NORTHERN ARIZONA UNIVERSITY
The W. A. Franke College of Business

Investor Sentiment and Oil Prices

Working Paper Series—15-02 | November 2015

Ding Du
Northern Arizona University
The W. A. Franke College of Business, PO Box 15066
Flagstaff, AZ 86011
ding.du@nau.edu
(928) 523-7274
Fax: (928) 523-7331

Ronald J Gunderson
Northern Arizona University
The W. A. Franke College of Business, PO Box 15066
Flagstaff, AZ 86011
ronald.gunderson@nau.edu
(928) 523-7358
Fax: (928) 523-7331

Xiaobing Zhao
Northern Arizona University
The W. A. Franke College of Business, PO Box 15066
Flagstaff, AZ 86011
xiaobing.zhao@nau.edu
(928) 523-7279
Fax: (928) 523-7331

Investor Sentiment and Oil Prices

Introduction

Empirical studies on the determinants of oil prices have explored a broad range of economic variables, including global real economic activity, gasoline/heating oil price spread, oil future prices, the spot price of industrial raw materials, crude oil production and inventories, oil-company stock prices, U.S. interest rates, and exchange rates (e.g., Hamilton, 2009; Alquist and Kilian, 2010; Chen *et al*, 2010; Reeve and Vigfusson, 2011; Baumeister and Kilian, 2012; Fattouh *et al*, 2013; Chen, 2014; Baumeister and Kilian, 2015; Baumeister *et al*, 2015). In this paper, we investigate if investor sentiment in financial markets is a determinant of oil prices. Our investigation is motivated by the following three observations.

First, speculation can play a role in the oil market. The structural model in Kilian (2009) suggests that oil prices are driven by crude oil supply shocks, global aggregate demand shocks, and precautionary oil demand shocks. Precautionary oil demand shocks "could arise because of concerns over unexpected growth of demand, over unexpected declines of supply, or over both" (Kilian, 2009, p1054). Since it is difficult to accurately forecast future oil demand and supply based on available fundamentals information, precautionary oil demand could lead to speculation. Masters (2008) hypothesizes that speculation leads to the oil price surge during 2003 - 2008. Empirically, Kilian and Murphy (2014) find that although the 2003 - 2008 oil price surge is not caused by speculation[1], speculative demand does help explain some fluctuations in oil prices.

Second, in general, investor sentiment in financial markets can have significant effects on speculative demand for assets, because (1) investors are subject to sentiment (Delong *et al*, 1990), and (2) arbitrage against sentimental investors can be costly and risky (Shleifer and Vishny, 1997; Wurgler and Zhuravskaya, 2002; Lamont and Thaler, 2003; Brunnermeier and Pedersen, 2005). Empirically, a growing literature in finance documents the effects of investor sentiment on equity returns. For instance, Baker and Wurgler (2006, 2007) find that investor sentiment helps explain the cross-section of U.S. stock returns. Baker *et al* (2012) provide similar international evidence. Stambaugh *et al* (2012) show that sentiment helps explain a broad set of stock-market anomalies.

Third, the "financialization" of commodity markets results in commodities including crude oil becoming an asset class alongside equities for financial investors (Arezki *et al*, 2014). Consequently, investor sentiment in financial markets could exert important influences on oil speculative demand and therefore oil prices. Empirically, Büyükşahin and Robe (2014) document that greater participation by financial investors in commodity futures markets results in the co-movement between commodities and equities returns.

---

[1] See also Fattouh *et al* (2013).

Taken together, the above three observations suggest that investor sentiment may be an important determinant of oil prices. To test our conjecture, we use the sentiment index developed by Baker and Wurgler (2006, 2007) (BW). The BW sentiment index is designed to capture the common or market-wide sentiment in financial markets by filtering out idiosyncratic noise in the six underlying sentiment measures through the principal component analysis. The six measures are the closed-end fund discount, the number and the first-day returns of IPOs, NYSE turnover, the equity share in total new issues, and the dividend premium.

In terms of empirical methodology, we follow Baker and Wurgler (2006, 2007) and focus on the predictability of investor sentiment. The idea is as follows. The contemporaneous correlation between investor sentiment and oil prices may not necessarily mean causality, because of reverse causality and confounding factors. For instance, global aggregate demand may drive both sentiment and oil prices, leading to a spurious (positive) correlation between oil prices and investor sentiment. To circumvent this problem, we instead identify the causal effects of investor sentiment by examining if the predictability in oil prices depends on *prior* sentiment. For instance, low oil returns, conditional on high prior sentiment, would be consistent with the ex-ante sentiment-driven overvaluation of oil and subsequent mispricing correction. Therefore, we organize our empirical analysis loosely around the following predictive specification:

$$R_{t+h} = \frac{P_{t+h} - P_t}{P_t} = \alpha_h + \beta_h BW_t + \sum_k \gamma_{kh} X_{k,t} + \varepsilon_{t+h} \qquad (1)$$

where R_{t+h} is the percentage change in the oil price (P) from month t to month $t + h$, BW_t is the BW sentiment index in month t, and X's are other predictive variables suggested by previous studies.

We start with ordinary least-squares regressions (OLS) as in Baker and Wurgler (2006). However, the relationship between oil returns and investor sentiment may not be the same across the entire conditional distribution of oil returns. Thus, we supplement our OLS analysis with the quantile regression (QR) proposed by Koenker and Bassett (1978). Median QR is also more robust to outliers than least squares regression, and is semi parametric as it avoids assumptions about the parametric distribution of the error process. Finally, we explore if investor sentiment also helps explain the movements in gasoline, heating oil, and oil-company stock prices, because they are all strongly correlated with crude oil prices.

Our findings can be easily summarized. Investor sentiment helps explain the fluctuations in oil prices (as well as gasoline, heating oil and oil-company stock prices). High/low sentiment predicts subsequent low/high oil returns particularly at longer horizons. Our findings have important theoretical as well as practical implications. In terms of theoretical implications, our findings suggest that future theoretical models of oil prices should take into account both fundamentals and investor sentiment. In terms of practical implications, our findings imply a new predictor of oil prices.

The remainder of the paper is organized as follows: Section 2 discusses our data. Section 3 presents the OLS results. Section 4 reports the QR results. Section 5 provides additional evidence. Section 6 concludes the paper with a brief summary.

2 Data

We consider both the price of West Texas Intermediate crude oil (WTI) and the Brent price of crude oil (Brent). While WTI is a benchmark for the U.S. oil market, Brent is a benchmark for global oil markets (Baumeister and Kilian, 2014). Following previous studies, we examine both nominal and real oil prices. We deflate nominal oil prices by the U.S. CPI to obtain real oil prices. The monthly nominal oil prices and the U.S. CPI from January 1986 to November 2014 are from the Federal Reserve Bank of St. Louis. Panels A and B of Figure 1 depict the nominal and real WTI and Brent. Since the inflation rate has been relatively stable over this sample period, nominal and real oil prices fluctuate in a similar fashion.

