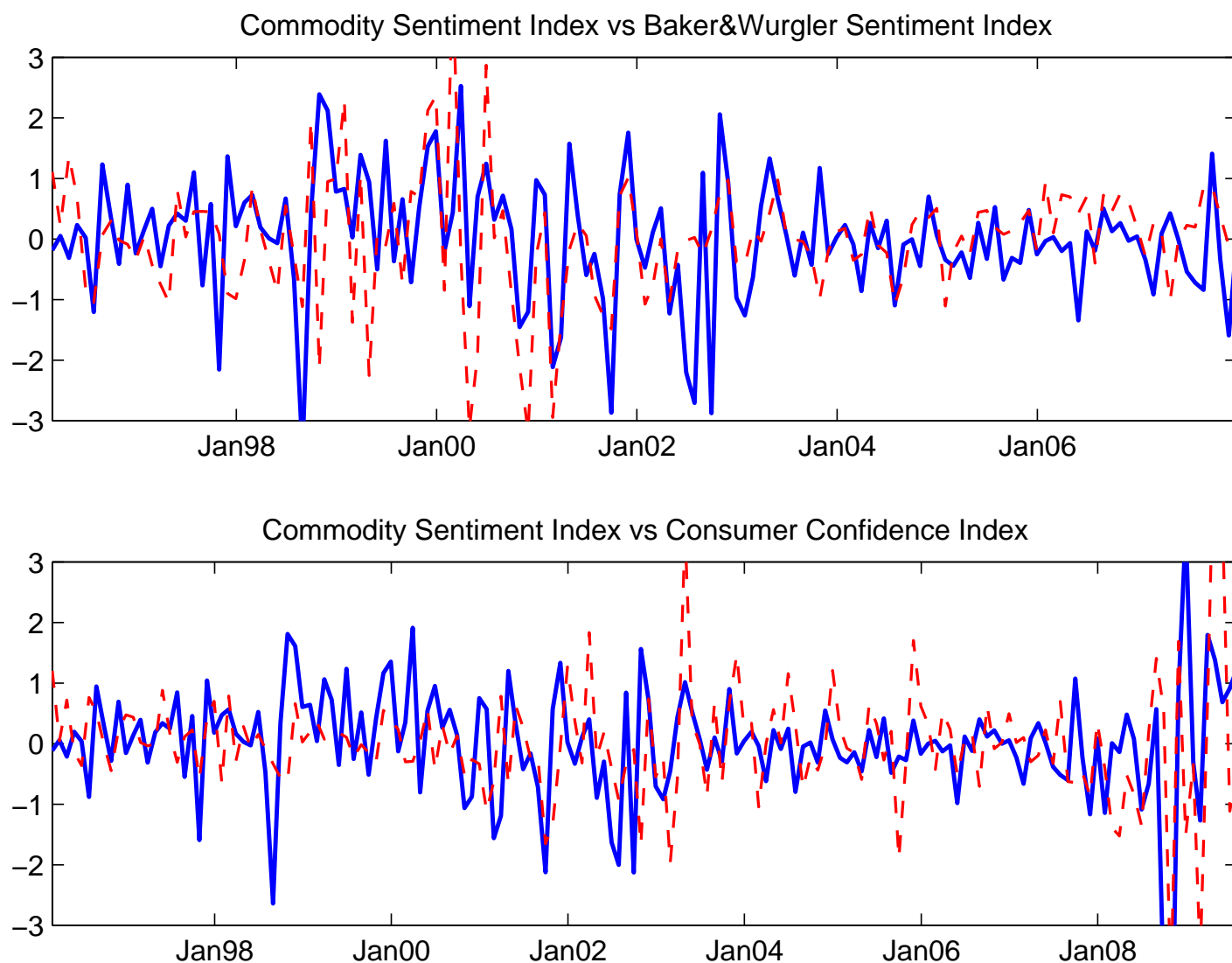
References

ACHARYA, V. V., L. A. LOCHSTOER, AND T. RAMADORAI (2013): "Limits to Arbitrage and Hedging: Evidence from Commodity Markets," *Journal of Finance Economics*, 109(2), 441–465.

BAILEY, W., AND K. C. CHAN (1993): "Macroeconomic Influences and the Variability of the Commodity Futures Basis," *The Journal of Finance*, 48(2), 555–573.

BAKER, M., AND J. WURGLER (2004a): "Appearing and Disappearing Dividends: The Link to Catering Incentives," *The Journal of Financial Economics*, 73(2), 271–288.

