Heterogeneous Beliefs and the Oil State Price Density*

Peter Christoffersen[†]
University of Toronto,
CBS and CREATES

Xuhui (Nick) Pan[‡]
Freeman School of Business,
Tulane University

October 8, 2015

Abstract

The oil state price density (SPD) is strongly U-shaped: Investors assign high state prices to both large increases and decreases in the oil price. We use data from the crude oil derivatives market, in which speculation data is available and short selling is not restricted, to document how the SPD varies with proxies for investor beliefs. Investors assign higher state prices to negative oil returns when demand for out-of-the-money puts is high, and when reported speculation is high during the post-financialization period. Higher state prices of negative (positive) returns are associated with higher trading volumes in out-of-the-money put (call) options.

JEL Classification: G12, G13

Keywords: State price density; skewness; investor beliefs; crude oil; speculator and hedger positions.

*We would like to thank Hitesh Doshi, Stephen Figlewski (CICF 2015 discussant), Kris Jacobs (2015 OU Energy Finance Research Conference discussant), Sang Baum Kang, Bingxin Li, Scott Linn (NFA 2012 discussant), Michel Robe, Sergei Sarkissian, Kenneth Singleton, Lars Stentoft (EFA 2012 discussant), Feng Zhao, and seminar participants at the Bank of Canada and Tulane University for their comments.

†Peter Christoffersen, Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, Ontario, Canada, M5P 3E6; Tel: (416) 946-5511; Fax: (416) 971-3048; E-mail: peter.christoffersen@rotman.utoronto.ca.

‡Xuhui (Nick) Pan, Freeman School of Business, Tulane University, 7 McAlister Dr., New Orleans, LA 70123; Tel: (504) 314-7031; Fax: (504) 862-8327; E-mail: xpan@tulane.edu.

Heterogeneous Beliefs and the Oil State Price Density

October 8, 2015

Abstract

The oil state price density (SPD) is strongly U-shaped: Investors assign high state prices to both large increases and decreases in the oil price. We use data from the crude oil derivatives market, in which speculation data is available and short selling is not restricted, to document how the SPD varies with proxies for investor beliefs. Investors assign higher state prices to negative oil returns when demand for out-of-the-money puts is high, and when reported speculation is high during the post-financialization period. Higher state prices of negative (positive) returns are associated with higher trading volumes in out-of-the-money put (call) options.

JEL Classification: G12, G13

Keywords: State price density; skewness; investor beliefs; crude oil; speculator and hedger positions.

1 Introduction

The state price density contains important information about investor preferences and determine expected returns and risk premia. From an assumption of no-arbitrage, we can price any asset as long as we know the SPD and the final payoff of the asset. Although several studies have estimated SPDs using equity market data,[1] few papers investigate other financial markets.[2] The crude oil market is the largest and most liquid commodity derivatives market and it is clearly important for the overall economy. We therefore estimate oil SPDs and use them to analyze investor preferences towards different states of the economy.

We estimate SPDs using crude oil futures and options data from 1990 to 2012 and discover that SPDs in the crude oil market display a time varying asymmetric U-shape pattern: Investors assign high state prices (per unit probability) to both large negative and large positive returns. This implies that either extreme low or high oil futures prices correspond to bad states of the economy. The slope of the U-shaped SPDs in both the decreasing and increasing regions also varies depending on the market condition. Investors' marginal value of payoffs at large negative and positive oil returns, exhibits significant variation across time. We also document that average returns on out-of-the-money (OTM) oil call and put options are negative, which is consistent with the SPDs having both decreasing and increasing regions and supports the theory in Bakshi, Madan, and Panayotov (2010).

A strand of the literature has argued that SPDs depend on differences in investor beliefs, and that this heterogeneity affects expected returns and the price of risk (e.g., Anderson, Ghysels, and Juergens, 2005; Beber, Buraschi, and Breedon, 2010). In particular, Bakshi, Madan, and Panayotov (2010) advocate that the pattern of the SPDs implied by index options is compatible with a theory in which risk-averse investors have heterogeneous beliefs and are able to short sell equities.

[1] An incomplete list includes Aït-Sahalia and Lo (2000), Jackwerth (2000), Rosenberg and Engle (2002), Chernov (2003), Chabi-Yo (2012), Christoffersen, Heston, and Jacobs (2013), Linn, Shive and Shumway (2014), and Song and Xiu (2014).

[2] Notable exceptions include Beber and Brandt (2006) and Li and Zhao (2009) who study SPDs in the fixed income market, and Kitsul and Wright (2013) who estimate SPDs from inflation options.

Heterogeneity of investor beliefs is difficult to measure precisely but the crude oil market provides an excellent laboratory to examine the dependence of SPDs on investor beliefs. First, data on speculators' positions are available in this market. The level of speculation can be interpreted as a measure of heterogeneous beliefs because speculators usually bet on certain price movements, and disagreements among investors induce speculative trading (e.g., Scheinkman and Xiong, 2003). Second, there are no restrictions on short sales in this market. While some investors trade crude oil futures to hedge risks according to an underlying demand or supply of crude oil, others may take positions simply based on beliefs about futures prices. We therefore test whether investor beliefs affect the SPDs. Third, the crude oil derivatives market is highly liquid, and historical data are now available for more than twenty years. We have a rich cross-section of futures and option prices, which allows us to accurately extract SPDs. Moreover, data on trading volumes and open interests of both futures and options enable us to construct various measures of investor belief heterogeneity as suggested in the literature (e.g., Kandel and Pearson, 1995; Buraschi and Jiltsov, 2006).

We investigate in detail if investor beliefs about futures prices embedded in trading activities in crude oil futures and options affect the slope of the SPDs, which is one of the fundamental characteristics of the SPDs. The slope is related to investors' risk aversion (Rosenberg and Engle, 2002). It also compares the marginal value of payoffs in different economic states, as measured by the level of futures returns. The SPDs are estimated as the log ratio of the risk-neutral and physical densities, the asymmetric U-shape pattern of the SPDs thus arises from the difference between risk-neutral skewness and physical skewness. As physical densities of oil futures returns are relatively symmetric, the slope of the SPDs are primarily characterized by risk-neutral skewness (Han, 2008). If risk-neutral skewness is more negative (positive), the decreasing (increasing) region of the SPDs has a steeper slope, and investors assign higher state prices to more negative (positive) returns. We first provide evidence that a direct measure of the slope of the SPDs, defined as the difference between the value of the SPDs of two return points, is affected by investor beliefs reflected in options and futures trades. However, this measure of slope is noisy in the distribution tails. We therefore focus on risk-neutral skewness, which provides a more reliable and comprehensive measure of variations in the marginal rate of substitution across states. We calculate risk-neutral skewness from two distinct approaches and regress it on investor beliefs. We also investigate whether the SPDs are affected by speculative activities in the crude oil market after the financialization of the commodity market. The selection of the sub-sample period is based on the evidence by Tang and Xiong (2012), who document the financialization of commodity markets beginning around 2004-2005 and how speculative investments affect commodity futures prices.[3]

Our empirical results indicate that the slope of the U-Shaped SPD is affected by various measures of investors beliefs. When the demand for OTM puts is high, the risk-neutral distribution is more negatively skewed. As such, the side of the SPD corresponding to negative oil returns has a steeper slope and investors assign higher state prices to more negative returns. Investor beliefs embedded in options trading volume also affect the shape of the SPDs. When OTM put (call) trading volume is high, the slope of the decreasing-return (increasing-return) side of the SPD is steeper, i.e. the marginal rate of substitution across negative (positive) oil futures returns is higher. We find that a high level of speculation after the financialization of the commodity market is associated with a steeper slope in the decreasing region of the SPD and investors overall are worried more about large negative returns. Our findings lend support to the argument from Singleton (2014) that flows from financial investors and speculative activities have had significant effects on the crude oil futures market during the post-financialization period.

We also investigate the impact of investor beliefs from the equity market and the aggregate economy on the SPDs. Although investors' expectations about the equity market significantly affect index option prices and return distributions in the equity market (Han, 2008), bearish stock market beliefs do not imply more negative risk-neutral skewness of crude

[3]More recent contributions on the financialization of commodity markets include Hamilton and Wu (2014), Cheng and Xiong (2014), Henderson, Pearson and Wang (2014), Ready (2015) and others.

oil futures returns, and do not have significant impact on the slope of the SPDs in the crude oil market. It is the belief dispersions from investors in the crude oil market, rather than investor beliefs about the equity market, that significantly affect state prices in the crude oil market. This finding aligns with Goldstein, Li and Yang (2013), who argue that although information is integrated and fast moving, financial markets can be relatively segmented due to the specialization and friction of investments.

This paper is part of a growing list of recent studies that examine how the activities of investors in the commodity market, both hedgers and speculators, affect futures prices and returns. Hamilton and Wu (2014) document significant changes in oil futures risk premia due to investments from financial institutions in recent years. Motivated by the coincident price rise and increased financial participation in the crude oil market, Büyükşahin and Harris (2011) analyze whether the crude oil price is driven by hedge funds and other speculators. Acharya, Lochstoer, and Ramadorai (2013) find that producers' hedging demand, captured by their default risk, predicts commodity returns. Hong and Yogo (2012) show that the high level of commodity market activity, measured by high growth in open-interest, predicts high commodity returns. Etula (2013) finds that the supply of speculator capital, captured by changes in broker-dealer balance sheets, predicts commodity returns, especially in energy commodities. Our paper examines how investor beliefs embedded in derivatives trading activities affect the SPDs, which determine not only the commodity futures prices and returns, but also risk premia.

The remainder of the paper proceeds as follows. Section 2 provides our methodological framework for extracting oil SPDs. Section 3 describes the data, SPD estimates, option returns, and skewness measures. Section 4 analyzes how the oil SPDs are affected by heterogeneity in investor beliefs, and Section 5 concludes.

2 The Oil Market State Price Density

In this section, we first discuss the economic framework to obtain the SPDs in the crude oil market based on the no-arbitrage principle. We then discuss the estimation methodology of the risk-neutral densities, physical densities and the SPDs using futures and option prices. Next we describe futures and option data and present the estimated SPDs, as well as option implied moments.

2.1 Economic Framework

If we denote the SPD by ξ, based on the no-arbitrage principle, as long as we know the final payoff p_T of an asset, the price of the asset at time t can be obtained by

$$p_t = E[\xi_{t,T} p_T | \mathcal{F}_t], \tag{1}$$

where $\mathcal{F}_t$ denotes the investors' information set at time t.

In theory, the SPD ξ is defined across all assets in the aggregate economy, and the economy-wide SPD potentially depends on many state variables and is generally unknown. However, when using particular assets of interest we can obtain the SPD projected onto those assets. For example, the SPD for the equity market is the projection of ξ onto the index returns. While several studies estimate the equity market SPD using index options (e.g., Jackwerth, 2000) and some study the bond market SPD using interest rate derivatives (e.g., Li and Zhao, 2009), we investigate the SPD for the crude oil market.

The crude oil market is important for the overall economy but relatively segmented from other financial markets which makes that projection interesting. Figure 1 shows the scatter plots of daily oil futures returns, daily changes in oil futures volatility, daily stock market returns, and daily changes in stock market volatility. The stock and oil markets are clearly quite different from one another.

2.2 The State Price Density from Oil Futures and Options

Consider a European call option written on a futures contract $F_{t,T}$ with the strike price K, which matures at time T.[4] Under the risk-neutral probability measure $\mathbb{Q}$, option prices discounted at the (non-stochastic) riskless rate are martingales. If we know the risk-neutral distribution, we can therefore price the time-t call option by

$$\begin{aligned} C(F_{t,T}, K, t, T) &= e^{-r(T-t)} E^{\mathbb{Q}}[Max(F_T - K, 0)|\mathcal{F}_t] \\ &= e^{-r(T-t)} \int_K^{\infty} (F_T(x) - K) P^{\mathbb{Q}}(F_T(x)|\mathcal{F}_t) dx, \end{aligned} \tag{2}$$

where $P^{\mathbb{Q}}(F_T(x)|\mathcal{F}_t)$ is the conditional density of F_T under the risk-neutral measure. Based on this equation, we can price any option with a known final payoff once we have $P^{\mathbb{Q}}(F_T(x)|\mathcal{F}_t)$.