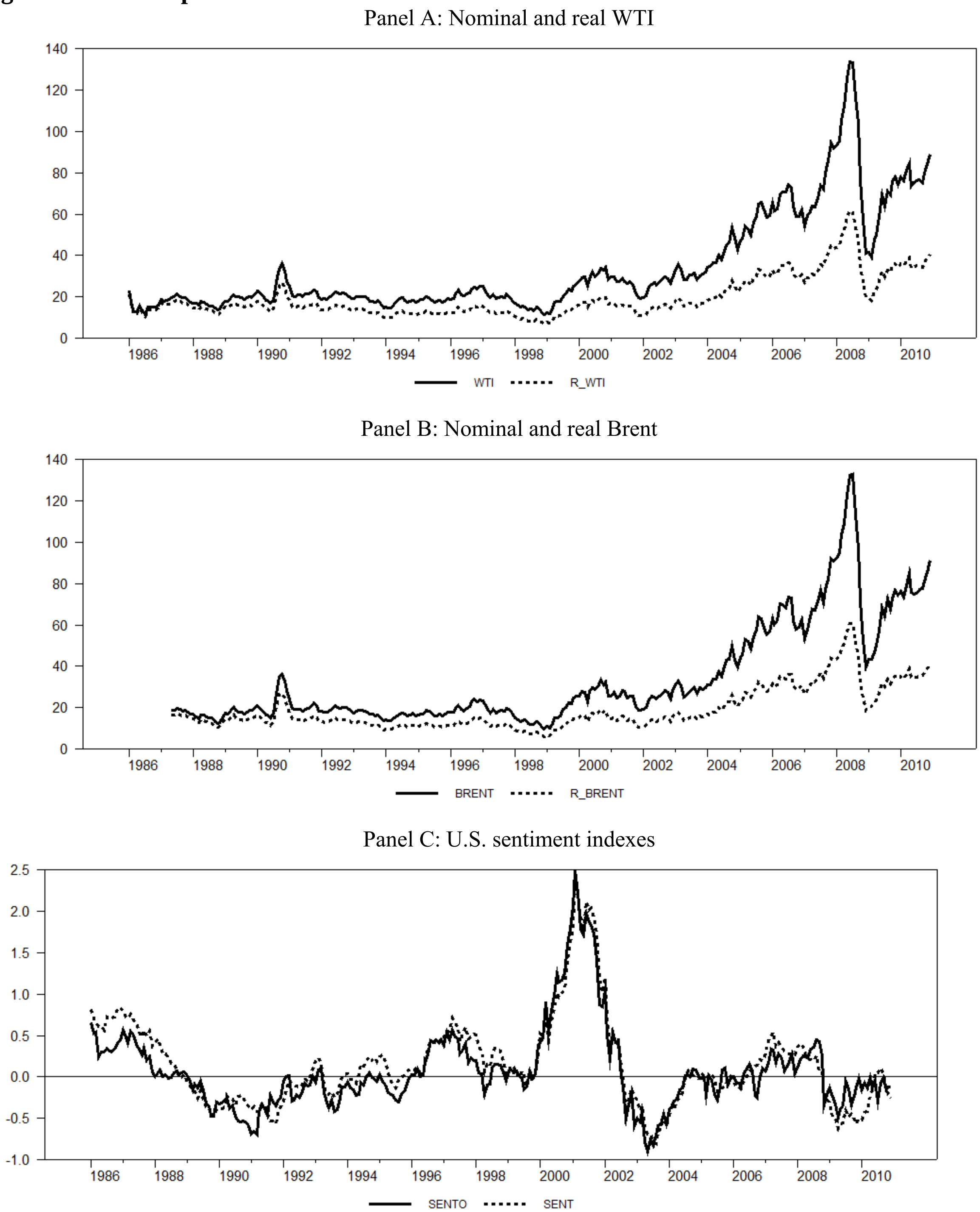
The monthly BW sentiment index data from July 1965 to December 2010 are downloaded from Professor Jeffrey Wurgler's web site.[2] Although the BW index is only updated to 2010, we still use it because (1) it is the dominant sentiment measure in the finance literature (e.g., McLean and Zhao, 2014; Neely *et al*; 2014), (2) it is designed specifically to capture the market-wide investor sentiment in financial markets, and is free of idiosyncratic noise in individual sentiment measures. The BW sentiment index has two versions. One is orthogonalized to the business cycle variables (SENTO), and the other is not (SENT). Following the sentiment literature, we focus on the orthogonalized sentiment index, a "pure" sentiment measure. However, we also examine the sentiment index that is not orthogonalized to the business cycle variables, since such analysis shed additional light on the role of sentiment. The two sentiment indexes, SENTO and SENT, are plotted in Panel C of Figure 1. Both indexes seem to capture most anecdotal accounts of movements in investor sentiment (e.g., substantial fluctuations in sentiment during the internet bubble).

Previous studies have suggested a broad set of economic variables as the determinants of oil prices, including global real economic activity, gasoline/heating oil price spread, oil future prices, the spot price of industrial raw materials, crude oil production and inventories, oil-company stock prices, U.S. interest rates, and exchange rates (e.g., Hamilton, 2009; Alquist and Kilian, 2010; Chen *et al*, 2010; Reeve and Vigfusson, 2011; Baumeister and Kilian, 2012; Fattouh *et al*, 2013; Chen, 2014; Baumeister and Kilian, 2015; Baumeister *et al*, 2015). In this paper, we focus on gasoline/heating oil spread, the

[2] We thank Professor Jeffrey Wurgler for making these data available at http://people.stern.nyu.edu/jwurgler/.

global real activity measure of Kilian (2009), exchange rates, interest rates, and oil-company stock prices. We chose these oil price

Figure 1 Crude oil prices and the U.S. sentiment



Panels A and B depict the nominal and real WTI and Brent. The BW sentiment index has two versions. One is orthogonalized to the business cycle variables (SENTO), and the other is not (SENT). The two sentiment measures are plotted in Panel C.

determinants, because (1) they are emphasized in recent empirical studies (e.g., Chen, 2014; Baumeister *et al*, 2015) and (2) their data are available to us. The gasoline price ($P_{Gasoline}$) and the heating oil price ($P_{Heating\ oil}$) are from U.S. Energy Information Administration (EIA). The global real activity measure (Kilian) is downloaded from Professor Lutz Kilian's web site.[3] Following Baumeister *et al* (2015), we use the trade weighted U.S. dollar index for major currencies (MCI) from the Federal Reserve Bank of St. Louis as our measure of exchange rates.[4] The U.S. risk-free rate (RF) from Professor Kenneth French's website is employed to proxy interest rates.[5] Following Chen (2014), we use the closing price of the NYSE ARCA index (ARCA) from *Yahoo Finance* to measure oil-company stock prices.

Table 1 Summary statistics

Panel A: Key statistics

	Obs	Mean	Std Error	Minimum	Maximum
WTI	300	34.29	24.45	11.35	133.88
BRENT	284	33.96	24.82	9.82	132.72
SENTO	300	0.08	0.51	-0.90	2.50
SENT	300	0.15	0.53	-0.81	2.32
GASOLINE	295	44.74	29.37	14.33	153.66
HEATING_OIL	295	44.71	31.76	14.19	177.41
KILIAN	300	0.86	24.34	-56.76	60.26
MCI	300	90.48	9.90	70.34	120.19
RF	300	0.34	0.18	0.00	0.79
ARCA	300	531.66	365.06	122.43	1569.59

Panel B: Correlation matrix

	WTI	BRENT	SENTO	SENT	GASOLINE	HEATING_OIL	KILIAN	MCI	RF
BRENT	0.999								
SENTO	0.009	0.007							
SENT	-0.097	-0.096	0.949						
GASOLINE	0.990	0.991	0.004	-0.104					
HEATING_OIL	0.995	0.994	0.014	-0.088	0.984				
KILIAN	0.750	0.746	-0.151	-0.224	0.767	0.743			
MCI	-0.682	-0.686	0.442	0.492	-0.677	-0.682	-0.688		
RF	-0.470	-0.466	0.125	0.193	-0.448	-0.453	-0.284	0.299	
ARCA	0.923	0.925	0.097	0.014	0.931	0.921	0.685	-0.561	-0.468

WTI is the price of West Texas Intermediate crude oil, BRENT is the Brent price of crude oil, SENTO is the Baker and Wurgler (2006, 2007) sentiment index that is orthogonalized to the business cycle variables, SENT is the Baker and Wurgler (2006, 2007) sentiment index that is not orthogonalized to the business cycle variables, GASOLINE is the gasoline price, HEATING_OIL is the heating oil price, KILIAN is the global real activity measure of Kilian (2009), MCI is the trade weighted U.S. dollar index for major currencies (MCI) from the Federal Reserve Bank of St. Louis, RF is the risk-free rate, and ARCA is the closing price of the NYSE ARCA index (ARCA) from *Yahoo Finance*. Panel A reports the summary statistics for these main variables used in the paper, while Panel B presents the corresponding correlations among these variables

[3] We thank Professor Lutz Kilian for making the data available at http://www-personal.umich.edu/~lkilian/.
[4] Exchange rates also affect financial markets (e.g., Du and Hu, 2012; Du, 2014; Du and Hu, 2014).
[5] We thank Professors Fama and French for making these data available at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/.

Our merged data cover the sample period from January 1986 to December 2010. The starting point for the sample period is dictated by the availability of the oil prices data from the Federal Reserve Bank of St. Louis. Table 1 reports the summary statistics. The strong correlations between oil prices and gasoline/heating oil as well as oil-company stock prices are worth noting. These strong correlations are expected from a theoretical perspective, and suggest that if sentiment can explain movements in oil prices, it should also be able to explain the fluctuations in gasoline, heating oil, and oil-company stock prices.

3 Ordinary least-squares regressions

3.1 Univariate regressions

Following Baker and Wurgler (2006), we first run univariate predictive regressions. The idea is to examine if investor sentiment has any explanatory power, and to compare its explanatory power to that of other determinants suggested by previous studies.