——— (2004b): "A Catering Theory of Dividends," *The Journal of Finance*, 59(3), 1125–1165.

——— (2006): "Investor Sentiment and the Cross-Section of Stock Returns," *The Journal of Finance*, 61(4), 1645–1680.

——— (2007): "Investor Sentiment in the Stock Market," *The Journal of Economic Perspectives*, 21(2), 129–151.

BAKSHI, G., N. KAPADIA, AND D. MADAN (2003): "Stock Return Charakteristics, Skew Laws, and the Differential Pricing of Individual Equity Options," *The Review of Financial Studies*, 16(1), 101–143.

BAKSHI, G., AND D. MADAN (2006): "A Theory of Volatility Spreads," *Management Science*, 52(12), 1945–1956.

BAKSHI, G., G. PANAYOTOV, AND G. SKOULAKIS (2011): "The Baltic Dry Index as a Predictor of Global Stock Returns, Commodity Returns, and Global Economic Activity," *SSRN Working Paper*.

BARBERIS, N., AND W. XIONG (2012): "Realized Utility," *Journal of Financial Economics*, 104(2), 251–271.

BEKAERT, G., AND E. ENGSTROM (2009): "Asset Return Dynamics under Bad Environment-Good Environment Fundamentals," *NBER Working Papers 15222*.

BODIE, Z., AND V. I. ROSANSKY (1980): "Risk and Return in Commodity Futures," *Financial Analysts Journal*, 27(3), 27–39.

BOLLEN, N., AND R. WHALEY (2004): "Does Net Buying Pressure Affect the Shape of Implied Volatility Functions?," *The Journal of Finance*, 59(2), 711–753.

BOLLERSLEV, T., G. TAUCHEN, AND H. ZHOU (2009): "Expected Stock Returns and Variance Risk Premia," *Review of Financial Studies*, 22(11), 4463–4492.

BORENSZTEIN, E., AND C. M. REINHART (1994): "The Macroeconomic Determinants of Commodity Prices," *IMF Staff Papers*.

BUYUKSAHIN, B., M. S. HAIGH, J. H. HARRIS, J. A. OVERDAHL, AND M. A. ROBE (2008): "Fundamentals, Trader Activity and Derivative Pricing," *SSRN Working Paper*.

BUYUKSAHIN, B., M. S. HAIGH, AND M. A. ROBE (2010): "Commodities and Equities: Ever a "Market of One"?," *The Journal of Alternative Investments*, 12(3), 76–95.

BUYUKSAHIN, B., AND M. A. ROBE (2011): "Fundamentals, Trader Activity and Derivative Pricing," *SSRN Working Paper*.

——— (2012): "Speculators, Commodities and Cross-Market Linkages," *SSRN Working Paper*.

CARR, P., AND L. WU (2009): "Variance Risk Premiums," *Review of Financial Studies*, 22(3), 1311–1341.

CHEN, J., H. HONG, AND J. STEIN (2002): "Breadth of Ownership and Stock Returns," *The Journal of Financial Economics*, 66(2-3), 171–205.

CHEN, Y.-C., K. ROGOFF, AND B. ROSSI (2010): "Can Exchange Rates Forecast Commodity Prices?," *Quarterly Journal of Economics*, 125(3), 1145–1194.

CHENG, I.-H., A. KIRILENKO, AND W. XIONG (2012): "Convective Risk Flows in Commodity Futures Markets," *Working Paper*.

CHOPRA, N., C. M. LEE, AND R. H. THALER (1993): "Yes, Disounts on closed-end funds are a sentiment index," *The Journal of Finance*, 48(2), 801–808.

CONNOLLY, R., C. STIVERS, AND L. SUN (2005): "Stock Market Uncertainty and the Stock-Bond Return Relation," *The Journal of Financial and Quantitative Analysis*, 40(1), 161–194.

COUDERT, V., AND M. GEX (2008): "Does Risk Aversion Drive Financial Crisis? Testing the Predictive Power of Empirical Indicators," *The Journal of Empirical Finance*, 15(2), 167–184.

DA, Z., Q. LIU, AND E. SCHAUMBURG (2012): "Short-Term Return Reversal: The Long and the Short of It," *Working Paper*.

DELONG, J. B., A. SHLEIFER, H. L. SUMMERS, AND J. R. WALDMANN (1990): "Noise Trader Risk in Financial Markets," *Journal of Political Economy*, 98(4), 703–738.