The call option price also is the expected value of the final payoff, discounted to time t, using the conditional SPD, $\xi_{t,T}$, as the (stochastic) discount factor,

$$\begin{aligned} C(F_{t,T}, K, t, T) &= E[\xi_{t,T} Max(F_T - K, 0)|\mathcal{F}_t] \\ &= \int_K^{\infty} \xi_{t,T}(x)(F_T(x) - K) P(F_T(x)|\mathcal{F}_t) dx, \end{aligned} \tag{3}$$

where $P(F_T(x)|\mathcal{F}_t)$ denotes the physical conditional density at time t.

Combining equations (3) and (2), we can estimate the projected SPD,

$$\xi_{t,T} = \xi(F_T|\mathcal{F}_t) = e^{-r(T-t)} \frac{P^{\mathbb{Q}}(F_T|\mathcal{F}_t)}{P(F_T|\mathcal{F}_t)}. \tag{4}$$

Defined as the Arrow-Debreu price of per unit of probability, the SPD reflects how investors evaluate possible states of economy and their expectations of the probability of those states happening. As shown in (4), SPDs depend on two components: risk-neutral densities and physical densities.

[4] Crude oil options expire three business days prior to the expiration of the underlying futures contract. To simplify the notation, we do not explicitly distinguish between the futures maturity date T and the option maturity date T' but they are of course both incorporated into our analysis.

Breeden and Litzenberger (1978) have shown that $P^{\mathbb{Q}}(F_T(x)|\mathcal{F}_t)$ can be obtained by taking the second-order derivative of call prices with respect to the strike price K,

$$P^{\mathbb{Q}}(F_T|\mathcal{F}_t) = e^{r(T-t)} \frac{\partial^2 C(F_{t,T}, K, t, T)}{\partial K^2}|_{K=F_T}. \tag{5}$$

Below we will provide details on the implementation of this second-order derivative on the grid of strike prices available for empirical work. Building on the large econometric literature on realized volatility, Maheu and McCurdy (2011) show how high-frequency intraday prices can be used to provide accurate physical conditional distributions at monthly and longer horizons. We will provide the details on this below as well.

The upshot is that option prices provide the risk-neutral distribution, high-frequency futures prices help estimate the physical distribution, and together they provide us with information on the oil price SPD.

2.3 Estimation of the Risk-Neutral Density

We compute conditional estimates of the risk-neutral density using option prices. More specifically, we adapt the semi-parametric approach introduced by Aït-Sahalia and Lo (1998) and further developed by Christoffersen and Jacobs (2004). The semi-parametric approach is designed to utilize all available information implicit in the entire cross-section of option prices, while keeping the parametric assumptions to a minimum. On a given day, we first fit Black (1976) implied volatilities from the cross-section of futures option prices to a second order polynomial in strike price and maturity. Then we construct a grid of strike prices and obtain at-the-money Black (1976) implied volatilities from the fitted polynomial function for each maturity $T - t$. With these implied volatilities, we back out fitted call prices $\widehat{C}\left(F_{t,T}, K, t, T, \sigma(K, T)\right)$ on the desired grid of strike prices. We are now ready to compute the second order derivative of the fitted option price with respect to the strike price

$$\widehat{P}^{\mathbb{Q}}(F_T|\mathcal{F}_t) = e^{r(T-t)} \cdot \frac{\partial^2 \widehat{C}\left(F_{t,T}, K, t, T, \sigma(K,T)\right)}{\partial K^2}|_{K=F_T}. \tag{6}$$

Let the return of longing a futures contract maturing at T be $R_{t,T} = \log(F_T/F_{t,T})$. We can then obtain the density of futures return over the period of $T-t$ from

$$\begin{aligned}\widehat{P}^{\mathbb{Q}}(R_{t,T}|\mathcal{F}_t) &= \frac{\partial}{\partial u}\Pr\left(\log(F_T/F_{t,T}) \le u|\mathcal{F}_t\right) = \frac{\partial}{\partial u}\Pr\left(F_T \le F_{t,T}\exp\left(u\right)|\mathcal{F}_t\right) \\ &= \widehat{P}^{\mathbb{Q}}(F_{t,T}\exp\left(u\right)|\mathcal{F}_t) \cdot F_{t,T}\exp\left(u\right), \end{aligned} \tag{7}$$

where Pr(.) denotes the cumulative distribution function. We compute a fixed one-month (21 business-day or 30 calendar-day) horizon option-implied density by interpolating the term structure of densities from (7) on each day.

As a robustness check, we construct the risk-neutral density of futures returns using the nonparametric approach from Aït-Sahalia and Duarte (2003) and Li and Zhao (2009), which we have not reported here as the results are similar.

2.4 Estimation of the Physical Density

We estimate the physical density from historical futures prices. Estimation of the historical distribution needs to take two considerations into account. First, we ought to use a time series of data as long as possible in order to increase precision of estimates. The longer the sample period is, the more efficient estimator we can obtain. Second, the estimation methodology needs to account for time-variation of the physical density, especially to allow for the presence of stochastic volatility in the crude oil market as documented by Trolle and Schwartz (2009). Bansal, Kiku, Shaliastovich, and Yaron (2014) further highlight the importance of time-varying volatility when estimating SPDs from financial market data.

We first calculate the time series of daily futures returns $\{R_{t,T}\}_{t=1}^{N}$ from 1990 to 2012 using futures prices. It is equivalent to the continuously compounded returns of holding the futures contract to maturity and realizing returns by closing out the position at the maturity date T. At each time t, we normalize the time series of returns with its sample mean $\overline{R}$ and conditional volatility $\sigma_{t,T}$, the estimation of which is described as below. This gives a time series of return innovations

$$z_{t,T} = \left(R_{t,T} - \overline{R}\right) / \sigma_{t,T}.$$

Then, similar to Jackwerth (2000), we estimate the density with a kernel function using the return innovation at t. The physical density of returns is then obtained as

$$\widehat{P}(R_{t,T}|\mathcal{F}_t) = \widehat{P}(\overline{R} + \sigma_{t,T} \cdot z_{t,T}).$$

We utilize high-frequency intraday oil futures prices to estimate conditional volatility, $\sigma_{t,T}$. Andersen and Bollerslev (1998), Andersen, Bollerslev, Diebold, and Labys (2003) and many others since have reported on the superiority of volatility estimated from intraday high-frequency data compared with daily data. We calculate daily volatility using the two-scale estimation approach from Zhang, Mykland and Ait-Sahalia (2005), which Andersen, Bollerslev and Meddahi (2011) have shown to be robust to the impact of microstructure noise in high-frequency data.

In order to match the forward looking horizon of option-implied risk neutral density, we compute expected one-month volatility using the heterogeneous autoregressive (HAR) model proposed by Corsi (2009). We first estimate daily regressions of the form

$$RVol_{t:t+21} = a + b_d RVol_{t-1:t} + b_w RVol_{t-5:t} + b_m RVol_{t-21:t} + e_{t:t+21},$$

where $RVol_{t-1:t}$, $RVol_{t-5:t}$, and $RVol_{t-21:t}$ denote the most recent daily, weekly, and monthly volatility, respectively, and we make sure we only use information up to day t when we run regressions. Then we use the HAR regression to predict volatility for the next month $\sigma_{t,T} =$

$E_t\left[RVol_{t:t+T}\right]$, with T equal to 21. The HAR regressions have been used in many studies including Busch, Christensen, and Nielsen (2011) to forecast volatilities in various financial markets. We estimate the HAR regression coefficients using a rolling window of 252 days.

3 Empirical Oil SPDs

In this section we first describe the oil futures and option data and estimate oil SPDs. We then analyze the oil SPDs using estimation-free option returns and model-free, option-implied skewness.

3.1 Futures and Option Data

We obtain daily crude oil futures and option data from January 2, 1990 to December 31, 2012 from the Chicago Mercantile Exchange (CME Group, formerly NYMEX); and we obtain high-frequency intraday oil futures prices from January 5, 1987 to December 31, 2012 from TickData. Crude oil traded on the CME is the largest and most liquid commodity future. The range of maturities of futures, and the range of strike prices of options are also larger than for other commodity derivatives. An advantage of the data is that crude oil futures and options have been traded on this exchange for more than 20 years, which allows us to study a long time series spanning recessions and many geopolitical events such as the Gulf wars, the 9/11 terrorist attacks, the 2008 crisis, and especially the recent boom and bust in commodity prices. This dataset also provides open interest and trading volume of both options and the underlying futures contracts. Relative demand and trading volume in futures and options reveal investor beliefs and expectations (e.g., Buraschi and Jiltsov, 2006), and therefore are informative when studying the SPDs.

The calculation of the risk-neutral density in (6) is based on European options, but the crude oil option data are American type. We convert American option prices into European option prices following Trolle and Schwartz (2009) who use the methodology of Barone-

Adesi and Whaley (1987). After obtaining European option prices, we exclude options with Black (1976) implied volatility less than 1% or greater than 200%, options with prices less than \$0.05, and contracts violating standard no-arbitrage conditions. Our empirical analysis is conducted at the weekly frequency and uses OTM calls and puts. We rely on OTM options because doing so minimizes the potential effects of approximation errors in the early exercise premium, and second, OTM options are usually more liquid than in-the-money (ITM) options. Each week, we sample options on Wednesdays since it is the weekday least likely to be a holiday and to be affected by day-of-the-week effects. This selection of data has been widely used in the literature (e.g., Bates, 1996, 2000; Heston and Nandi, 2000). As the computation of risk-neutral densities is based on call prices, we utilize traded OTM puts and transform them into synthetic ITM calls. Our call option data thus effectively span the entire moneyness range when implementing equation (6).

Panel A of Table 1 provides descriptive statistics on the futures data by maturity. Although the number of contracts and average prices are relatively constant across maturities, average open interest and trading volume decrease sharply beyond the two-month maturity. While open interest of six-month futures contracts is around 20% of one-month contracts, the trading volume of six-month contracts is only about 4% of one-month contracts. Long-maturity futures often lack liquidity, which is also true for options as reported in Panel B. The trading volume of all option contracts beyond six months (or 180 calendar days) is only about 5% of the one-month contract volume. Panel C of Table 1 reports option data across moneyness. Although we have many observations on deep OTM (ITM) options, at-the-money options are the most heavily traded. Across moneyness, the average Black (1976) implied volatility displays a smile pattern with deep OTM (ITM) options trading at higher implied volatility than ATM options. Across maturities, short maturity options on average have a higher implied volatility than long maturity options as shown in Panel B.

3.2 The Shape of Oil SPDs

We compute conditional risk-neutral densities and physical densities on each Wednesday using the approach outlined in sections (2.3) and (2.4). We find that while the risk-neutral density can be either negatively or positively skewed, the physical density is relatively more symmetric. While the sample mean of the risk-neutral variance and physical variance are comparable and the sample mean of the ratio of risk-neutral and physical variance $\frac{|Var^{\mathbb{Q}}|}{|Var|}$ is 1.34, the sample mean of the ratio of absolute risk-neutral and absolute physical skewness $\frac{|Skew^{\mathbb{Q}}|}{|Skew|}$ is over 3.23. Moreover, risk-neutral skewness exhibits much higher time variation than physical skewness: the standard deviation of risk-neutral skewness and physical skewness is 0.191 and 0.039 respectively. The magnitude and time variation of risk-neutral skewness dominate physical skewness. Therefore, the shape of the SPD, which we calculate as the log difference between the risk-neutral and physical densities, is mainly driven by the risk-neutral density (Han, 2008). Figure 2 shows the estimated one-month SPDs as a function of futures returns in the crude oil market on each Wednesday during 1998 to 2012. The horizontal axis denotes returns, and the sample year is indicated in the title of each graph.

Figure 2 shows that the SPDs are nonlinear and generally display an asymmetric U-shape as a function of returns. At the aggregate level, investors in the crude oil market regard states with extreme declines or increases in the oil price as bad states and assign a high value for payoffs received in those states. This might be due to the heterogeneous nature of investors in the crude oil derivative market: Investors with long positions, will bear losses in the case of futures price decreasing if their positions are not protected. They regard negative returns as bad states and highly value payoffs received in these states. Investors with short positions, will suffer from increasing futures price and consider those states with positive returns as bad states. They assign a high value to payoffs received when oil futures returns are extremely high. The oil futures market is different from the equity market in which the representative agent is always long the underlying asset.