For nominal oil prices, we run the following regressions.

$$R_{t+h} = \alpha_h + \beta_h SENTO_t + \varepsilon_{t+h} \quad (2a)$$

$$R_{t+h} = \alpha_h + \gamma_{kh} X_{k,t} + \varepsilon_{t+h} \quad (2b)$$

where *R* is the percentage change in WTI or Brent, *SENTO* is the BW sentiment index that is orthogonalized to the business cycle variables, *X* is the oil price determinant suggested by previous studies such as the gasoline/heating oil price spread, the return on oil-company stocks, the global real activity change, the exchange rate change, and the interest rate change (see Kilian, 2009; Chen *et al*, 2010; Chen, 2014; Baumeister *et al*, 2015), and *h* = 1, 3, 6, 9, 12, 15, 18, 21 and 24 for our monthly data. Following previous studies (e.g. Baumeister *et al*, 2015), we define the gasoline price spread as $S\_Gas_t = \log(P_{Gasoline,\,t}) - \log(P_t)$, the heating oil price spread as $S\_Heating_t = \log(P_{Heating\ oil,\,t}) - \log(P_t)$, the return on the oil-company stocks as $R_{ARCA,t} = \frac{ARCA_t - ARCA_{t-1}}{ARCA_{t-1}}$, the global real activity change as $\Delta Kilian_t = \frac{Kilian_t - Kilian_{t-h}}{Kilian_{t-h}}$, the exchange rate change as $\Delta MCI_t = \frac{MCI_t - MCI_{t-h}}{MCI_{t-h}}$, and the interest rate change as $\Delta RF_t = RF_t - RF_{t-h}$.

For real oil price regressions, we modify our regressions accordingly. More specifically, the dependent variable is $RR_{t+h} = \frac{P_{t+h} / CPI_{t+h} - P_t / CPI_t}{P_t / CPI_t}$, where *P* is WTI or Brent. Because the BW sentiment index and the Kilian global real activity index are not in dollars, we do not adjust them. Since the gasoline and heating oil spreads are log price differences, the real and nominal spreads are the same. We define the real return on the oil-company stocks as $RR_{ARCA,t} = \frac{ARCA_t / CPI_t - ARCA_{t-1} / CPI_{t-1}}{ARCA_{t-1} / CPI_{t-1}}$, the real exchange rate change as $\Delta RMCI_t = \dfrac{RMCI_t - RMCI_{t-h}}{RMCI_{t-h}}$ where *RMCI* is the real trade weighted U.S. dollar index for major currencies (RMCI) from the Federal Reserve Bank of St. Louis, and the real interest rate change as $\Delta RRF_t = (RF_t - I_t) - (RF_{t-h} - I_{t-h})$ where *I* is the CPI inflation rate.

If high sentiment leads to overvaluation and subsequent mispricing correction or price reversals, we expect the coefficient on SENTO to be negative in predictive regressions. Previous studies also suggest that the gasoline and heating oil spreads, the Kilian index change, the interest rate change, and the oil-company stock return should all have positive coefficients (e.g., Kilian, 2009; Chen *et al*, 2010; Chen, 2014; Baumeister and Kilian, 2015). Thus, our discussion of statistical significance will be based on one-sided tests in this paper.

The results for the nominal WTI are reported in Panel A of Table 2. The t-statistics are based on Newey-West HAC standard errors with the lag parameter set equal to 12 for the monthly regressions. We report the coefficient estimates of the explanatory variables as well as the adjusted R^2. First, consistent with Chen (2014) and Baumeister *et al* (2015), the financial variables such as oil-company stock returns and interest rate changes have significant power to explain the fluctuations in the nominal WTI, although they lose their explanatory power at longer horizons. Second, consistent with our conjecture, the coefficient of sentiment is negative and statistically significant particularly at longer horizons. Furthermore, the sentiment coefficient monotonically increases from -0.019 (t = -2.55) at horizon of 1 month to -0.184 (t = -1.77) at horizon of 21 months. This is expected, since mispricing correction is not instantaneous. Third, in terms of explanatory power or adjusted R^2, it appears that sentiment has nontrivial explanatory power, particularly at longer horizons. For instance, the adjusted R^2s are 2.8%, 4.3%, and 4.4% at horizons of 6 months, 12 months and 18 months for sentiment, respectively. In contrast, the adjusted R^2s are 0.6%, 0.1%, and -0.3% at same horizons for oil-company stock returns, respectively. The evidence thus suggests that sentiment is important for understanding the movements in nominal WTI.

The results for the nominal Brent are reported in Panel B of Table 2, and are generally consistent with those for the nominal WTI. First, the financial variables such as oil-company stock returns and interest rate changes have significant power to explain the fluctuations in the nominal Brent at shorter horizons. Second, the coefficient of sentiment is negative and statistically significant particularly at longer horizons. Third, sentiment has important explanatory power, especially at longer horizons. For instance, the adjusted R^2s are 2.7%, 4.8%, 4.5%, and 3.9% at horizons of 6 months, 12 months, 18 months, and 24 months for sentiment, respectively. In contrast, the adjusted R^2s are 1.1%, 0.5%, -0.2%, and -0.3% at same horizons for oil-company stock returns, respectively. The evidence thus suggests that sentiment helps understand the fluctuations in nominal Brent.

Table 2 Univariate regressions for nominal oil prices

Panel A: WTI									
	1	3	6	9	12	15	18	21	24
SENTO	-0.019	-0.048	-0.085	-0.114	-0.143	-0.152	-0.168	-0.184	-0.170
	(-	(-	(-	(-	(-2.15)	(-	(-	(-	(-
R^2	0.009	0.016	0.028	0.038	0.043	0.039	0.044	0.048	0.036
S Gas	0.115	0.466	0.445	0.219	0.566	0.761	0.282	0.025	0.721
	(1.22)	(1.61)	(1.54)	(0.63)	(1.31)	(1.76)	(0.68)	(0.06)	(1.47)
R^2	0.010	0.047	0.020	0.001	0.016	0.026	0.000	-0.003	0.016
S Heatin	-0.015	0.015	0.120	0.179	-0.019	-0.391	-0.749	-0.709	-0.632
	(-	(0.08)	(0.37)	(0.41)	(-0.04)	(-	(-	(-	(-
R^2	-0.003	-0.003	-0.002	-0.001	-0.003	0.002	0.014	0.011	0.007
ΔKilian	0.002	-0.001	-0.001	0.000	0.000	0.002	0.001	0.001	0.002
	(0.65)	(-	(-	(0.08)	(0.12)	(1.02)	(0.38)	(0.54)	(0.82)
R^2	0.003	-0.003	-0.003	-0.003	-0.003	-0.000	-0.003	-0.003	-0.002
ΔMCI	-0.635	-0.441	-0.085	0.087	0.004	-0.034	-0.170	-0.069	0.107
	(-	(-	(-	(0.12)	(0.00)	(-	(-	(-	(0.12)
R^2	0.014	0.005	-0.003	-0.003	-0.003	-0.003	-0.002	-0.003	-0.003
ΔRF	0.247	0.057	-0.002	0.053	0.167	0.220	0.151	-0.060	-0.231
	(2.32)	(0.32)	(-	(0.17)	(0.52)	(0.75)	(0.59)	(-	(-
R^2	0.015	-0.003	-0.003	-0.003	0.000	0.004	0.001	-0.003	0.007
R_{ARCA}	0.371	0.502	0.425	0.271	0.401	0.333	0.051	-0.084	0.031
	(3.26)	(2.16)	(2.09)	(1.15)	(1.45)	(1.12)	(0.15)	(-	(0.09)
R^2	0.055	0.023	0.006	-0.001	0.001	-0.001	-0.003	-0.003	-0.003
					Panel B:				
	1	3	6	9	12	15	18	21	24
SENTO	-0.018	-0.048	-0.090	-0.125	-0.160	-0.173	-0.189	-0.208	-0.199
	(-	(-	(-	(-	(-2.29)	(-	(-	(-	(-
R^2	0.007	0.013	0.027	0.040	0.048	0.043	0.045	0.049	0.039
S Gas	0.092	0.402	0.435	0.155	0.360	0.723	0.475	0.270	0.865
	(1.16)	(1.29)	(1.23)	(0.45)	(0.84)	(1.41)	(0.83)	(0.47)	(1.31)
R^2	0.007	0.041	0.024	-0.001	0.007	0.028	0.009	-0.000	0.028
S Heatin	-0.022	0.012	0.241	0.288	0.011	-0.087	-0.233	-0.130	0.035
	(-	(0.06)	(0.76)	(0.80)	(0.03)	(-	(-	(-	(0.06)
R^2	-0.003	-0.004	0.001	0.002	-0.004	-0.003	-0.002	-0.003	-0.004
ΔKilian	0.003	-0.001	-0.001	-0.000	0.000	0.002	0.001	0.001	0.002
	(0.87)	(-	(-	(-	(0.18)	(1.07)	(0.32)	(0.43)	(0.82)
R^2	0.005	-0.002	-0.003	-0.004	-0.003	-0.000	-0.003	-0.003	-0.002
ΔMCI	-0.888	-0.394	0.053	0.097	-0.083	-0.087	-0.150	-0.077	0.084
	(-	(-	(0.11)	(0.13)	(-0.10)	(-	(-	(-	(0.08)
R^2	0.025	0.002	-0.003	-0.003	-0.003	-0.003	-0.003	-0.003	-0.003
ΔRF	0.241	0.104	0.104	0.076	0.200	0.306	0.188	-0.085	-0.309
	(1.74)	(0.52)	(0.48)	(0.24)	(0.64)	(1.07)	(0.73)	(-	(-
R^2	0.011	-0.003	-0.002	-0.003	0.001	0.009	0.002	-0.002	0.012
R_{ARCA}	0.522	0.637	0.575	0.455	0.626	0.529	0.301	0.215	0.222
	(4.02)	(2.56)	(2.69)	(1.76)	(2.15)	(1.58)	(0.80)	(0.51)	(0.57)
R^2	0.095	0.030	0.011	0.003	0.005	0.001	-0.002	-0.003	-0.003