DERRIEN, F. (2005): "IPO Pricing in "Hot" Market Conditions: Who Leaves Money on the Table?," *The Journal of Finance*, 60(1), 487–522.

DOMANSKI, D., AND A. HEATH (2007): "Financial Investors and Commodity Markets," *BIS Quarterly Review*.

DRECHSLER, I., AND A. YARON (2011): "What's Vol Got to Do with It," *Review of Financial Studies*, 24(1), 1–45.

DUFFIE, D., L. SAITA, AND K. WANG (2007): "Multi-Period Corporate Default Prediction with Stochastic Covariates," *The Journal of Financial Economics*, 83(3), 635–665.

DUSAK, K. (1973): "Futures Trading and Investor Returns: An Investigation of Commodity Market Risk," *The Journal of Political Economy*, 81(6), 1387–1406.

ERB, C., AND C. HARVEY (2006): "The Strategic and Tactical Value of Commodity Futures," *Financial Analysts Journal*, 62(2), 69–97.

ETULA, E. (2013): "Broker-Dealer Risk Appetite and Commodity Returns," *Journal of Financial Econometrics*, 11(3), 486–521.

FAMA, E., AND K. FRENCH (1987): "Commodity Futures Prices: Some Evidence on Forecast Power," *The Journal of Business*, 60(1), 55–73.

——— (2001): "Disappearing Dividends: Changing Firm Characteristics or Lower Propensity to Pay?," *The Journal of Financial Economics*, 60(1), 3–43.

FIGLEWSKI, S. (1989): "Options Arbitrage in Imperfect Markets," *The Journal of Finance*, 44(5), 1289–1311.

FRANKEL, J. A., AND J. WURGLER (2010): "The Effect of Monetary Policy on Real Commodity Prices," *NBER Working Paper*.

FUERTES, A.-M., J. MIFFRE, AND G. RALLIS (2010): "Tactical Allocation in Commodity Futures Markets: Combining Momentum and Term Structure Signals," *The Journal of Banking and Finance*, 34(10), 2530–2548.

GORTON, G., F. HAYASHI, AND G. ROUWENHORST (2013): "The Fundamentals of Commodity Futures Returns," *Review of Finance*, 17(1), 35–105.

GORTON, G., AND G. ROUWENHORST (2006): "Facts and Fantasies about Commodity Futures," *Financial Analysts Journal*, 62(2), 47–68.

GROEN, J., AND P. PESENTI (2010): "Commodity Prices, Commodity Currencies, and Global Economic Developments," *SSRN Working Paper*.

HAMILTON, J. D., AND J. C. WU (2014): "Risk Premia in Crude Oil Futures Prices," *Journal of International Money and Finance*, 42, 9–37.

HAN, B. (2008): "Investor Sentiment and Option Prices," *The Review of Financial Studies*, 21(1), 387–414.

HASTIE, T., R. TIBSHIRANI, AND J. FRIEDMAN (2009): *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer Series in Statistics, 2 edn.

HENDERSON, B. J., N. D. PEASON, AND L. WANG (2015): "New Evidence on the Financialization of Commodity Markets," *The Review of Financial Studies*, 28(5), 1285–1311.

HICKS, J. R. (1939): *Value and Capital: An Inquiry into Some Fundamental Principles of Economic Theory.* Clarendon Press, Oxford.

HONG, H., AND J. STEIN (2003): "Differences of Opinion, Short-Sales Constraints, and Market Crashes," *The Review of Financial Studies*, 16(2), 487–525.

HONG, H., AND M. YOGO (2010): "Commodity Market Interest and Asset Return Predictability," *SSRN Working Paper*.

———— (2012): "What Does Futures Market Interest Tell Us about the Macroeconomy and Asset Prices?," *The Journal of Financial Economics*, 105(3), 473–490.

HUANG, D., F. JIANG, J. TU, AND G. ZHOU (2014): "Investor Sentiment Aligned: A Powerful Predictor of Stock Returns," *The Review of Financial Studies*, forthcoming.

JONES, C. (2001): "A Century of Stock Market Liquidity and Trading Costs," *Working Paper, Columbia University*.

KELLY, B., AND S. PRUITT (2013): "Market Expectations in the Cross-Section of Present Values," *The Journal of Finance*, 68(5), 1721–1756.

KEYNES, J. M. (1923): "Some Aspects of Commodity Markets," *Manchester Guardian Commercial*, 13(4), 784–786.

KONDOR, P. (2009): "Risk in Dynamic Arbitrage: Price Effects of Convergence Trading," *The Journal of Finance*, 64(2), 638–658.