The U-shaped SPDs display considerable variations across time. First, we observe that investors assign different state prices across time to the same level of returns. For example, for a given level of returns, the state prices for very large negative or positive returns are higher in 2003 than in 2005. The asymmetric U-shape is also wider in the $2008-2009$ years than in other years. Second, the U-shape has a different level of dispersion across time. For a given range of returns, the range of state prices significantly differs. For example, consider the year of 2000 and 2007: while investors assign similar state price to certain returns and have constant preferences towards those returns during 2000, their value for the same level of returns varies much more during 2007. Third, the dispersion in both the decreasing and the increasing regions start to increase around 2004 or 2005, when speculative activity–and investor belief heterogeneity–in the crude oil market dramatically increase after the financialization of the commodity market. The slopes in both the decreasing and increasing regions can become steeper or flatter depending on the state of nature. Investors' marginal value of payoffs in different states, when returns are negative or positive, exhibits significant variations across time. We conclude that it is crucial to allow for conditional SPDs in the oil market.

3.3 Option Returns and Oil SPDs

Even though we have estimated the oil SPDs imposing a minimum level of modeling assumptions, it is worth assessing the robustness of the U-shape found in the SPDs so far. Therefore, we next use a completely model-free approach, namely the returns of OTM puts and calls, to support the SPDs are non-monotonic and U-shaped. Investors in the crude oil futures market can hedge potential losses due to low (high) oil prices using OTM puts (calls). In particular, if net long futures investors are worried about extremely negative oil returns, they will demand OTM puts. OTM puts will be expensive and their returns will be negative and increase in the strike price. If net short investors consider extremely positive returns as bad states, they will demand OTM calls. Returns of OTM calls will be negative and will decrease in the strike price. Therefore, by analyzing returns of OTM puts and OTM calls that pay off when futures have large negative or positive returns, we can detect whether investors generally regard both extremely low and high oil futures prices as bad states of economy. Such a finding will in turn provide supporting evidence for U-shaped SPDs.

We compute option returns as follows. On the third Wednesday of every month, we take a long position in available calls and puts with maturity as close to 30 calendar days as possible. We compute the hold-to-maturity returns of calls and puts as

$$\begin{aligned} r_{t,T}^{call} &= max(F_{t,T}e^{R_{t,T}} - K, 0)/c_{t,T} - 1, \\ r_{t,T}^{put} &= max(K - F_{t,T}e^{R_{t,T}}, 0)/p_{t,T} - 1, \end{aligned} \tag{8}$$

where $F_{t,T}$ is the price of underlying futures, $R_{t,T} = \log(F_T/F_{t,T})$ is the return of the underlying futures contract during $T - t$, K is the strike price, and $c_{t,T}$ and $p_{t,T}$ are European style call and put option prices. We use observed options and we do not create artificial prices by interpolation or extrapolation. On each Wednesday, we assign available options to various bins by moneyness, defined by $X = K/F_{t,T}$, and option returns are then averaged within each bin. This procedure provides a non-overlapping return time series for various moneyness levels. In Table 2 we show the average return, its t-statistics, the 10^{th}, 50^{th} and 90^{th} percentiles in each moneyness bin using data from January 1990 to December 2012, as well as average returns for two sub-sample periods: pre-financialization January 1990 to December 2004, and post-financialization January 2005 to December 2012.[5] We consider five moneyness bins: $(0.5, 0.85]$, $(0.85, 0.95]$, $(0.95, 1.05]$, $(1.05, 1.15]$ and $(1.15, 1.5)$. OTM puts have moneyness of $(0.5, 0.85]$ and $(0.85, 0.95]$; OTM calls have moneyness of $(1.05, 1.15]$ and $(1.15, 1.5)$. We are mostly interested in returns of deep OTM puts ($X \in (0.5, 0.85]$) and deep OTM calls ($X \in (1.15, 1.5)$).

Panel A of Table 2 reports put option returns. For the period of 1990 to 2012, put options mostly earn negative returns and returns increase in moneyness (strike). The negative

[5] Our results also obtain if we use American option prices and/or if we date the financialization one year ealier.

returns on OTM puts are larger in magnitude than those for ITM puts and are statistically significant. When we compare OTM and ITM puts, returns from OTM puts are significantly lower (more negative). Call option returns are somewhat different from put returns as shown in Panel B of Table 2. There is no clear monotonic pattern in returns across moneyness. Although ITM and ATM call options can have positive returns, the returns on deep OTM calls ($X \in (1.15, 1.5)$) are negative and statistically significant. In addition, returns on OTM calls are always lower than returns on ITM calls. Combining the negative returns on OTM puts and OTM calls in Panels A and B, we conclude that both deep OTM puts and deep OTM calls are more expensive than other options. Bakshi, Madan, and Panayotov (2010) show that negative call option returns are compatible with a U-shaped state price density. Chaudhuri and Schroder (2015) further show that the SPD is monotonic if and only if option returns increase in strike prices, i.e. OTM calls should have higher returns than ITM calls.

When we compare the pre- and post-2005 periods, we find that the negative return pattern of put options becomes weaker in the post-2005 period than in the pre-2005 period, as only deep OTM puts have significantly negative returns after 2005. Interestingly, call option returns *decrease* across moneyness after 2005, which is not compatible with the monotonically downward-sloping SPDs but compatible with a U-shape. OTM call options become more expensive in the post-2005 period as returns of OTM calls are more negative, which implies that investors become worried about extremely high future oil prices. This could be because short position speculators (who do not have natural hedge and use options to hedge their futures positions) have exposure to the futures price risk, and demand OTM call options to hedge.[6] Alternatively it could be speculators getting leverage through OTM options.

We next analyze if the expensiveness of OTM calls (and puts) is due to illiquidity premia. We do not find supporting evidence for this. In our sample period, both trading volume and

[6] In general speculators have net long positions in the aggregate and provide liquidity for commodity producers (hedgers). However, short-position speculators do exist as evidenced by the Morningstar Short/Flat Commodity Index and the Morningstar Short-Only Commodity Index. In addition, some long positions by financial traders require short positions from other speculators, who may themselves need to hedge using options.

the ratio of trading volume over open interest of OTM calls (puts) are higher than ITM calls (puts), and OTM calls and puts are more liquid than ITM calls and puts. Therefore, OTM calls and puts are actively traded, and the negative returns cannot obviously be generated by illiquidity.[7]

Finally, we analyze returns on butterfly spreads. On the third Wednesday of every month, we take a long position in a butterfly spread constructed from call options with maturity close to 30 calendar days. We buy an ITM call option and an OTM call option with strike prices K_1 and K_3, and we sell two ATM call options with a strike price of K_2 such that $K_2 - K_1 = K_3 - K_2$. We find that the average return of the butterfly spread is significantly negative when K_3 is above a certain threshold. This is consistent with the above argument that OTM call options are more expensive relative to other calls, which in turn supports the increasing SPDs in the region of large positive returns. The fact that returns on call options *decrease* with strike prices, along with negative returns of the butterfly spread, supports the evidence for a non-monotonic SPDs.[8]

In summary, we find that put option returns are negative and increase across the strike price. OTM call option returns are negative and decrease across the strike price when the strike price is sufficiently high, and especially after 2005. Negative returns of OTM puts and calls imply that investors are worried about both extreme negative and positive oil futures returns and consider both as bad states, which lends support to the U-shaped SPDs. The difference in patterns between put and call returns supports the asymmetry of the U-shaped SPDs. Our evidence of negative returns on OTM call options is consistent with the model of Bakshi, Madan and Panayotov (2010) where investors have heterogeneous beliefs and can take short positions.

[7] Our dataset does not include bid-ask spreads and so our conjectures about liquidity are based on our volume data.

[8] Chaudhuri and Schroder (2015) show that the SPDs are monotonically downward-sloping *if and only if* returns on calls are *increasing* in the strike price.

3.4 Risk-Neutral Moments

We are not interested only in the empirical shape of SPD in the crude oil market, but also how SPDs are affected by investors' heterogeneous beliefs. This is related to the economic question of how much more investors in the aggregate are willing to pay for securities that pay off in one state of the economy versus another. As the shape of the SPDs is mainly driven by the properties of the risk-neutral distribution, we compute risk-neutral moments and link their time variation to investor belief heterogeneity and trading activities.[9] Risk-neutral moments are calculated from the estimated risk-neutral densities:

$$Var^{\mathbb{Q}}_{t,T} = E^{\mathbb{Q}}_t\left[\left(R_{t,T} - E^{\mathbb{Q}}_t[R_{t,T}]\right)^2\right], \tag{9}$$

$$Skew^{\mathbb{Q}}_{t,T} = \frac{E^{\mathbb{Q}}_t\left[\left(R_{t,T} - E^{\mathbb{Q}}_t[R_{t,T}]\right)^3\right]}{\left\{E^{\mathbb{Q}}_t\left[\left(R_{t,T} - E^{\mathbb{Q}}_t[R_{t,T}]\right)^2\right]\right\}^{3/2}}, \tag{10}$$

where $E^{\mathbb{Q}}_t[x] = \int_{-\infty}^{\infty} xP^{\mathbb{Q}}(x)dx$ is the expected value under the risk-neutral measure and $P^{\mathbb{Q}}(x)$ is the density estimated from option prices. Since risk-neutral moments are calculated using the option implied densities and available returns on each Wednesday, they are model free and conditional on current market conditions.

Figure 3 displays the time series for the weekly option-implied variance (upper panel) and skewness (lower panel) from 1990 to 2012. From left to right, the first two vertical dotted lines denote two significant weeks of the first Gulf War: the Iraq invasion of Kuwait on August 2, 1990, and the liberation of Kuwait on January 17, 1991. Other vertical dotted lines denote the week of the September 11, 2001 terrorist attacks; the second Gulf War on March 20, 2003; the week when the WTI crude oil spot price reached its highest level in history (July 23, 2008 [$145.31]); the week when the Lehman Brothers filed for Chapter 11 bankruptcy protection (September 15, 2008); the week when the crude oil spot price reached its lowest

[9] Datta, Londono and Ross (2014) investigate how option-implied moments change around important events in the crude oil market.

level during the recent crisis period (December 23, 2008 [$30.28]); the week when the Libyan Civil War began and oil and gas production in Libya fell by more than 60%; and the week when Standard & Poor's downgraded the U.S. long-term sovereign credit rating from AAA to AA+, respectively. We note that the variance rises on those days. Consistent with Robe and Wallen (2014), these hikes in the upper panel of Figure 3 indicate that oil variance is affected by not only the oil market fundamentals, but also macroeconomic conditions. However, when we compare the upper and lower panels of Figure 3, it appears that skewness reacts more to oil market-specific information. For example, skewness significantly increases during the week of the second Gulf War in March 2003 and the week of Libyan War in February 2011, but it does not change much when the U.S. sovereign credit is downgraded.

The empirical results so far have shown that the SPDs in the crude oil market display an asymmetric U-shape as a function of returns, and exhibit considerable variation across time. Returns of OTM options are consistent with this nonlinear pattern of the SPDs. We also find some preliminary evidence that the risk-neural skewness of oil futures returns is more likely to be affected by oil market-specific information. We next investigate how the shape of the SPDs depend on investors' heterogeneous beliefs.

4 Heterogeneous Beliefs and Oil SPDs

Bakshi, Madan, and Panayotov (2010) and others have shown that the SPDs depend on investor disagreement. Heterogeneity is represented not only in asset prices and returns, but also in the relative positions and trading volumes in equilibrium. Dependence of the SPDs on heterogeneous investor beliefs is present in the equity market (e.g., Anderson, Ghysels and Juergens, 2005; Buraschi and Jiltsov, 2006), as well as in the foreign currency market (Beber, Buraschi and Breedon, 2010). When agents have different beliefs about asset prices, they engage in trading either for speculation or hedging.

In this section, we investigate how heterogeneous beliefs embedded in crude oil derivatives trading affect the slope of the SPD. The SPD slope is a natural object of interest because it can be directly related to the investors' risk aversion. The steeper the slope of the SPD, the higher state prices investors assign to economic states corresponding to more extreme returns. We first describe how investors' trading positions have evolved over the past twenty years in crude oil futures and options and we then construct measures of heterogeneous beliefs in these markets. Subsequently, we document that the slope of the SPD in both the negative and the positive return regions are affected by heterogeneous beliefs. We complement the SPD analysis by investigating how heterogeneous beliefs affect risk-neutral oil skewness.