The dependent variable is $RR_{t+h} = \frac{P_{t+h} - P_{t\,t}}{P_t}$, where *P* is WTI or Brent. *SENTO* is the BW sentiment index that is orthogonalized to the business cycle variables. We define the gasoline price spread as $S\_Gas_t = \log(P_{Gasoline,\,t}) - \log(P_t)$, the heating oil price spread as $S\_Heating_t = \log(P_{Heating\ oil,\,t}) - \log(P_t)$, the return on the oil-company stocks as $R_{ARCA,t} = \frac{ARCA_t - ARCA_{t-1}}{ARCA_{t-1}}$, the global real activity change as $\Delta Kilian_t = \frac{Kilian_t - Kilian_{t-h}}{Kilian_{t-h}}$, the exchange rate change as $\Delta MCI_t = \frac{MCI_t - MCI_{t-h}}{MCI_{t-h}}$, and the interest rate change as $\Delta RF_t = RF_t - RF_{t-h}$. The t-statistics are based on Newey-West HAC standard errors with the lag parameter set equal to 12 for the monthly regression.

Table 3 Univariate regressions for real oil prices

					Panel A: WTI				
	1	3	6	9	12	15	18	21	24
SENTO	-0.018	-0.046	-0.081	-0.107	-0.133	-0.140	-0.153	-0.166	-0.150
	(-	(-	(-	(-	(-2.05)	(-	(-	(-	(-
R^2	0.008	0.015	0.027	0.036	0.041	0.037	0.041	0.044	0.032
S Gas	0.031	0.129	0.203	0.249	0.384	0.505	0.492	0.487	0.646
	(1.13)	(1.61)	(1.62)	(1.55)	(2.03)	(2.78)	(2.60)	(2.13)	(2.67)
R^2	0.002	0.017	0.023	0.026	0.046	0.067	0.058	0.051	0.084
S Heatin	0.006	0.033	0.116	0.200	0.238	0.273	0.303	0.344	0.384
	(0.22)	(0.48)	(0.93)	(1.20)	(1.20)	(1.26)	(1.39)	(1.58)	(1.66)
R^2	-0.003	-0.002	0.006	0.018	0.019	0.020	0.023	0.028	0.032
ΔKilian	0.002	-0.001	-0.001	0.000	0.000	0.002	0.001	0.001	0.002
	(0.63)	(-	(-	(0.13)	(0.15)	(1.05)	(0.44)	(0.59)	(0.90)
R^2	0.002	-0.003	-0.003	-0.003	-0.003	0.000	-0.003	-0.003	-0.002
ΔMCI	-0.569	-0.419	-0.170	-0.024	-0.108	-0.103	-0.190	-0.086	0.107
	(-	(-	(-	(-	(-0.14)	(-	(-	(-	(0.13)
R^2	0.010	0.004	-0.002	-0.003	-0.003	-0.003	-0.002	-0.003	-0.003
ΔRF	-0.003	-0.068	0.006	0.029	0.096	0.178	0.142	0.076	0.079
	(-	(-	(0.18)	(0.60)	(2.00)	(2.32)	(2.16)	(1.09)	(1.07)
R^2	-0.003	0.019	-0.003	-0.002	0.011	0.028	0.017	0.002	0.002
R_{ARCA}	0.352	0.476	0.428	0.277	0.413	0.352	0.081	-0.050	0.067
	(3.38)	(2.19)	(2.15)	(1.20)	(1.54)	(1.23)	(0.25)	(-	(0.20)
R^2	0.051	0.021	0.007	-0.000	0.002	-0.000	-0.003	-0.003	-0.003
					Panel B:				
	1	3	6	9	12	15	18	21	24
SENTO	-0.017	-0.046	-0.086	-0.118	-0.149	-0.159	-0.172	-0.187	-0.176
	(-	(-	(-	(-	(-2.19)	(-	(-	(-	(-
R^2	0.006	0.013	0.025	0.038	0.045	0.040	0.042	0.045	0.035
S Gas	0.068	0.259	0.427	0.483	0.647	0.863	0.859	0.847	1.130
	(2.20)	(2.86)	(3.70)	(2.76)	(2.91)	(3.59)	(3.05)	(2.56)	(3.24)
R^2	0.011	0.047	0.068	0.066	0.088	0.125	0.110	0.096	0.154
S Heatin	0.025	0.095	0.279	0.430	0.454	0.514	0.567	0.642	0.741
	(0.68)	(1.11)	(1.98)	(2.26)	(1.98)	(1.86)	(1.94)	(2.20)	(2.33)
R^2	-0.001	0.004	0.031	0.059	0.048	0.048	0.053	0.062	0.073
ΔKilian	0.003	-0.001	-0.001	-0.000	0.000	0.002	0.001	0.001	0.002
	(0.85)	(-	(-	(-	(0.20)	(1.10)	(0.37)	(0.47)	(0.89)
R^2	0.004	-0.002	-0.003	-0.004	-0.003	0.000	-0.003	-0.003	-0.002
ΔRMCI	-0.828	-0.392	-0.021	-0.023	-0.215	-0.168	-0.193	-0.124	0.051
	(-	(-	(-	(-	(-0.26)	(-	(-	(-	(0.06)
R^2	0.021	0.002	-0.004	-0.004	-0.002	-0.002	-0.002	-0.003	-0.004
ΔRRF	-0.009	-0.045	0.035	0.047	0.110	0.181	0.159	0.081	0.065
	(-	(-	(0.92)	(0.85)	(2.03)	(1.97)	(2.32)	(1.08)	(0.82)
R^2	-0.003	0.005	-0.001	-0.000	0.012	0.024	0.018	0.002	-0.001
RR_{ARCA}	0.505	0.616	0.583	0.463	0.640	0.548	0.332	0.250	0.270
	(4.24)	(2.64)	(2.80)	(1.83)	(2.27)	(1.70)	(0.92)	(0.64)	(0.74)
R^2	0.091	0.029	0.012	0.004	0.007	0.002	-0.002	-0.003	-0.003

The dependent variable is $RR_{t+h} = \frac{P_{t+h}/CPI_{t+h} - P_t/CPI_t}{P_t/CPI_t}$, where P is WTI or Brent. Because the BW sentiment index and the Kilian global real activity index are not in dollars, we do not adjust them. Since the gasoline and heating oil spreads are log price differences, the real and nominal spreads are the same. We define the real return on the oil-company stocks as $RR_{ARCA,t} = \frac{ARCA_t/CPI_t - ARCA_{t-1}/CPI_{t-1}}{ARCA_{t-1}/CPI_{t-1}}$, the real exchange rate change as $\Delta RMCI_t = \frac{RMCI_t - RMCI_{t-h}}{RMCI_{t-h}}$ where $RMCI$ is the real trade weighted U.S. dollar index for major currencies (RMCI) from the Federal Reserve Bank of St. Louis, and the real interest rate change as $\Delta RRF_t = (RF_t - I_t) - (RF_{t-h} - I_{t-h})$ where I is the CPI inflation rate.. The t-statistics are based on Newey-West HAC standard errors with the lag parameter set equal to 12 for the monthly regression.