LEE, C., A. SHLEIFER, AND R. THALER (1990): "Anomalies: Closed-End Mutual Funds," *The Journal of Economic Perspectives*, 4(4), 154–164.

——— (1991): "Investor Sentiment and the Closed-End Fund Puzzle," *The Journal of Finance*, 46(1), 75–109.

LEE, W. Y., C. X. JIANG, AND D. C. INDRO (2002): "Stock Market Volatility, Excess Returns, and the Role of Investor Sentiment," *The Journal of Banking and Finance*, 26(12), 2277–2299.

LUDVIGSON, S. C., AND S. NG (2009): "Macro Factors in Bond Risk Premia," *The Review of Financial Studies*, 22(12), 5027–5067.

MASTERS, M. W. (2008): "Testimony before the Committee on Homeland Security and Government Affairs, U.S. Senate," http://www.hsgac.senate.gov//imo/media/doc/052008Masters.pdf?attempt=2.

MAYER, J. (2009): "The Growing Interdependence between Financial and Commodity Markets," *UNCTAD Working Paper*.

MILLER, E. (1977): "Risk, Uncertainty, and Divergence of Opinion," *The Journal of Finance*, 32(4), 1151–1168.

MORRISON, D. F. (1967): *Multivariate Statistical Methods*. New York: McGraw-Hill.

MOSKOWITZ, T. J., Y. H. OOI, AND L. H. PEDERSEN (2012): "Time Series Momentum," *The Journal of Financial Economics*, 104(2), 228–250.

MOU, Y. (2011): "Limits to Arbitrage and Commodity Index Investment: Front-Running the Goldman Roll," *Working Paper.*

NAGEL, S. (2011): "Evaporating Liquidity," *NBER Working Paper No. 17653.*

NEAL, R., AND S. WHEATLEY (1998): "Do Measures of Investor Sentiment Predict Returns?," *The Journal of Financial and Quantitative Analysis*, 33(4), 523–548.

OFEK, E., M. RICHARDSON, AND R. WHITELAW (2004): "Limited Arbitrage and Short Sales Restrictions: Evidence From the Options Markets," *The Journal of Financial Economics*, 74(2), 305–342.

PINDYCK, R., AND J. ROTEMBERG (1990): "The Excess-Comovement of Commodity Prices," *Economic Journal*, 100(403), 1173–1189.

PONTIFF, J. (1996): "Costly Arbitrage: Evidence from Closed-End Funds," *The Quarterly Journal of Economics*, 111(4), 1135–1151.

REHMAN, Z., AND G. VILKOV (2012): "Risk-Neutral Skewness: Return Predictability and Its Sources," *SSRN Working Paper.*

SCHENKMAN, J., AND W. XIONG (2003): "Overconfidence and Speculative Bubbles," *Journal of Political Economy*, 111(6), 1183–1219.

SINGLETON, K. J. (2011): "Investor Flows and the 2008 Boom/Bust in Oil Prices," *SSRN Working Paper.*

STOCK, J. H., AND M. W. WATSON (2002): "Macroeconomic Forecasting Using Diffusion Indexes," *The Journal of Business & Economic Statistics*, 20(2), 147–162.

SZYMANOWSKA, M., F. DE ROON, T. NIJMAN, AND R. VAN DEN GOORBERGH (forthcoming): "An Anatomy of Commodity Futures Risk Premia," *The Journal of Finance*.

TANG, K., AND W. XIONG (2012): "Index Investment and Financialization of Commodities," *Financial Analysts Journal*, 68(6), 54–74.

TETLOCK, P. C. (2007): "Giving Content to Investor Sentiment," *The Journal of Finance*, 62(3), 1139–1168.

WHALEY, R. (2000): "The Investor Fear Gauge," *The Journal of Portfolio Management*, 26(3), 12–17.

ZWEIG, M. (1973): "An Investor Expectations Stock Price Predictive Model Using Closed End Fund Premiums," *The Journal of Finance*, 28(1), 67–87.

Appendix A: Commodity Group Composition and Data Sources

This table provides information about the composition of our commodity groups. In total, we have 22 different commodities in seven groups: Animal, Energy, Grains, Industrials, Metals, Precious, and Softs. The table also gives information about the market and about the source of the price information.