4.1 Market Participation and Measure of Beliefs

The CFTC classifies large traders in the crude oil derivatives market into commercial traders or hedgers and non-commercial traders or speculators.[10] Figure 4 shows long and short positions taken by hedgers and speculators in the futures market, as well as their ratios, obtained from the CFTC futures-only Commitments of Traders (COT) report. Although participation in the futures market by both hedgers and speculators has experienced steady growth from 1990 onwards, the increases in positions have been faster since 2004-2005, as shown in the top panel of Figure 4. When we look at the ratio of long positions taken by speculators over hedgers, we see the ratio typically fluctuates around 0.2 before 2004-2005, but displays constant growth afterwards. The speculator-to-hedger ratio of short positions also displays steady growth, although not as substantial as the ratio for long positions. The rapid growth in positions from speculators and the historical boom and bust of futures price in 2008 (as shown in the bottom panel of Figure 4) have drawn significant attention from the academicians, practitioners and policy makers. Some papers (e.g., Singleton, 2014, and Cheng and Xiong, 2014) have attributed position growth of speculators to commodity linked investments from financial institutions.

[10] We acknowledge that this classification of hedgers and speculators is not perfect and does not cover detailed information on trading purposes by all investors in the crude oil derivatives market. However, Büyükşahin and Robe (2014) confirm that the speculative activities inferred from the public CFTC position data are able to capture speculative activities inferred from non-public, trader-level CFTC data quite well.

The crude oil option market has experienced even more dramatic growth. Figure 5 shows the number of OTM options and their open interest and trading volume aggregated within each week. OTM calls are defined as call options with moneyness (K/F)>1.05, and OTM puts are put options with moneyness (K/F)<0.95. We only consider options with maturity below 180 calendar days. We also report the weekly average annualized 30 calendar-day fixed-maturity oil option-implied volatility, which is constructed using CBOE's VIX methodology. While there is an increasing number of OTM options traded in the market, the growth of the number of options after $2004-2005$, especially in the first half of 2008 is indeed dramatic. In particular, the number of OTM call option contracts jumps in 2008 as shown in the top panel of Figure 5. When oil prices are most volatile, investors appear to demand OTM options to hedge their positions. We observe that while open interest in options has decreased a bit after 2010 (as shown in the second panel), OTM options has been traded more actively over time (as shown in the third panel).

Given the active participation in futures and option markets, we consider measures of investor beliefs in both markets and investigate how they affect the SPDs. We will focus on variables measuring differences in beliefs, as well as variables that directly relate to either net long or net short positions. Although these measures are different, they are all broadly related to heterogeneity in investor beliefs, as suggested by the literature we discuss below.

First we consider the speculation index, which quantifies the level of speculation activity, as a measure of investor beliefs in the crude oil market. Several studies such as Scheinkman and Xiong (2003) and Xiong and Yan (2010) argue that agents with heterogeneous beliefs engage in speculative trading with each other. Singleton (2014) discusses how disagreements among investors induce speculative activities, price drift, and high volatility in the crude oil market.

The speculation index is constructed to gauge the intensity of speculation relative to hedging (Working, 1960; Büyükşahin and Harris, 2011). If we denote SS (SL) as the speculative short (long) position, and HS (HL) as the hedging short (long) position, we can define

$$\text{Speculation Index}_t = \begin{cases} 1 + \frac{SS_t}{HL_t + HS_t} & \text{if } HS_t \geq HL_t, \\ 1 + \frac{SL_t}{HL_t + HS_t} & \text{if } HS_t < HL_t. \end{cases} \tag{11}$$

The speculation index has to be interpreted together with hedging activities in the futures market. It measures the extent to which speculative positions exceed the level necessary to meet hedging demand. Panel A of Figure 6 plots the speculation index from 1990 to 2012. We observe that there has been a high level of speculative activities in recent years. Before 2000s, the speculation index is below 1.05 meaning less than 5% speculation net of hedging. However, it rises steadily over time to 1.2 in 2010 and drops to around 1.1 afterwards. As such, the premia that hedgers pay for insurance against future price risk are likely affected by the active participation of speculators. For example, Hamilton and Wu (2014) document significant changes in oil futures risk premia since 2005 due to index-fund investing.

Besides positions taken by speculators, actual trading volume of futures may reflect the degree of heterogeneity and how investors speculate and share risks among each other. The literature (e.g., Kandel and Pearson, 1995) has documented the positive relationship between investor belief heterogeneity and volume of trade. Buraschi and Jiltsov (2006) show that, the trading volume of stocks and options is positively correlated with the dispersion in beliefs. Carlin, Longstaff, and Matoba (2014) show that increased disagreement is associated with higher trading volume. Therefore, we use the trading volume of futures as another measure of investor belief dispersion in the underlying futures market. Panel B plots the trading volume of 30-day futures contracts, which grows steadily during the entire period but has become very volatile during the recent years.

Next we discuss two measures of heterogeneous beliefs in the option market used by Han (2008). One is the open interest ratio of OTM puts to calls, which measures the relative demand for insurance against downside risks and reflects the hedging needs of different agents. A higher open interest ratio of OTM puts to calls suggests investors are overall more pessimistic, and they tend to demand put options either to get protection against future price drops or to pursue potential returns on put positions. Panel C shows that the open interest ratio of OTM puts to calls having several spikes during our sample period. The other measure is the trading volume of options, which is a proxy for the level of disagreement among options investors. Open interest and trading volume capture different information as open interest only increases when both sides of a trade open a new position, whereas any trade (both open, both close, one closes and one opens) increases volume. To separate the factor that specifically affects net long and net short futures investors, we consider trading volume of OTM puts and OTM calls individually. Panels D and E shows that, the trading volume of OTM puts and OTM calls is relatively stable before 2004-2005 but begins to rise steadily afterwards. We adjust the time series of futures and option trading volume with a Hodrick-Prescott filter to remove the time trend.[11] Since some trading in the near maturity futures and option contracts can be due to liquidation or rollover to the next maturity contract, trading volume and open interest measured from the near contracts may have little information content. We therefore interpolate the closest-to-maturity contract and the next available contract, and obtain futures and options trading volume and open interest with a maturity of fixed 30 calendar days, which matches our time-horizon of the SPDs and risk-neutral moments.

To what extent do investor beliefs of the economy and the equity market affect the SPDs in the crude oil market? We consider two types of proxies of beliefs to address this question: investor beliefs about the equity market and economy, as well as investors' expectations about market volatility. The first measure is the bull-bear spread based on the survey of Investors Intelligence which has been used by Brown and Cliff (2004, 2005) and Han (2008). Every week, Investors Intelligence sends out 150 surveys to institutional investment advisors and collects their expectations of future market movements as bullish, bearish, or neutral. The bull-bear spread is then calculated as the percentage of bullish investors minus the percentage of bearish investors, and it is often used as a proxy for beliefs of institutional

[11] Han (2008) and Buraschi and Jiltsov (2006) remove the time trend in trading volume as well.

investors about the equity market. Secondly, we consider the consumer sentiment index from the University of Michigan as plotted in Panels G.[12] The literature has used this index to study how sentiment affects stock prices (e.g., Lemmon and Portniaquina, 2006; Stambaugh, Yu, and Yuan, 2012). The third proxy we use to measure investor beliefs is the CBOE's VIX index. VIX has become a benchmark for measuring investors' expectations of market volatility and investor sentiment as a fear indicator. As plotted in Panels H and F, VIX is in general negatively correlated with the bull-bear spread. The VIX index tends to be higher when more investors are less confident in the market.

4.2 Other Control Variables

We consider two sets of control variables in the regression analysis. One set includes oil market-specific variables that determine futures returns and therefore may affect the SPDs. We first follow Hong and Yogo (2012) and calculate the basis as

$$\text{Basis}_{t,T} = \left(\frac{F_{t,T}}{S_t}\right)^{\frac{1}{T-t}} - 1, \tag{12}$$

where $F_{t,T}$ is the price of the futures contract with maturity $T - t$ at time t, and S_t is the spot price at time t. The basis can be interpreted as the implied net convenience yield.

Next we include option-implied oil volatility. First, Trolle and Schwartz (2009) present strong evidence of stochastic volatility in the crude oil market; and stochastic volatility has to be incorporated into the model to price commodity derivatives. Therefore, volatility could affect futures prices and return distributions. Second, contrary to the traditional theory of storage, which claims volatility tends to be high when the futures price is in backwardation, Carlson, Khokher, and Titman (2007) and Kogan, Livdan, and Yaron (2008) document that the relationship between the futures term structure slope and volatility is non-monotonic. While the sign of the slope of the futures term structure implies investors' demand for

[12] When we include the investor sentiment index in Baker and Wurgler (2006), which restricts our analysis up to 2010, our main results remain qualitatively unchanged.

convenience, volatility captures complementary information about investors' exposure to risk.

We also consider the storage level of crude oil and historical returns of futures, both of which affect futures returns as documented in the literature (e.g., Bessembinder, 1992; Bessembinder and Chan, 1992; de Roon, Nijman, and Veld, 2000). We use the storage level defined as U.S. total stocks of crude oil, excluding strategic petroleum reserves, based on the report from Energy Information Administration (EIA). We apply a Hodrick-Prescott filter to remove the trend, where the smoothing parameter is set to a number appropriate for weekly data. The historical return is the moving average of the daily return during the past week on a long position in the futures contract.

The other set of control variables captures macroeconomic conditions and investment opportunities. They include the 1-month Treasury bill rate, the spread between yields on the 10-year Treasury bond and the 3-month Treasury bill, the difference in yields on Baa and Aaa corporate bonds, the Chicago Fed National Activity Index and the log growth in aggregate industrial production during the past 12 months.[13] These variables are used by Han (2008) and others as well. We obtain the one-month Treasury bill rate from the online data library of Kenneth French, and the other macro variables are from the Federal Reserve Bank of Chicago and St. Louis.

Table 3 reports summary statistics of risk-neutral skewness computed from (10), as well as the main regressors and control variables in the crude oil market. We apply necessary filters to some regressors to remove time trends. The final column of Table 3 shows that the unit-root tests are rejected at the 5% level for all the time series we use in the regressions. Note also that overall skewness is negative, investors on average demand more OTM puts than OTM calls, and historical annualized return of holding the one-month futures contract is 7.46% with a Sharpe ratio of 0.47. Table 4 reports the weekly correlation matrix of regressors, as well as control variables in the crude oil market. We observe that the correlations are

[13] Our results are robust to using the Philadelphia Fed ADS index instead of the CFNAI index from the Chicago Fed.

rather low, except for oil volatility (IVOil) which has a correlation of 65.33% with VIX, and OTM put/call ratio which has a correlation of -41.04% with the basis. Given the high correlation between IVOil and VIX, IVOil is orthogonalized with VIX before it is included in the regression analysis. That is, we use the residual in the regression of $IVOil_t = a + bVIX_t + e_t$. Note also that the trading volumes of futures, call and put options not surprisingly are positively correlated among each other.

4.3 Regression Results

We consider three types of measures to examine how investor beliefs affect the slope of the SPD and thus the marginal rate of substitution across states in the crude oil market: 1) belief dispersion in the crude oil futures market, 2) belief dispersion in the crude oil option market, and 3) investor beliefs of the equity market and the economy. In our regressions we add control variables discussed above that may affect crude oil futures returns and reflect macroeconomic conditions. The full regression model is given by

$$\begin{aligned} \text{Slope of the SPD}_t \;=\; & \alpha_t + \beta_t \cdot \text{Beliefs in the crude oil futures market}_t \\ & + \gamma_t \cdot \text{Beliefs in the crude oil option market}_t \\ & + \delta_t \cdot \text{Beliefs in the equity market}_t \\ & + \eta_t \cdot \text{Control variables}_t. \end{aligned} \tag{13}$$

To assess the dependence of the slope of the SPD on investor beliefs, we run the regression in (13) using different dependent variables. We first measure the slope of the SPD curve directly using the differences in SPD values across return intervals. Second, we substitute risk-neutral skewness for the direct slope measure, since the asymmetric U-shaped pattern of the SPDs is mainly determined by the level of risk-neutral skewness. Before performing regressions, we standardize and winsorize all variables at the 1^{st} and 99^{th} percentile. To minimize estimation noise from the limited option data, we only include days with at least two OTM calls and two OTM puts in the regression sample to ensure the estimate of risk-neutral skewness is reliable.