Table 4 Multivariate regressions for nominal oil prices

	Panel A: WTI								
	1	3	6	9	12	15	18	21	24
SENTO	-0.009	-0.033	-0.078	-0.128	-0.162	-0.175	-0.190	-0.207	-0.171
	(-1.52)	(-1.42)	(-1.72)	(-2.19)	(-2.18)	(-2.06)	(-1.95)	(-1.85)	(-1.58)
S_Gas	0.090	0.463	0.421	0.097	0.403	0.532	0.004	-0.197	0.665
	(1.06)	(1.68)	(1.48)	(0.33)	(1.07)	(1.41)	(0.01)	(-0.44)	(1.20)
S_Heating	0.028	0.048	0.108	0.111	-0.064	-0.396	-0.819	-0.782	-0.611
	(0.44)	(0.26)	(0.36)	(0.28)	(-0.12)	(-0.68)	(-1.56)	(-1.34)	(-0.93)
ΔKilian	0.002	-0.001	-0.001	0.000	0.001	0.003	0.002	0.001	0.002
	(0.69)	(-0.62)	(-0.60)	(0.13)	(0.36)	(1.35)	(0.73)	(0.56)	(0.86)
ΔMCI	-0.548	-0.433	0.004	0.365	0.389	0.388	0.332	0.394	0.428
	(-1.50)	(-1.13)	(0.01)	(0.54)	(0.53)	(0.54)	(0.40)	(0.43)	(0.46)
ΔRF	0.244	-0.109	-0.168	-0.041	0.078	0.160	0.195	0.042	-0.197
	(2.59)	(-0.58)	(-0.69)	(-0.13)	(0.24)	(0.56)	(0.74)	(0.15)	(-0.59)
R_{ARCA}	0.322	0.427	0.390	0.287	0.332	0.247	-0.008	-0.057	0.031
	(3.58)	(2.36)	(2.04)	(1.34)	(1.39)	(0.90)	(-0.02)	(-0.16)	(0.09)
R^2	0.089	0.072	0.038	0.033	0.053	0.064	0.054	0.047	0.051
	Panel B: Brent								
	1	3	6	9	12	15	18	21	24
SENTO	-0.011	-0.038	-0.089	-0.138	-0.172	-0.189	-0.209	-0.235	-0.196
	(-1.67)	(-1.69)	(-1.93)	(-2.22)	(-2.19)	(-2.15)	(-2.03)	(-1.97)	(-1.68)
S_Gas	0.064	0.412	0.367	0.020	0.251	0.589	0.359	0.185	0.902
	(0.82)	(1.39)	(1.10)	(0.06)	(0.66)	(1.27)	(0.67)	(0.36)	(1.43)
S_Heating	-0.012	-0.100	0.094	0.212	-0.137	-0.321	-0.415	-0.232	-0.212
	(-0.19)	(-0.67)	(0.36)	(0.62)	(-0.38)	(-0.91)	(-1.14)	(-0.52)	(-0.43)
ΔKilian	0.002	-0.002	-0.002	0.000	0.001	0.003	0.002	0.001	0.002
	(0.98)	(-0.96)	(-0.70)	(0.04)	(0.42)	(1.50)	(0.75)	(0.44)	(0.69)
ΔMCI	-0.581	-0.340	0.162	0.418	0.399	0.365	0.355	0.467	0.416
	(-1.41)	(-0.87)	(0.30)	(0.55)	(0.50)	(0.48)	(0.40)	(0.47)	(0.42)
ΔRF	0.202	-0.116	-0.092	-0.013	0.128	0.233	0.191	-0.025	-0.311
	(1.75)	(-0.56)	(-0.38)	(-0.04)	(0.38)	(0.78)	(0.68)	(-0.08)	(-0.80)
R_{ARCA}	0.456	0.535	0.498	0.417	0.466	0.271	0.073	0.111	0.121
	(4.44)	(2.78)	(2.32)	(1.77)	(1.72)	(0.88)	(0.20)	(0.28)	(0.30)
R^2	0.114	0.069	0.043	0.035	0.047	0.070	0.047	0.037	0.066

The dependent variable is $R_{t+h} = \frac{P_{t+h} - P_{t\,t}}{P_t}$, where *P* is WTI or Brent. *SENTO* is the BW sentiment index that is orthogonalized to the business cycle variables. We define the gasoline price spread as $S\_Gas_t = \log(P_{Gasoline,\,t}) - \log(P_t)$, the heating oil price spread as $S\_Heating_t = \log(P_{Heating\ oil,\,t}) - \log(P_t)$, the return on the oil-company stocks as $R_{ARCA,t} = \frac{ARCA_t - ARCA_{t-1}}{ARCA_{t-1}}$, the global real activity change as $\Delta Kilian_t = \frac{Kilian_t - Kilian_{t-h}}{Kilian_{t-h}}$, the exchange rate change as $\Delta MCI_t = \frac{MCI_t - MCI_{t-h}}{MCI_{t-h}}$, and the interest rate change as $\Delta RF_t = RF_t - RF_{t-h}$. The t-statistics are based on Newey-West HAC standard errors with the lag parameter set equal to 12 for the monthly regression.

Table 5 Multivariate regressions for real oil prices

	Panel A: WTI								
	1	3	6	9	12	15	18	21	24
SENTO	-0.012	-0.037	-0.081	-0.128	-0.151	-0.161	-0.171	-0.212	-0.220
	(-1.75)	(-1.87)	(-2.05)	(-2.35)	(-2.22)	(-2.20)	(-1.93)	(-2.03)	(-1.95)
S_Gas	0.073	0.296	0.337	0.245	0.681	0.962	0.770	0.610	1.190
	(1.19)	(1.74)	(1.68)	(0.96)	(2.17)	(3.84)	(3.15)	(1.73)	(3.30)
S_Heating	-0.042	-0.186	-0.108	0.096	-0.178	-0.377	-0.239	-0.041	-0.367
	(-0.77)	(-1.25)	(-0.54)	(0.34)	(-0.55)	(-1.24)	(-0.81)	(-0.15)	(-1.05)
ΔKilian	0.002	-0.001	-0.001	0.000	0.000	0.003	0.002	0.001	0.002
	(0.71)	(-0.32)	(-0.84)	(0.07)	(0.22)	(1.55)	(0.68)	(0.56)	(1.27)
ΔRMCI	-0.451	-0.238	0.036	0.263	0.063	0.050	0.044	0.382	0.593
	(-1.34)	(-0.73)	(0.08)	(0.42)	(0.09)	(0.08)	(0.06)	(0.44)	(0.67)
ΔRRF	0.009	-0.064	0.013	0.005	0.101	0.185	0.153	0.072	0.085
	(0.57)	(-1.19)	(0.33)	(0.13)	(2.43)	(3.03)	(2.71)	(1.14)	(1.38)
RR_{ARCA}	0.324	0.456	0.444	0.256	0.344	0.150	-0.153	-0.139	-0.145
	(3.86)	(2.15)	(2.27)	(1.11)	(1.43)	(0.60)	(-0.50)	(-0.45)	(-0.49)
R^2	0.065	0.080	0.059	0.080	0.133	0.164	0.109	0.084	0.137
	Panel B: Brent								
	1	3	6	9	12	15	18	21	24
SENTO	-0.014	-0.046	-0.098	-0.146	-0.169	-0.184	-0.207	-0.265	-0.287
	(-2.20)	(-2.50)	(-2.74)	(-2.77)	(-2.43)	(-2.45)	(-2.22)	(-2.35)	(-2.29)
S_Gas	0.091	0.422	0.514	0.320	0.738	1.126	1.002	0.836	1.527
	(1.37)	(1.99)	(2.43)	(1.11)	(2.11)	(3.85)	(3.27)	(2.07)	(3.85)
S_Heating	-0.034	-0.199	-0.072	0.239	-0.043	-0.270	-0.095	0.189	-0.112
	(-0.57)	(-1.11)	(-0.32)	(0.76)	(-0.13)	(-0.82)	(-0.30)	(0.67)	(-0.28)
ΔKilian	0.003	-0.001	-0.002	-0.000	0.000	0.003	0.001	0.000	0.002
	(1.07)	(-0.60)	(-0.83)	(-0.14)	(0.24)	(1.71)	(0.59)	(0.17)	(0.98)
ΔRMCI	-0.496	-0.193	0.081	0.254	0.041	0.008	0.040	0.436	0.657
	(-1.34)	(-0.53)	(0.17)	(0.41)	(0.06)	(0.01)	(0.05)	(0.47)	(0.68)
ΔRRF	0.008	-0.038	0.047	0.031	0.123	0.175	0.156	0.059	0.060
	(0.39)	(-0.79)	(1.25)	(0.72)	(2.76)	(2.38)	(2.79)	(0.83)	(0.92)
RR_{ARCA}	0.452	0.515	0.423	0.304	0.404	0.151	-0.162	-0.053	-0.165
	(4.75)	(2.18)	(1.92)	(1.21)	(1.55)	(0.55)	(-0.47)	(-0.15)	(-0.51)
R^2	0.108	0.097	0.103	0.117	0.157	0.203	0.167	0.151	0.224

The dependent variable is $RR_{t+h} = \frac{P_{t+h}/CPI_{t+h} - P_t/CPI_t}{P_t/CPI_t}$, where *P* is WTI or Brent. Because the BW sentiment index and the Kilian global real activity index are not in dollars, we do not adjust them. Since the gasoline and heating oil spreads are log price differences, the real and nominal spreads are the same. We define the real return on the oil-company stocks as $RR_{ARCA,t} = \frac{ARCA_t/CPI_t - ARCA_{t-1}/CPI_{t-1}}{ARCA_{t-1}/CPI_{t-1}}$, the real exchange rate change as $\Delta RMCI_t = \frac{RMCI_t - RMCI_{t-h}}{RMCI_{t-h}}$ where *RMCI* is the real trade weighted U.S. dollar index for major currencies (RMCI) from the Federal Reserve Bank of St. Louis, and the real interest rate change as $\Delta RRF_t = (RF_t - I_t) - (RF_{t-h} - I_{t-h})$ where *I* is the CPI inflation rate. The t-statistics are based on Newey-West HAC standard errors with the lag parameter set equal to 12 for the monthly regression.