No.	Commodity	Market	Group
1	Live Cattle	Chicago Mercantile Exchange	Animal
2	Lean Hogs	Chicago Mercantile Exchange	Animal
3	Feeder Cattle	Chicago Mercantile Exchange	Animal
4	Heating Oil	New York Mercantile Exchange	Energy
5	Crude Oil	New York Mercantile Exchange	Energy
6	Wheat	Chicago Board of Trade	Grains
7	Corn	Chicago Board of Trade	Grains
8	Soybeans	Chicago Board of Trade	Grains
9	Soybean Oil	Chicago Board of Trade	Grains
10	Soybean meal	Chicago Board of Trade	Grains
11	Oats	Chicago Board of Trade	Grains
12	Cotton	Coffee, Sugar, and Cocoa Exchange	Industrials
13	Lumber	Chicago Mercantile Exchange	Industrials
14	Copper	New York Commodities Exchange	Metals
15	Silver	New York Commodities Exchange	Precious
16	Platinum	New York Mercantile Exchange	Precious
17	Gold	New York Commodities Exchange	Precious
18	Palladium	New York Mercantile Exchange	Precious
19	Cocoa	Coffee, Sugar and Cocoa Exchange	Softs
20	Sugar	Coffee, Sugar and Cocoa Exchange	Softs
21	Orange Juice	New York Commodities Exchange	Softs
22	Coffee	Coffee, Sugar and Cocoa Exchange	Softs

Electronic copy available at: https://ssrn.com/abstract=1934397

Appendix B: Construction of the Daily Closed-End Fund Discounts (CEFD) and Dividend Premium Proxies

B.1 Dividend Premium

We follow Fama and French (2001) and Baker and Wurgler (2004b) to compute the daily and monthly dividend premium. Our sample is derived from aggregations of COMPUSTAT data (provided through Wharton Research Data Services). It includes all firms listed on a U.S. stock exchange having the following data (COMPUSTAT data items are given in parantheses): total assets (AT), closing stock prices (PRCCD), number of common shares outstanding (CSHOC), [cash] dividends per share by ex date (DVPSX), and preferred dividends (DVP). We exclude companies with assets below \$500,000 and equity values lower than \$250,000. Consistent with the literature, we do not consider utilities (SIC codes 4900 to 4949) or financial firms (SIC codes 6000 to 6999). Finally, all firms in our sample have a quarterly reporting periodicity (RP) and are publicly traded. Our sample results in a total of 5801 firms in the cross section.

Market equity is determined as the end of month stock price (PRCCM) times shares outstanding (CSHOM). Book equity is collected at a quarterly frequency. It is defined as stockholders' equity (SEQ) [or first available of common equity (CEQ) plus preferred stock at carrying value (UPSTK) or book assets (AT) minus liabilities (LT)] minus preferred stock at liquidating value (PSTKL) [or redemption value (PSTKRV)] plus balance sheet deferred taxes and investment tax credit (TXDITCQ) if available. If values are given in two successive years, i.e., in t and $t + 1$, we allocate the data linearly over the missing quarters unless a quarterly value of better quality is available. The market value of equity is determined on a daily basis by multiplying the closing price with the number of shares outstanding. [16] This allows the calculation of the market-to-book ratio, which is defined as book assets minus

---

[16] As the daily samples of Compustat only start in 2002, we decided to retrieve daily closing prices and the number of outstanding shares from CRSP.

book equity plus market value of equity all divided by book assets.

Our approach differs from the procedure in the literature in that we identify the dividend-paying firms in a more rigorous and timely fashion. Similar to Baker and Wurgler (2004b), we classify each sample firm as dividend "payer" or "nonpayer". In previous work, a firm was regarded as dividend payer if it had ever issued a dividend in the previous 12 months (1 year). This yearly frequency is indeed low relative to the speed of information transmission and its reflection in the price. Moreover, if a firm accidentally pays a dividend in a year, its inclusion in the dividend payer group tends to be inconsistent. Hence, we classify a firm as a dividend payer if we observe a dividend payment within both the six months previous to and the six months after the price observation month. Similarly, firms not paying dividends during that period enter the sample as nonpayers. All other firms remain unclassified. Following Baker and Wurgler (2004b, 2006), we take the equally-weighted averages of the market-to-book values separately for payers and nonpayers. The dividend premium is then defined as the difference of the logs of these averages.

B.2 Closed-End Fund Discounts

We identify closed-end funds by the CRSP daily stock files (Share Code *4). This results in a total sample of 1068 funds. Next, we download daily data on net asset values (NAVs) from the Bloomberg database. For each trading day, we calculate equally-weighted closed-end fund discounts by calculating the average difference between NAVs and the corresponding fund share market prices.