4.3.1 SPD Slopes and Investor Beliefs

The SPD in the crude oil is U-shaped so that investors assign high state prices to large negative and positive returns. The decreasing (increasing) region of the SPD captures investors' preferences towards different level of negative (positive) returns. On each Wednesday, we compute the slope as the difference between the value of the SPDs at two return points:

$$\text{Decreasing-Region Slope}_t = \text{SPD}_t(R_2) - \text{SPD}_t(R_1), \text{ where } R_2 < R_1 << 0;$$

$$\text{Increasing-Region Slope}_t = \text{SPD}_t(R_2) - \text{SPD}_t(R_1), \text{ where } R_2 > R_1 >> 0.$$

This direct measure of SPD slope is similar to that in Xing, Zhang and Zhao (2010) who quantify the steepness of stock option smirks as the difference of implied volatilities of two moneyness points. To provide a broad and consistent picture of the slope, we consider two measures of the slope in both the decreasing- and increasing regions. Taking the asymmetry of the oil SPD into account, the decreasing-region slopes are based on return intervals $(-0.27, -0.21)$ and $(-0.25, -0.19)$, and the increasing-region slopes are measured from return intervals of $(0.11, 0.17)$ and $(0.12, 0.18)$.

Table 5 presents the regression results. It shows that the open interest ratio of OTM puts over OTM calls positively (negatively) affect the steepness of the decreasing- (increasing-) region slope. In other words, when investors demand more OTM puts, as suggested by higher OTM put open interest, they overall assign higher state prices to large negative returns and lower state prices to large positive returns. We find that, if investors trade more OTM call options, the increasing region slope becomes steeper and investors assign higher state prices to large positive returns. There is evidence that the level of speculation is positively (negatively) associated with the steepness of the decreasing- (increasing-) slope, indicating that investors' marginal utility over negative returns increases (they assign higher state price to negative returns) when there is more speculative activity. We also see that our proxies for investor beliefs about the equity market and the economy have no effects on the slope of oil SPDs. These proxies therefore do not seem to matter for how investors compare relative state price across returns in the oil market.

Although we present evidence that the slope of the SPD is affected by investor beliefs in the crude oil market, we have to discuss the decreasing and increasing SPD regions separately. In addition, there are multiple return intervals on each trading day and the choice of return interval may change the magnitude of the slope. Therefore, this slope measure only represents investors' preferences over certain return regions but not all possible economic states. Next, we argue that risk-neutral skewness is a more comprehensive measure of the slope and investors' preference, and our later analyses focus on the risk neutral-skewness.

4.3.2 Option-Implied Skewness and Investor Beliefs

Since the SPD is the ratio of the risk-neutral and physical density, the asymmetric U-shaped pattern of the SPD mainly arises from the differences between risk-neutral and physical skewness. Because the skewness of the risk-neutral density has more pronounced time variation compared with the physical density, as we have discussed in section (3.2), the slope of the U-shaped SPDs mostly depends on the skewness of the risk-neutral density (Han, 2008). Furthermore, the risk-neutral skewness measure utilizes all data available with valid closing quotes. It provides a thorough picture of the slope of the SPDs and investors' overall attitudes across economic states. Therefore, risk-neutral skewness is arguably a more reliable and comprehensive measure of the slope (or the marginal rate of substitution across states). We next rely on risk-neutral skewness and investigate whether it is affected by investor beliefs. Risk-neutral skewness is calculated from risk-neutral densities and the formula (10). To assess the impact of speculative activities on the SPDs during the post-financialization period, we use a dummy variable that equals to one for the period of 2005 to 2012 and equals to zero for the period of 1990 to 2004. The coefficient of Speculation Index$\times$Dummy will tell us whether speculation has different impacts on risk-neutral skewness or the SPDs during the two sub-period periods.

Table 6 contains the basic results from regressing risk-neutral skewness on belief measures from the oil futures and options market. In Panel A, the dependent variable is risk-neutral skewness. We include lagged skewness to control for its positive autocorrelation in all specifications. Regression results show that the increasing level of speculation after 2005 is associated with a more negative risk-neutral skewness, i.e., the decreasing-region slope is steeper and investors' marginal utility over negative returns is higher. This relationship between speculation and the steepness of the decreasing-region slope is consistent with Table 5, but intensifies after financialization. We observe that open interest ratio between OTM puts and calls is negatively related with skewness. A high open interest ratio of OTM puts to calls suggests investors are overall more concerned about large negative returns, and is therefore associated with high state prices of large negative returns. We also observe that OTM put volume (OTM call volume) is negatively (positively) associated with risk-neutral skewness, which means that higher OTM put volume (OTM call volumes) suggests a steeper decreasing- (increasing-) region slope of the SPDs, and investors are worried about large negative (positive) returns.

Panel B of Table 6 reports the regression results when we use the AR(1) residual of risk-neutral skewness as the dependent variable. In other words, we regress the residual of $Skew_t = a + b * Skew_{t-1} + \varepsilon_t$ on various measures of investor beliefs, which are also the residual from the AR(1) regressions. The results are consistent with Panel A. The higher the open interest ratio of OTM puts to calls and the higher the OTM put trading volume are, the more negative the risk-neutral skewness is and investors overall assign higher state prices to large negative returns. The higher the trading volume of OTM call options is, the more positive the risk-neutral skewness is and investors overall assign higher state prices to large positive returns.

Two comparisons with the existing literature can be made here. First, the empirical evidence that the OTM puts to calls ratio affects the risk-neutral skewness aligns with findings in the stock and index option literature (Dennis and Mayhew, 2002; Han, 2008). Since heterogeneous investors have different demands for OTM options due to their various expectations of market fundamentals, pessimists can share risks with others by buying insurance from optimists. The larger the difference in beliefs, the higher the demand for OTM puts which drives up the prices of options with low strike prices, and the distribution is therefore more negatively skewed. Second, consistent with Buraschi and Jiltsov (2006) we confirm that option trading volume can affect the slope of the SPDs as a proxy for dispersion in beliefs. OTM put (call) trading volume negatively (positively) affects risk-neutral skewness, meaning a higher level of belief dispersion of put (call) investors is related with steeper slope in the decreasing- (increasing-) region, or higher state price of more negative (positive) returns. This is also consistent with the expensiveness of OTM put and call options we document in Table 2. While net long oil futures investors use OTM puts to hedge large futures price drops, net short investors trade OTM calls to hedge their potential loss due to large futures price increases.

4.3.3 Robustness Check

To further illustrate the dependence of risk-neutral skewness on investor beliefs, we perform two types of robustness exercises. The first exercise is to include belief measures of the equity market and the economy, control variables in the crude oil market that can affect futures returns, as well as macroeconomic variables which represent economic conditions and investment opportunities, in the regression model (13). In the other robustness check, we replace the dependent variable by another model-free measure of risk-neutral skewness (Bakshi, Kapadia, and Madan 2003; BKM hereafter), and we check whether the results obtained from density-based skewness hold.

Table 7 reports regression results when we include other belief measures and control variables. We include the lagged dependent variable to control for positive autocorrelation in all specifications. For ease of comparison, we only report coefficients of the belief measures, i.e., β_t, γ_t, and δ_t in (13). The coefficients of the additional control variables are omitted for brevity. Our conclusions from Table 6 hold when we consider belief measures in the equity market and the economy in column (1). Belief measures in the equity market have no significant impacts on the risk-neutral skewness and the shape of the SPDs in the crude oil market. In column (2), we include macro control variables, which are the 1-month Treasury-bill rate, the spread between yields on the 10-year Treasury bond and the 3-month Treasury bill, the difference in yields on Baa and Aaa corporate bonds, the Chicago Fed National Activity Index, and the log growth in aggregate industrial production over the last 12 months. These variables reflect the general economic condition and may affect the demand and supply of oil futures and therefore their expected returns. In column (3), we include control variables that determine futures returns in the crude oil market, namely, basis, storage level, historical returns, and IVOil. In column (4), we include both sets of control variables. Although adding control variables marginally increases R^2 in columns (2)-(4), the main results in column (1) and Table 6 still hold for both the significance level and the magnitude of coefficients of interest. We observe that the risk-neutral skewness is more negative (the marginal value of payoffs at large negative returns are higher) when there are higher OTM put open interest over OTM calls and when investors trade more OTM puts. The marginal value of payoffs at large positive returns are higher when there are higher level of belief dispersion in the call options. Lastly, a higher level of speculative activities after 2005 is related to more negative skewness and induces a higher probability of negative expected returns. It suggests that the increasing speculative activities after $2004-2005$ do have a significant impact on the futures market.

We next use another measure of risk-neutral skewness computed from OTM option prices (BKM, 2003). The numerical implementation of extracting risk-neutral moments follows Duan and Wei (2009). While our previous calculation of risk-neutral skewness in section

(3.4) is based on the densities, risk-neutral skewness of BKM (2003) is calculated by only utilizing OTM option prices. The two measures of risk-neutral variance have a correlation of about 95%; the two measures of skewness have a correlation of 70.04% after 2005.

Table 8 reports the regression results using the alternative measure of risk-neutral skewness. All columns in Table 8 have the same model specification as the regressions in Table 7. The main conclusions still remain valid: Higher OTM put demand implies more negatively skewed risk-neutral densities. OTM put (call) trading volume is negatively (positively) associated with risk-neutral skewness and therefore the steepness of the decreasing- (increasing-) region slope of the SPDs. The significantly negative coefficients of Speculation Index$\times$Dummy suggest that a higher level of speculation after the financialization around 2005 is associated with more negative skewness, indicating a higher marginal value of payoffs at large negative returns.

Overall, the regression results indicate that the slope of the SPDs is dependent on investor beliefs embedded in futures and options trade. In the aggregate investors assign higher state prices to more negative returns (i.e., the decreasing-region slope of the SPDs is steeper) when OTM put demand and OTM put trading volume are higher. More speculative positions after the financialization around 2005 induce more negative risk-neutral skewness and reinforce this effect. Higher level of belief dispersion in call options is associated with higher state prices in the increasing region, which means that the relative state price change with respect to positive returns becomes higher when investors overall trade more OTM calls. Investors' sentiment of the equity market such as bull-bear spread does not have a strong impact on the SPDs in the crude oil market. The findings are consistent and robust, regardless of whether we consider more control variables or use another measure of risk-neutral skewness. Our regression results based on risk-neutral skewness are also consistent with the results based on the direct measure of the slope of the SPDs using individual return intervals.

5 Conclusion

Using more than twenty years of futures and option data, we back out investor preferences towards different states in the crude oil market. We characterize the time variation in the SPDs, which are important for pricing contingent claims on crude oil. Moreover, we investigate the dependence of the SPDs on investor beliefs by utilizing the informative features of the crude oil derivatives market. Comparing with other financial markets, speculation data is available in this market and investors can take long or short positions based on their beliefs without being subject to regulations. We investigate how investor beliefs about oil futures prices affect the slope of the SPDs, which can be directly related to the marginal rate of substitution across states. Our regression analysis also examines the impact of speculation on the SPDs during the post-financialization period, when the level of speculative activities in the crude oil market is believed to be high.

We obtain three main results. First, we find that the SPDs in the crude oil market display a (time varying) U-shaped pattern. This implies that investors with various hedging or speculative purposes regard both extreme low or high oil futures prices as bad states of economy. We show that returns on both OTM calls and puts are negative and are consistent with the U-shaped SPDs. Second, we document the dependence of the SPDs on investor beliefs. When there are higher demands for OTM puts, the decreasing-region slope of the SPDs is steeper, and investors assign higher state prices to more negative returns. The slope of the SPDs is also affected by belief dispersions measured by options trading volume. These findings are robust under different model specifications and when we use the alternative measure of risk-neutral skewness. These findings support Bakshi, Madan, and Panayotov (2010) who show that the SPDs display a U-shaped pattern when investors have heterogeneous beliefs. Third, the increase in speculation after 2005 is associated with more negative risk-neutral skewness and higher marginal value of payoffs in those states of large negative returns.