Table 3 presents the results for real oil prices, which are similar as those for nominal oil prices. First, real financial variables such as real returns of oil-company stocks and real interest rate changes still have significant power to explain real oil price movements. Second, high sentiment predicts low future real oil returns particularly at longer horizons. Third, sentiment has reasonable explanatory power, in particular for longer horizons. The evidence thus suggests that sentiment helps understand the fluctuations in oil prices.

3.2 Multivariate regressions

Univariate regressions examine if sentiment has any explanatory power. But the explanatory power of sentiment can be due to its correlation with other explanatory variables. To investigate the marginal explanatory power of investor sentiment, we run multivariate regressions that include all explanatory variables. The t-statistics are based on Newey-West HAC standard errors with the lag parameter set equal to 12 for the monthly regressions.

These results are reported in Table 4 for nominal oil prices. Even with the presence of other explanatory variables, sentiment is still a statistically significant factor to explain the changes in nominal oil prices, in particular at longer horizons. Furthermore, its explanatory power tends to increase monotonically with the time horizon. For instance, for the nominal WTI, the coefficients on the sentiment index are -0.078 (t = -1.72), -0.162 (t = -2.18), and -0.190 (t = -1.95) at horizons of 6 months, 12 months and 15 months, respectively. A comparison of the coefficient estimates in Table 4 with those in Table 2 suggests that sentiment is a complementary factor relative to other explanatory variables, because its coefficients do not tend to decrease even with the presence of other explanatory variables.

Table 5 presents the results for real oil prices, which are generally consistent with those for nominal oil prices. Sentiment is a statistically significant factor to explain the fluctuations in real oil prices even with the presence of other explanatory variables. Furthermore, its explanatory power tends to increase with the time horizon. For instance, for the real Brent, the coefficients on the sentiment index are -0.014 (t = -2.20), -0.098 (t = -2.74), -0.169 (t = -2.43), -0.207 (t = -2.22), and -0.287 (t = -2.29) at horizons of one month, 6 months, 12 months, 18 months, and 24 months, respectively. Thus, the evidence suggests that investor sentiment does have marginal explanatory power. This is the central finding of the paper. Because the results based on Brent are always consistent with those based on WTI, to save space, we only report the results based on WTI in the rest of the paper.

One potential concern is multicollinearity, which can result in unreliable parameter estimates. To address this concern, we examine the correlations among the explanatory variables in our regressions and report the results in Table 6. Panel A shows the correlations among the explanatory variables used in the nominal oil price regressions. To save space, we only report the correlations at three forecast horizons,

Table 6 Correlations among the oil price determinants

	Panel A: Nominal oil price determinants							Panel B: Real oil price determinant					
	h = 1							h = 1					
	SENTO	S_Gas	S_Heating	ΔKilian	ΔMCI	ΔRF		SENTO	S_Gas	S_Heating	ΔKilian	ΔRMCI	ΔRRF
S_Gas	-0.08						S_Gas	-0.08					
S_Heating	0.00	-0.02					S_Heating	0.00	-0.02				
ΔKilian	-0.01	0.00	0.07				ΔKilian	-0.01	0.00	0.07			
ΔMCI	0.12	0.06	0.06	0.03			ΔRMCI	0.12	0.08	0.05	0.03		
ΔRF	-0.06	0.04	-0.06	0.09	0.07		ΔRRF	0.02	-0.12	-0.05	0.01	0.16	
R_{ARCA}	-0.04	0.07	-0.05	0.05	-0.20	0.05	RR_{ARCA}	-0.04	0.06	-0.05	0.05	-0.20	-0.09
	h = 12							h = 12					
	SENTO	S_Gas	S_Heating	ΔKilian	ΔMCI	ΔRF		SENTO	S_Gas	S_Heating	ΔKilian	ΔRMCI	ΔRRF
S_Gas	-0.08						S_Gas	-0.08					
S_Heating	0.00	-0.03					S_Heating	0.00	-0.03				
ΔKilian	0.05	0.02	0.05				ΔKilian	0.05	0.02	0.05			
ΔMCI	0.33	0.08	0.13	0.04			ΔRMCI	0.38	0.07	0.14	0.04		
ΔRF	-0.06	0.17	0.02	0.00	-0.01		ΔRRF	-0.04	-0.08	0.17	-0.03	0.27	
R_{ARCA}	-0.04	0.08	-0.05	0.04	-0.06	0.10	RR_{ARCA}	-0.05	0.07	-0.05	0.04	-0.04	-0.05
	h = 24							h = 24					
	SENTO	S_Gas	S_Heating	ΔKilian	ΔMCI	ΔRF		SENTO	S_Gas	S_Heating	ΔKilian	ΔRMCI	ΔRRF
S_Gas	-0.07						S_Gas	-0.07					
S_Heating	0.01	-0.03					S_Heating	0.01	-0.03				
ΔKilian	0.02	-0.02	0.01				ΔKilian	0.02	-0.02	0.01			
ΔMCI	0.42	0.03	-0.05	0.09			ΔRMCI	0.48	0.03	-0.02	0.10		
ΔRF	0.14	0.20	0.09	0.00	-0.05		ΔRRF	0.02	-0.03	0.07	0.00	0.22	
R_{ARCA}	-0.05	0.09	-0.05	-0.03	-0.09	0.14	RR_{ARCA}	-0.05	0.07	-0.05	-0.03	-0.07	0.07

Panel A shows the correlations among the explanatory variables used in the nominal oil price regressions, while Panel B presents those employed in the real oil price regressions. To save space, we only report the correlations at three forecast horizons, namely one month (h = 1), one year (h = 12), and two years (h = 24). *SENTO* is the BW sentiment index, $S\_Gas_t = \log(P_{Gasoline,\,t}) - \log(P_t)$, $S\_Heating_t = \log(P_{Heating\ oil,\,t}) - \log(P_t)$, $R_{ARCA,t} = \frac{ARCA_t - ARCA_{t-1}}{ARCA_{t-1}}$, $\Delta Kilian_t = \frac{Kilian_t - Kilian_{t-h}}{Kilian_{t-h}}$, $\Delta MCI_t = \frac{MCI_t - MCI_{t-h}}{MCI_{t-h}}$, $\Delta RF_t = RF_t - RF_{t-h}$, $RR_{ARCA,t} = \frac{ARCA_t / CPI_t - ARCA_{t-1} / CPI_{t-1}}{ARCA_{t-1} / CPI_{t-1}}$, $\Delta RMCI_t = \frac{RMCI_t - RMCI_{t-h}}{RMCI_{t-h}}$ where *RMCI* is the real trade weighted U.S. dollar index for major currencies (RMCI) from the Federal Reserve Bank of St. Louis, and $\Delta RRF_t = (RF_t - I_t) - (RF_{t-h} - I_{t-h})$ where *I* is the CPI inflation rate.

Table 7 Quantile regressions

Panel A: Quantile regressions for nominal WTI

	1	3	6	9	12	15	18	21	24
$\tau = 0.25$									
SENTO	-0.003	-0.026	-0.071	-0.092	-0.189	-0.227	-0.230	-0.256	-0.111
	(-0.26)	(-1.06)	(-1.99)	(-2.18)	(-3.77)	(-3.88)	(-3.58)	(-3.81)	(-1.51)
$\tau = 0.50$									
SENTO	-0.005	-0.033	-0.048	-0.157	-0.164	-0.128	-0.117	-0.105	-0.083
	(-0.56)	(-1.65)	(-1.64)	(-4.56)	(-3.93)	(-2.69)	(-2.31)	(-1.90)	(-1.39)
$\tau = 0.75$									
SENTO	-0.016	-0.037	-0.092	-0.217	-0.216	-0.194	-0.154	-0.186	-0.153
	(-1.42)	(-1.62)	(-2.63)	(-5.22)	(-4.55)	(-3.57)	(-2.72)	(-3.02)	(-2.20)

Panel B: Quantile regressions for real WTI

	1	3	6	9	12	15	18	21	24
$\tau = 0.25$									
SENTO	-0.005	-0.023	-0.060	-0.096	-0.081	-0.189	-0.188	-0.285	-0.233
	(-0.46)	(-1.00)	(-1.86)	(-2.50)	(-1.77)	(-3.58)	(-3.19)	(-4.49)	(-3.57)
$\tau = 0.50$									
SENTO	-0.007	-0.024	-0.055	-0.117	-0.115	-0.093	-0.093	-0.096	-0.194
	(-0.69)	(-1.21)	(-1.99)	(-3.55)	(-2.91)	(-2.08)	(-1.92)	(-1.84)	(-3.53)
$\tau = 0.75$									
SENTO	-0.016	-0.016	-0.082	-0.124	-0.105	-0.154	-0.112	-0.160	-0.158
	(-1.45)	(-0.74)	(-2.57)	(-3.26)	(-2.34)	(-3.15)	(-2.08)	(-2.71)	(-2.51)

In Panel A, the dependent variable is $R_{t+h} = \frac{P_{t+h} - P_{t\,t}}{P_t}$, where *P* is WTI. *SENTO* is the BW sentiment index that is orthogonalized to the business cycle variables. In Panel B, the dependent variable is $RR_{t+h} = \frac{P_{t+h} / CPI_{t+h} - P_t / CPI_t}{P_t / CPI_t}$, where *P* is WTI and CPI is the U.S. CPI.

Table 8 Further Evidence

Panel A: Alternative sentiment index and oil prices									
	Nominal WTI								
	1	3	6	9	12	15	18	21	24
SENT	-0.008	-0.035	-0.077	-0.126	-0.155	-0.164	-0.168	-0.187	-0.173
	(-1.16)	(-1.50)	(-1.73)	(-2.23)	(-2.06)	(-1.91)	(-1.83)	(-1.80)	(-1.62)
R^2	0.088	0.074	0.040	0.037	0.054	0.063	0.047	0.040	0.051
	Real WTI								
	1	3	6	9	12	15	18	21	24
SENT	-0.009	-0.032	-0.070	-0.115	-0.131	-0.134	-0.135	-0.174	-0.205
	(-1.18)	(-1.50)	(-1.77)	(-2.16)	(-1.92)	(-1.85)	(-1.65)	(-1.81)	(-1.89)
R^2	0.063	0.077	0.055	0.075	0.126	0.155	0.096	0.068	0.131

Panel B: Investor sentiment and gasoline, heating oil, and oil-company stock prices							
Horizon		Nominal prices			Real prices		
		Gas	Heating	ACRA	Gas	Heating	ACRA
1	SENTO	-0.018	-0.015	-0.003	-0.017	-0.014	-0.002
		(-2.21)	(-1.86)	(-0.62)	(-2.19)	(-1.84)	(-0.49)
	R^2	0.004	0.004	-0.003	0.004	0.003	-0.003
6	SENTO	-0.088	-0.092	-0.031	-0.084	-0.088	-0.027
		(-2.32)	(-2.54)	(-1.54)	(-2.27)	(-2.49)	(-1.38)
	R^2	0.030	0.034	0.011	0.029	0.033	0.008
12	SENTO	-0.152	-0.153	-0.089	-0.142	-0.144	-0.081
		(-2.24)	(-2.17)	(-2.24)	(-2.14)	(-2.09)	(-2.06)
	R^2	0.054	0.044	0.053	0.051	0.042	0.047
18	SENTO	-0.161	-0.198	-0.172	-0.146	-0.182	-0.157
		(-1.77)	(-1.86)	(-2.93)	(-1.66)	(-1.77)	(-2.76)
	R^2	0.039	0.048	0.129	0.036	0.046	0.121
24	SENTO	-0.151	-0.211	-0.246	-0.132	-0.189	-0.223
		(-1.35)	(-1.68)	(-3.04)	(-1.25)	(-1.58)	(-2.90)
	R^2	0.028	0.046	0.176	0.024	0.042	0.168

In Panel A, we report the results based on the alternative sentiment index. For the nominal WTI regressions, the dependent variable is $R_{t+h} = \frac{P_{t+h} - P_{t\,t}}{P_t}$, where *P* is nominal WTI, *SENT* is the BW sentiment index that is not orthogonalized to the business cycle variables. For the real WTI regressions, the dependent variable is $RR_{t+h} = \frac{P_{t+h} / CPI_{t+h} - P_t / CPI_t}{P_t / CPI_t}$, where *P* is nominal WTI and CPI is the consumer price index.

In Panel B, we run the following regressions for nominal gasoline, heating oil, and oil-company stock prices.

$$R'_{t+h} = \alpha_h + \beta_h SENTO_t + \varepsilon_{t+h}$$

where $R'_{t+h} = \frac{P_{Gasoline,t+h} - P_{Gasoline,t}}{P_{Gasoline,t}}$, $\frac{P_{Heating,t+h} - P_{Heating,t}}{P_{Heating,t}}$, or $R'_{t+h} = \frac{ARCA_{t+h} - ARCA_t}{ARCA_t}$, and *SENTO* is the BW sentiment index that is orthogonalized to the business cycle variables. We deflate nominal gasoline, heating oil, and oil-company stock prices by the U.S. CPI to run the corresponding real price regressions. . The t-statistics are based on Newey-West HAC standard errors with the lag parameter set equal to 12 for the monthly regression.

namely one month (h = 1), one year (h = 12), and two years (h = 24). As we can see, all the correlations are less than 0.5, suggesting that there is no severe multicollinearity. In Panel B, we report the correlations among the explanatory variables employed in the real oil price regressions. Again, the results imply that severe multicollinearity is not present in our regressions.

4 Quantile regressions

The least squares regression results reported in Tables 2 to 5 provide estimates of the *average* effects of the various independent variables on the dependent variable, the future changes in oil prices. They depict the impact of the independent variables (e.g., sentiment) on the dependent variable near the center of the dependent variable conditional distribution. However, the effects of the various economic variables on oil price changes may not be the same across different portions of the conditional distribution of oil price changes. For instance, when oil prices experience dramatic movements, sentiment may play a more or less significant role compared to modest prices movements. Least-squares regressions are incapable of revealing such heterogeneity. Quantile regression, in contrast, is designed for identifying such differential effects (Koenker & Bassett, 1978; Du *et al*, 2013). In addition, median QR is also more robust to outliers than least squares regression, and is semi parametric as it avoids assumptions about the parametric distribution of the error process. Therefore, we perform QR tests in this section.

The multivariate quantile regression model for nominal oil prices can be specified as:

$$R_{t+h} = \alpha_h^{\tau} + \beta_h^{\tau} SENTO_t + \sum_k \gamma_{kh}^{\tau} X_{k,t} + \varepsilon_{t+h} \qquad (3)$$

where α^{τ}'s, β^{τ}'s and γ^{τ}'s are the τ-th quantile regression coefficients. The quantile regression coefficient for a particular τ measures the impact of a one unit change in the corresponding independent variable on the τ-th quantile of the dependent variable holding constant the effects of all the other independent variables. The model for real oil prices can be specified in the same way. We focus on τ = 0.25, 0.50, and 0.75 for the ease of exposition. Note that all data points are used in estimating the quantile regressions. That is, 25% of all the data points will fall below the τ = 0.25 quantile regression hyperplane while 50% will fall below the τ = 0.50 quantile regression hyperplane, and so forth. Hence, the median (τ = 0.50) quantile regression hyperplane bisects all the data points into two halves, each conditioned on the included independent variables.

The QR results for the nominal WTI are presented in Panel A of Table 7.[6] First, the median (τ = 0.50) QR results suggest that although our OLS results are robust to outliers, outliers do affect coefficient estimates. Specifically, although the coefficient estimates on the sentiment measure for τ = 0.50 are still

[6] To save space, we don't report the coefficient estimates for other explanatory variables in Table 7. The results are available from the authors upon request.

significantly negative at longer horizons, they tend to be different from OLS estimates. For instance, the QR estimates at horizons of 6 months, 12 months and 18 months are -0.048 (t = -1.64), -0.164 (t = -3.93), and -0.117 (t = -2.31), while those for OLS regressions from Table 4 are -0.078 (t = -1.72), -0.162 (t = -2.18), and -0.190 (t = -1.95). Second, the impact of sentiment on oil prices seems to be different across different portions of the conditional distribution of the nominal WTI change, particularly at longer horizons. For instance, at horizon of 18 months, the QR coefficient estimates on sentiment are -0.230 (t = -3.58), -0.117 (t = -2.31), and -0.154 (t = -2.72) for $\tau = 0.25$, 0.50 and 0.75, respectively. The evidence suggests that the effects of sentiment on the nominal WTI seem to be stronger when movements in nominal WTI are near the two tails of the conditional distribution, particularly at longer horizons. The key take-away is that quantile regressions shed new insight on the effects of sentiment on the nominal WTI.