The empirical findings in this paper suggest some extensions in the broader context of commodity markets. First, it would be interesting to develop theoretical commodity pricing models with heterogeneous beliefs. Prices are likely determined in an equilibrium where investors share risks and speculate based on their beliefs about the market. Sockin and Xiong (2014) spearhead this literature by analyzing how informational frictions imply heterogenous beliefs which in turn affect commodity prices. Second, our results highlight the importance of exploring the implications of heterogeneous beliefs for the relative demand and trading volume of commodity futures and options.

References

[1] Acharya, V., Lochstoer, L., Ramadorai, T., 2013. Limits to Arbitrage and Hedging: Evidence from Commodity Markets. *Journal of Financial Economics* 109, 441-465.

[2] Aït-Sahalia, Y., Lo, A., 1998. Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices. *Journal of Finance* 53, 499-547.

[3] Aït-Sahalia, Y., Lo, A., 2000. Nonparametric Risk Management and Implied Risk Aversion. *Journal of Econometrics* 94, 9-51.

[4] Aït-Sahalia, Y., Duarte, J., 2003. Nonparametric Option Pricing under Shape Restrictions. *Journal of Econometrics* 116, 9-47.

[5] Andersen, T.G., Bollerslev, T., 1998. Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts, *International Economic Review* 39, 885-905.

[6] Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2003. Modeling and Forecasting Realized Volatility, *Econometrica* 71, 579-625.

[7] Andersen, T, Bollerslev, T., Meddahi, N., 2011. Realized Volatility Forecasting and Market Microstructure Noise. *Journal of Econometrics* 160, 220-234

[8] Anderson, E., Ghysels, E., Juergens, J., 2005. Do Heterogenous Beliefs Matter for Asset Pricing? *Review of Financial Studies* 18, 875-924.

[9] Baker, M., Wurgler, J., 2006. Investor Sentiment and the Cross-Section of Stock Returns. *Journal of Finance* 61, 1645–1680.

[10] Bakshi, G., Kapadia, N., Madan, D., 2003. Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options. *Review of Financial Studies* 16, 101-143.

[11] Bakshi, G., Madan, D., Panayotov, G., 2010. Returns of Claims on the Upside and the Viability of U-Shaped Pricing Kernels. *Journal of Financial Economics* 97, 130-154.

[12] Bansal, R., Kiku, D., Shaliastovich, I., Yaron, A., 2014. Volatility, the Macroeconomy, and Asset Prices. *Journal of Finance* 69, 2471-2511.

[13] Barone-Adesi, G., Whaley, R., 1987. Efficient Analytic Approximation of American Option Values. *Journal of Finance* 42, 301-320.

[14] Bates, D. 1996. Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options. *Review of Financial Studies* 9, 69-107.

[15] Bates, D. 2000. Post-'87 Crash Fears in the S&P 500 Futures Option Market, *Journal of Econometrics* 94, 181-238.

[16] Beber, A. Brandt, M., 2006. The effects of Macroeconomic News on Beliefs and Preferences: Evidence from the Options Market. *Journal of Monetary Economics* 53, 1997-2039.

[17] Beber, A., Buraschi, A., Breedon, F., 2010. Differences in Beliefs and Currency Risk Premiums. *Journal of Financial Economics* 98, 415-438.

[18] Bessembinder, H. 1992. Systematic Risk, Hedging Pressure, and Risk Premia in Futures Markets. *Review of Financial Studies* 5, 637-667.

[19] Bessembinder, H., Chan, K., 1992. Time Varying Risk Premia and Forecastable Returns in Futures Markets. *Journal of Financial Economics* 32, 169-193.

[20] Black, F. 1976. The Pricing of Commodity Contracts. *Journal of Financial Economics* 3, 167-79.

[21] Breeden, D., Litzenberger, R., 1978. Prices of State-Contingent Claims Implicit in Option Prices. *Journal of Business* 51, 621-651.

[22] Brown, G., Cliff, M., 2004. Investor Sentiment and the Near-Term Stock Market. *Journal of Empirical Finance* 11, 1-27.

[23] Brown, G., Cliff, M., 2005. Investor Sentiment and Asset Valuation. *Journal of Business* 78, 405-440.

[24] Busch, T., Christensen B., Nielsen, M., 2011. The Role of Implied Volatility in Forecasting Future Realized Volatility and Jumps in Foreign Exchange, Stock, and Bond Markets. J*ournal of Econometrics* 160, 48-57.

[25] Buraschi, A., Jiltsov, A., 2006. Model Uncertainty and Option Markets with Heterogeneous Agents. *Journal of Finance* 61, 2841-2897.

[26] Büyükşahin, B., Harris, J., 2011. Do Speculators Drive Crude Oil Futures Prices? *The Energy Journal* 32, 167-202.

[27] Büyükşahin, B., Robe, M., 2014. Speculators, Commodities, and Cross-Market Linkages. *Journal of International Money and Finance* 42, 38–70.

[28] Carlin, B., Longstaff, F., Matoba, K., 2014. Disagreement and Asset Prices. *Journal of Financial Economics* 114, 226-238

[29] Carlson, M., Khoker, Z., Titman, S., 2006. Equilibrium Exhaustible Resource Price Dynamics. *Journal of Finance* 42, 1663-1703.

[30] Chabi-Yo, F., 2012. Pricing Kernels with Stochastic Skewness and Volatility Risk. *Management Science* 58, 624-640.

[31] Chaudhuri, R., Schroder, M., 2015. Monotonicity of the Stochastic Discount Factor and Expected Option Returns. *Review of Financial Studies,* forthcoming.

[32] Cheng, I.-H., Xiong, W., 2014. The Financialization of Commodity Markets. *Annual Review of Financial Economics.*

[33] Chernov, M., 2003. Empirical Reverse Engineering of the Pricing Kernel, *Journal of Econometrics* 116, 329-364.

[34] Christoffersen, P. Jacobs, K., 2004. Which GARCH Model for Option Valuation? *Management Science* 50, 1204-1221.

[35] Christoffersen, P., Heston, S., Jacobs, K., 2013. Capturing Option Anomalies with a Variance-Dependent Pricing Kernel. *Review of Financial Studies* 26, 1962-2006.

[36] Corsi, F., 2009. A Simple Approximate Long Memory Model of Realized Volatility. *Journal of Financial Econometrics* 7, 174–196.

[37] Datta, D., Londono, J., Ross, L., 2014. Generating Options-Implied Probability Densities to Understand Oil Market Events. Working Paper, Federal Reserve Board of Governors.

[38] Dennis, P., Mayhew, S., 2002. Risk-Neutral Skewness: Evidence from Stock Options. *Journal of Financial and Quantitative Analysis* 37, 471-493.

[39] De Roon, F., Nijman, T., Veld, C., 2000. Hedging Pressure Effects in Futures Markets. *Journal of Finance* 55, 1437-1456.

[40] Duan, J.-C., Wei, J., 2009. Systematic Risk and the Price Structure of Individual Equity Options. *Review of Financial Studies* 22, 1981-2006.

[41] Etula, E. 2013. Broker-Dealer Risk Appetite and Commodity Returns. *Journal of Financial Econometrics* 11, 486-521.

[42] Goldstein, I., Li, Y., Yang, L., 2013. Speculation and Hedging in Segmented Markets. *Review of Financial Studies* forthcoming.

[43] Hamilton, J., Wu, J., 2014. Risk Premia in Crude Oil Futures Prices. *Journal of International Money and Finance* 42, 9-37.

[44] Han, B. 2008. Investor Sentiment and Option Prices. *Review of Financial Studies* 21, 387-414.

[45] Henderson, B., Pearson N., Wang, L., 2014. New Evidence on the Financialization of Commodity Markets. Working paper, University of Illinois at Urbana-Champaign.

[46] Heston, S., Nandi, S., 2000. A Closed-Form GARCH Option Pricing Model. *Review of Financial Studies* 13, 585-626.

[47] Hong, H., Yogo, M., 2012. What Does Futures Market Interest Tell Us about the Macroeconomy and Asset Prices? *Journal of Financial Economics* 105, 473-490.

[48] Jackwerth, J. 2000. Recovering Risk Aversion from Option Prices and Realized Returns. *Review of Financial Studies* 13, 433-451.

[49] Kandel, E., Pearson, N., 1995. Differential Interpretation of Public Signals and Trade in Speculative Markets. *Journal of Political Economy* 103, 831-872.

[50] Kitsul, Y., Wright, J., 2013. The Economics of Options-Implied Inflation Probability Density Functions. *Journal of Financial Economics* 110, 696–711

[51] Kogan, L., Livdan, D., Yaron, A., 2008. Oil Futures Prices in a Production Economy with Investment Constraints. *Journal of Finance* 64, 1345-1375.

[52] Lemmon, M., Portniaquina, E., 2006. Consumer Confidence and Asset Prices: Some Empirical Evidence. *Review of Financial Studies* 19, 1499–1529.

[53] Li, H., Zhao, F., 2009. Nonparametric Estimation of State-Price Densities Implicit in Interest Rate Cap Prices. *Review of Financial Studies* 22, 4335-4376.

[54] Linn, M., Shive, S., Shumway, T., 2014. Pricing Kernel Monotonicity and Conditional Information. Working Paper, University of Michigan.

[55] Maheu, J., McCurdy, T., 2011. Do High-Frequency Measures of Volatility Improve Forecasts of Return Distributions? *Journal of Econometrics* 160, 69-76.

[56] Newey, W., West, K., 1987. A Simple Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55, 703-708.

[57] Ready, R., 2015. Oil Consumption, Economic Growth, and Oil Futures: A Fundamental Alternative to Financialization. Working paper, University of Rochester.

[58] Robe, M., Wallen, J., 2014. Fundamentals, Derivatives Market Information and Oil Price Volatility. Working paper, American University.

[59] Rosenberg, J., Engle, R., 2002. Empirical Pricing Kernels. *Journal of Financial Economics* 64, 341-372.

[60] Scheinkman, J., Xiong, W., 2003. Overconfidence and Speculative Bubbles. *Journal of Political Economy* 111, 1183-1219

[61] Singleton, K., 2014. Investor Flows and the 2008 Boom/Bust in Oil Prices. *Management Science* 60, 300-318.

[62] Sockin, M., Xiong, W., 2014. Informational Frictions and Commodity Markets. *Journal of Finance*, forthcoming.

[63] Song, Z., Xiu, D., 2014. A Tale of Two Option Markets: Pricing Kernels and Volatility Risk. Working paper, University of Chicago.

[64] Stambaugh, R., Yu, J., Yuan, Y., 2012. The Short of It: Investor Sentiment and Anomalies. *Journal of Financial Economics* 104, 288–302.

[65] Tang, K., Xiong, W., 2012. Index Investment and the Financialization of Commodities. *Financial Analysts Journal* 68, 54-74.

[66] Trolle, A., Schwartz, E., 2009. Unspanned Stochastic Volatility and the Pricing of Commodity Derivatives. *Review of Financial Studies* 22, 4423-4461.

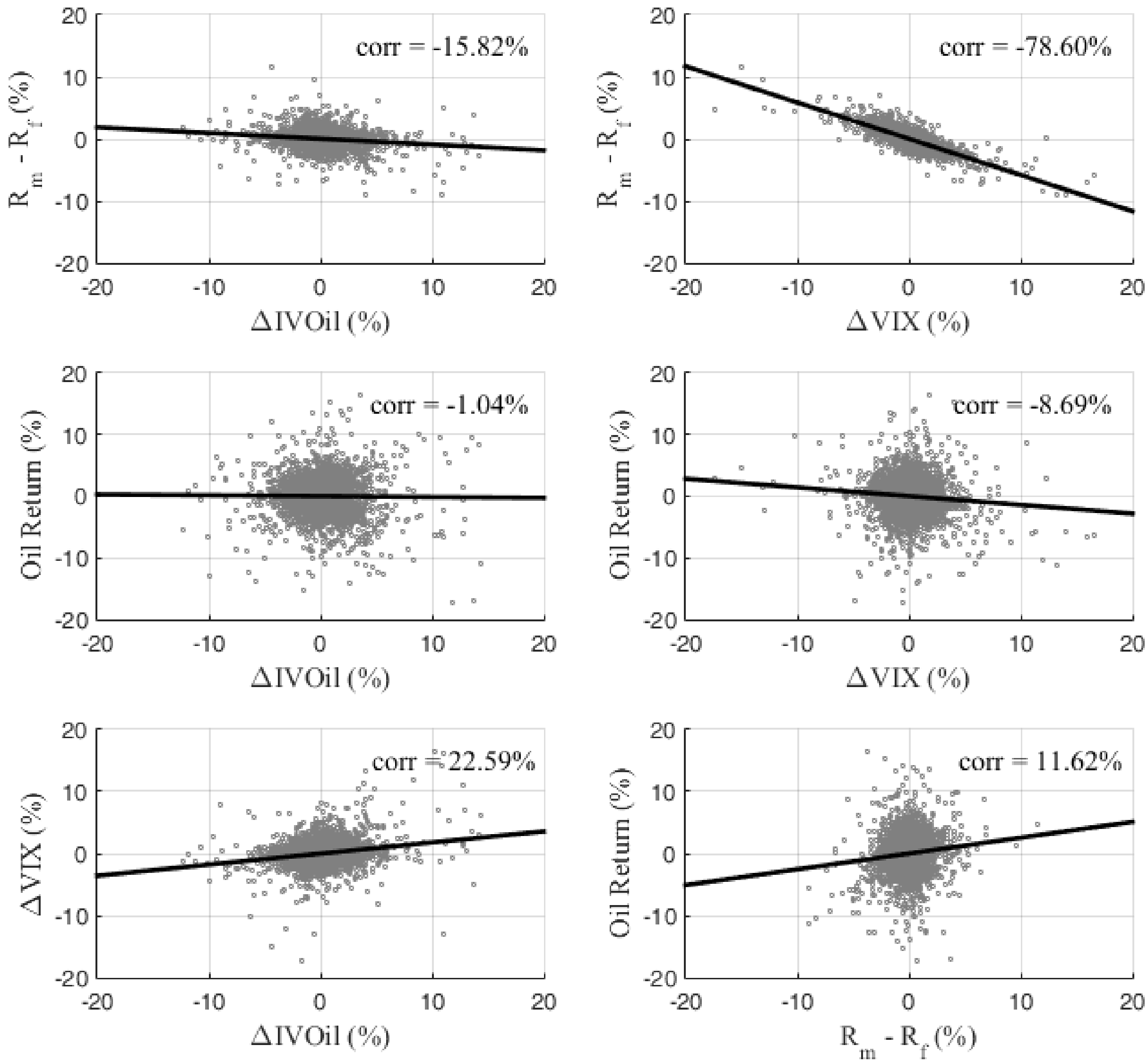
[67] Working, H. 1960. Speculation on Hedging Markets. Stanford University Food Research Institute Studies 1, 185-220.

[68] Xing, Y., Zhang, X., Zhao, R., 2010. What Does the Individual Option Volatility Smirk Tell Us About Future Equity Returns? *Journal of Financial and Quantitative Analysis* 45, 641-662.

[69] Xiong, W., Yan, H., 2010. Heterogeneous Expectations and Bond Markets. *Review of Financial Studies* 23, 1433-1466.

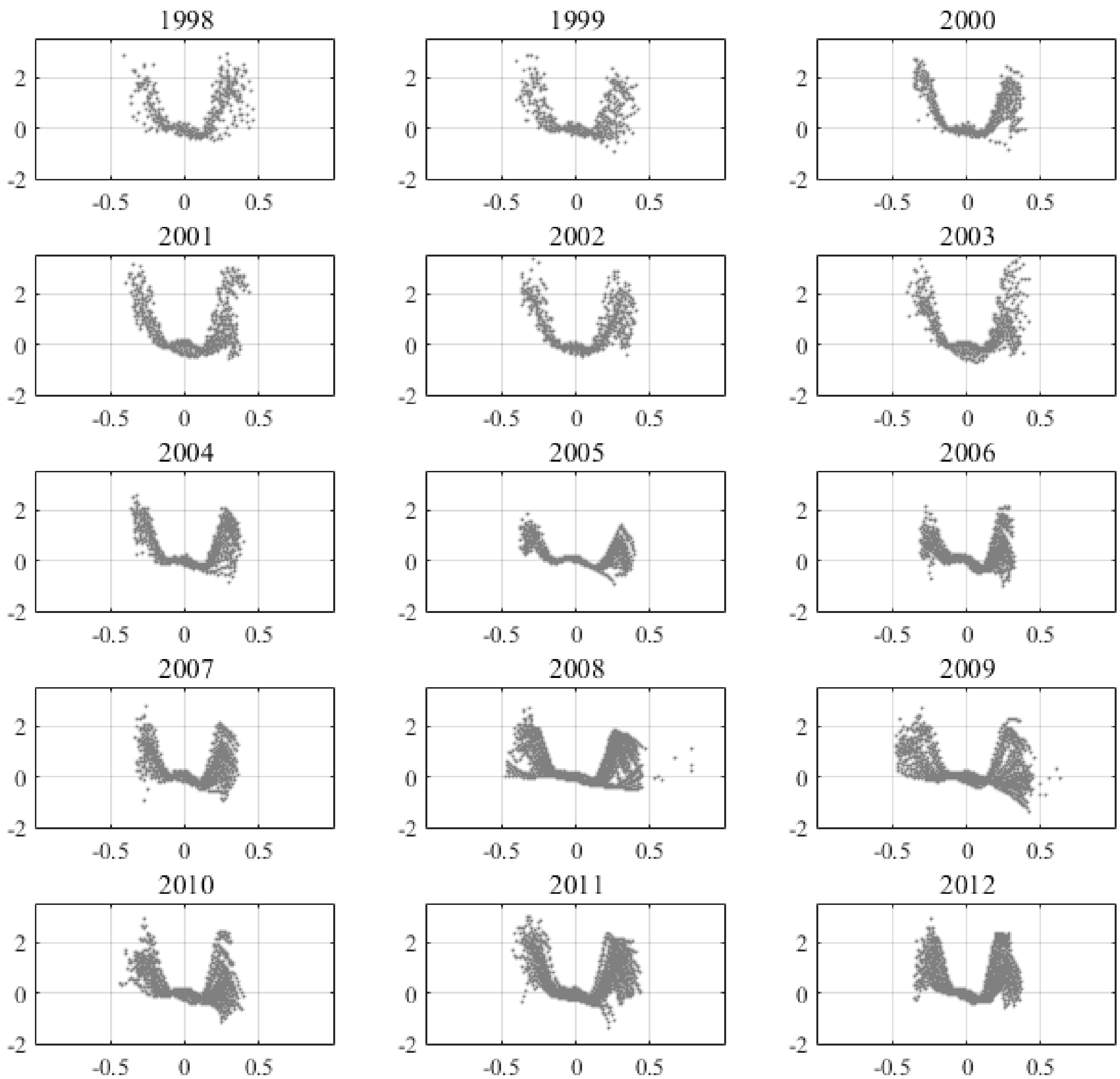
[70] Zhang, L., Mykland, P., Ait-Sahalia, Y., 2005. A tale of two time scales: determining integrated volatility with noisy high-frequency data. Journal of the American Statistical Association, 100, 1394-1411.

Figure 1: Scatter Plots of Stock Market Return, Oil Return, VIX, and Oil Volatility



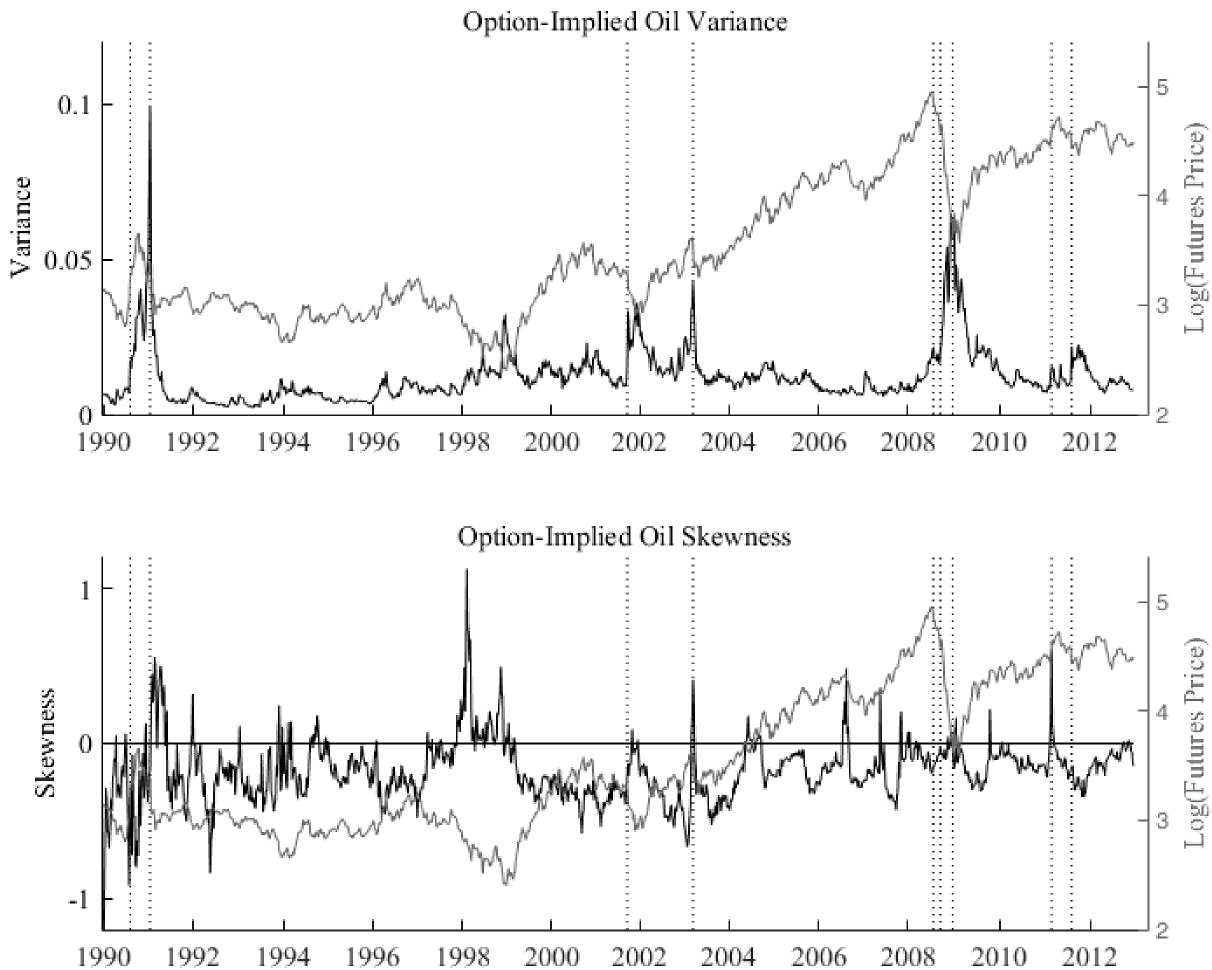
Notes: We report scatter plots and the linear regression line of daily changes in option-implied oil price volatility (ΔIVOil), daily changes in option-implied stock market volatility (ΔVIX), daily oil return calculated from the closest-to-maturity futures price and daily stock market excess return. The data period covers January 2, 1990 to December 31, 2012.