The results for the real WTI are reported in Panel B of Table 7, and are consistent with those based on the nominal WTI. First, the median QR results suggest that although our OLS results are robust to outliers, outliers do affect coefficient estimates. Second, the effects of sentiment on the real WTI seem to be stronger when movements in real WTI are near the two tails of the conditional distribution, particularly at longer horizons.

5 Further evidence

5.1 Alternative sentiment index

We use the BW sentiment index that is orthogonalized to the business cycle variables (SENTO) in Sections 3 and 4. This index, in theory, is a "pure" sentiment index. In contrast, SENT is not orthogonalized to the business cycle variables. Thus, a high value of SENT may not imply high sentiment. As a result, the subsequent mispricing correction may be weaker. The weaker results associated with SENT, thus, provide additional evidence that it is sentiment that drives price reversals. However, given the strong correlation between SENTO and SENT that appears in Panel C of Figure 1, and also in Table 1 (the correlation is 0.95), we should not expect dramatically different results.

Panel A of Table 8 reports the results for the nominal and real WTI based on SENT.[7] The t-statistics are based on Newey-West HAC standard errors with the lag parameter set equal to 12 for the monthly regressions. Consistent with our conjecture, although SENT still has significant power to explain the fluctuations in the nominal and real WTI particularly at longer horizons, its coefficients tend to be smaller than those based on SENTO. For instance, for the nominal WTI, the coefficient estimates on SENT at horizons 6 months, 12 months and 18 months are -0.077 (t = -1.73), -0.155 (t = -2.06) and -0.168

[7] To save space, we don't report the coefficient estimates for other explanatory variables in Table 8. The results are available from the authors upon request.

(t = -1.83), while those based on SENTO in Table 4 are -0.078 (t = -1.72), -0.162 (t = -2.18) and -0.190 (t = -1.95). Thus, these results suggest that it is sentiment that drives oil price reversals at longer horizons.

5.2 Gasoline, heating oil, and oil-company stock prices

Table 1 suggests that oil prices are strongly correlated with gasoline, heating oil, and oil-company stock prices. Thus, if sentiment helps explain the movements in oil prices, it should help explain the movements in gasoline, heating oil, and oil-company stock prices too. This again would provide addition evidence that sentiment matters for oil prices. To test our conjecture, we run the following simple regressions for nominal gasoline, heating oil, and oil-company stock prices.

$$R'_{t+h} = \alpha_h + \beta_h SENTO_t + \varepsilon_{t+h} \quad (4)$$

where $R'_{t+h} = \frac{P_{Gasoline,t+h} - P_{Gasoline,t}}{P_{Gasoline,t}}$, $\frac{P_{Heating,t+h} - P_{Heating,t}}{P_{Heating,t}}$, or $R'_{t+h} = \frac{ARCA_{t+h} - ARCA_t}{ARCA_t}$. We also deflate nominal gasoline, heating oil, and oil-company stock prices by the U.S. CPI to run the corresponding real price regressions. If high sentiment leads to overvaluation and subsequent mispricing correction in oil prices, we expect similar relationship in gasoline, heating oil, and oil-company stock prices.

The results are reported in Panel B of Table 8. The t-statistics are based on Newey-West HAC standard errors with the lag parameter set equal to 12 for the monthly regressions. In general, the BW sentiment index significantly predicts future price reversals across all three prices, particularly at longer horizons. For instance, for nominal gasoline, heating oil, and oil-company stock prices, at a horizon of 12 months, the sentiment coefficient estimates are -0.152 (t = -2.24), -0.153 (t = -2.17), and -0.089 (t = -2.24), respectively. Thus, the evidence supports the notion that sentiment matters for oil prices.

6 Conclusions

Oil prices play a critical role in the global economy. Thus, it is important to understand the determinants of oil prices. Although both theoretical models (e.g., Kilian, 2009) and anecdotal accounts (e.g., Masters, 2008) suggest that investor sentiment in financial markets may be a potential determinant of oil prices, there has been no empirical research addressing this question directly. We fill the gap. Our findings can be easily summarized. Investor sentiment significantly explains the movements in oil prices (as well as gasoline, heating oil and oil-company stock prices). High/low sentiment predicts subsequent low/high oil returns at horizons from six months to two years. Our findings have important theoretical as well as practical implications. In terms of theoretical implications, our findings suggest that future theoretical models of oil prices should take into account both fundamentals and investor sentiment. In terms of practical implications, our findings imply a new predictor of oil prices.

References

Alquist, R., and Kilian, L. (2010) What do we learn from the price of crude oil futures?, *Journal of Applied Econometrics* 25: 539-573.

Arezki, R., Loungani, P., van der Ploeg, R. and Venables, A. J. (2014) Understanding international commodity price fluctuations. *Journal of International Money and Finance* 42: 1-8.

Baker, M. and Wurgler, J. (2006) Investor sentiment and the cross-section of stock returns. *Journal of Finance* 61: 1645–80.

Baker, M. and Wurgler J. (2007) Investor sentiment in the stock market. *Journal of Economic Perspectives* 21: 129–151.

Baker, M., Wurgler, J. and Yuan, Y. (2012) Global, local, and contagious investor sentiment. *Journal of Financial Economics* 104: 272-287.

Baumeister, C. and Kilian, L. (2012) Real-time forecasts of the real price of oil. *Journal of Business and Economic Statistics* 30: 326-336.

Baumeister, C. and Kilian, L. (2014). What central bankers need to know about forecasting oil prices. *International Economic Review* 55, 869-889.

Baumeister, C. and Kilian, L. (2015) Forecasting the real price of oil in a changing world: a forecast combination approach. *Journal of Business and Economic Statistics* 33: 338-351.

Baumeister, C., Guérin, P. and Kilian, L. (2015), Do high-frequency financial data help forecast oil prices? The MIDAS Touch at Work. *International Journal of Forecasting* 31: 238-252.

Brunnermeier, M. and Pedersen, L. (2005) Predatory trading. *Journal of Finance* 60: 1825–1863.

Büyükşahin, B. and Robe, M. A. (2014) *Speculators, commodities and cross-market linkages*. *Journal of International Money and Finance* 42: 38-70.

Chen, S. S. (2014) Forecasting crude oil price movements with oil-sensitive stocks. *Economic Inquiry* 52: 830-844.

Chen, Y-C., Rogoff, K. S., and Rossi, B. (2010) Can exchange rates forecast commodity prices?. *Quarterly Journal of Economics* 125: 1145-1194.

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

Du, D. (2014) Persistent exchange rate movements and stock returns. *Journal of International Financial Markets, Institutions & Money* 28: 36–53.

Du, D. and Hu, O. (2012) Foreign exchange volatility and stock returns. *Journal of International Financial Markets, Institutions & Money* 22: 1202– 1216.

Du, D. and Hu, O. (2014) Cash flow, currency risk, and the cost of capital. *Journal of Financial Research* 37: 139–158.

Du, D., Ng, P. and Zhao, X. (2013) Measuring currency exposure with quantile regression. *Review of Quantitative Finance and Accounting* 41: 549-566.

Fattouh, B., Mahadeva, L. and Kilian, L. (2013) The role of speculation in oil markets: what have we learned so far?. *Energy Journal* 34: 7-33.

Hamilton, J. D. (2009) Understanding crude oil prices. *Energy Journal* 30: 179-206.

Kilian, L. (2009) Not all oil price shocks are alike: disentangling demand and supply shocks in the crude oil market. *American Economic Review* 99: 1053-1069.

Kilian, L. and Murphy, D. P. (2014) The role of inventories and speculative trading in the global market for crude oil. *Journal of Applied Econometrics* 29: 454-478.

Koenker, R., and Bassett, G. (1978) Regression quantiles. *Econometrica* 46: 33-50.

Lamont, O. A. and Thaler, R. H. (2003) Can the market add and subtract? Mispricing and tech stock carve-outs. *Journal of Political Economy* 111: 227–268.

Masters, M. W. (2008) Testimony before the U.S. Senate Committee on Homeland Security and Governmental Affairs, May 20.

Mclean, R. D. and Zhao, M. (2014) The business cycle, investor sentiment, and costly external finance. *Journal of Finance* 69: 1377–1409.

Neely, C.J., Rapach, D.E., Tu, J. and Zhou, G. (2014) Forecasting the equity risk premium: the role of technical indicators. *Management Science* 60(7): 1772-1791.

Reeve, T. A., and Vigfusson, R. J. (2011) Evaluating the forecasting performance of commodity futures prices, International Finance Discussion Paper No. 1025, Board of Governors of the Federal Reserve System.

Shleifer, A. and Vishny, R. (1997) The limits of arbitrage. *Journal of Finance* 52: 35–55.

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

Wurgler, J. and Zhuravskaya, K. (2002) Does arbitrage flatten demand curves for stocks?. *Journal of Business* 75: 583–608.