Figure 2: State Price Densities in the Crude Oil Market



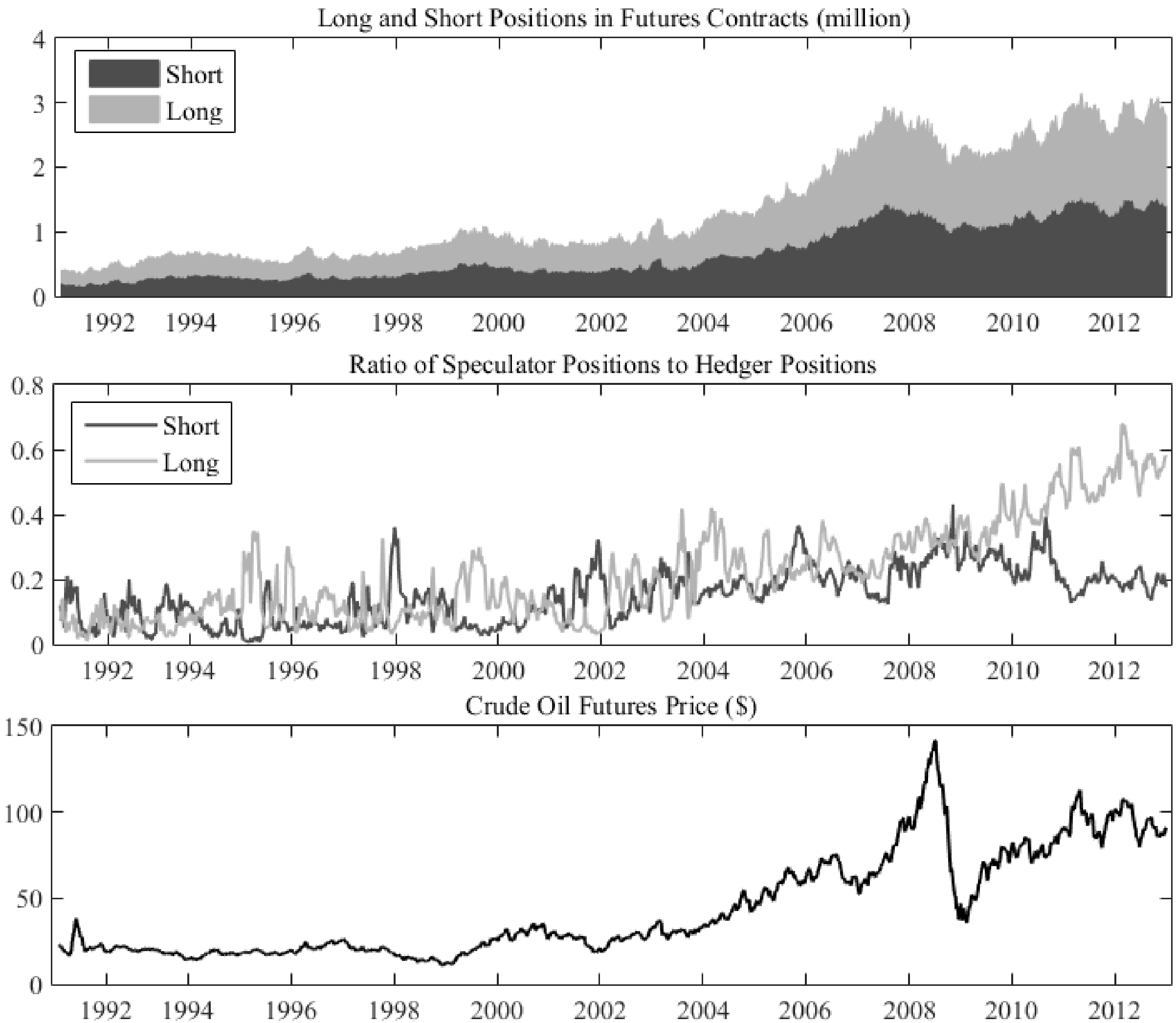
Notes: This figure shows the estimates of the state price densities (SPDs) in the crude oil market for the one-month horizon. For each year from 1998 to 2012 we plot the log ratio of the risk-neutral densities and physical densities of oil futures returns on each Wednesday. The horizontal axis is futures returns defined as $\log(F_T/F_{t,T})$.

Figure 3: Option-Implied Variance and Skewness



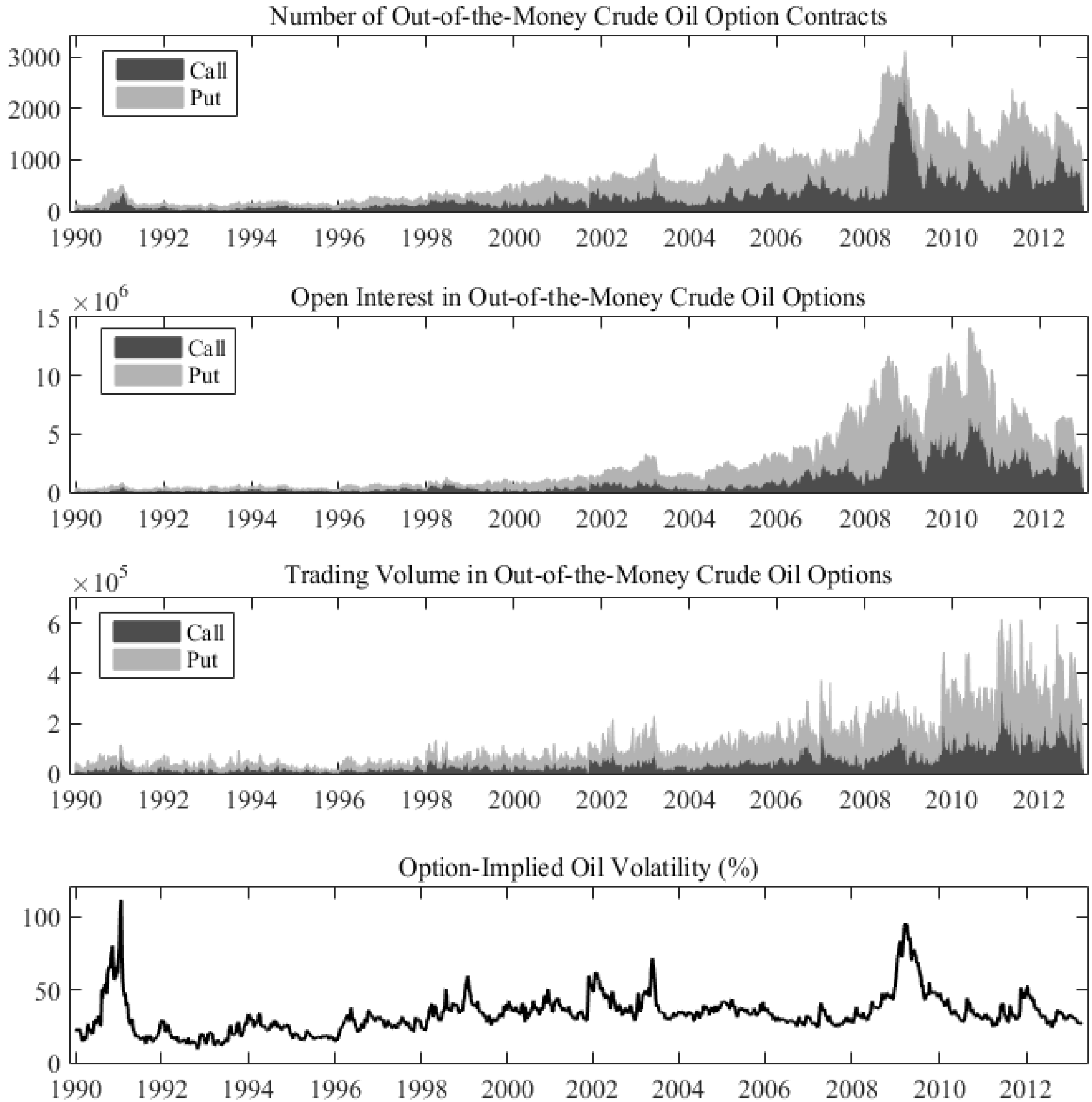
Notes: This figure shows weekly option-implied variance and skewness of one-month horizon crude oil futures returns on the left scale (in black) from January 1990 to December 2012. One the right scale we report the closest-to-maturity log futures price (in grey). We first obtain risk-neutral densities on each Wednesday. Variance and skewness are then calculated based on the estimated densities. From left to right, the first two vertical dotted lines denote the weeks of the Iraq invasion of Kuwait on August 2, 1990, and the liberation of Kuwait on January 17, 1991. Other vertical dotted lines denote the weeks of the September 11, 2001 attacks; the invasion of Iraq on on March 20, 2003; the WTI crude oil spot price peak (July 23, 2008 at $145.31); the Lehman Brothers bankruptcy (September 15, 2008); the WTI crude oil spot low point (December 23, 2008 at $30.28); the week when Libyan Civil War began and oil and gas production in Libya fell by more than 60%; and Standard & Poor's downgrade of the U.S. long-term sovereign credit rating from AAA to AA+.

Figure 4: Crude Oil Futures Positions Taken by Hedgers and Speculators



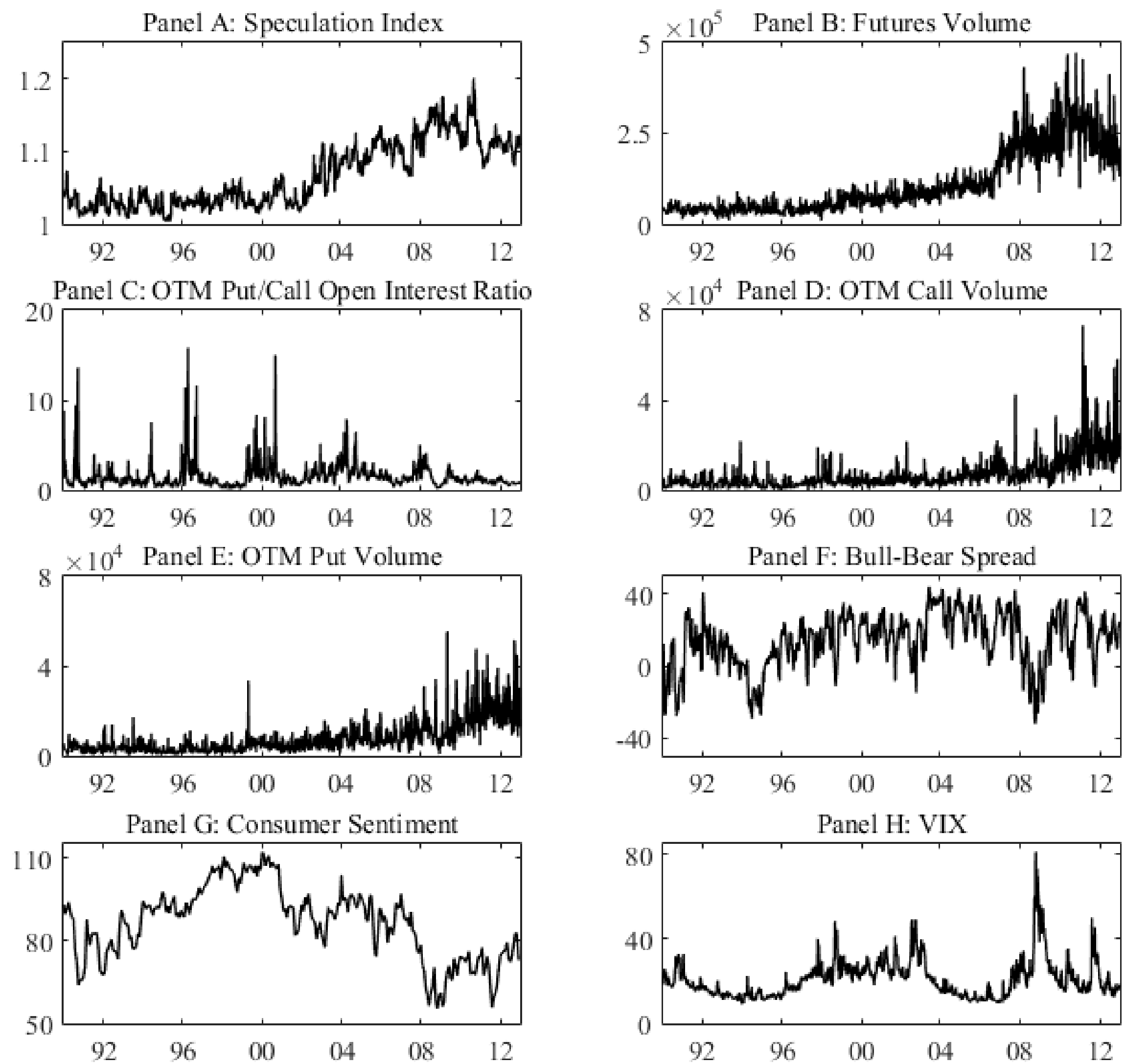
Notes: This figure shows long and short futures positions taken by hedgers and speculators from 1990 to 2012. Positions of hedgers and speculators are from the U.S. Commodity Futures Trading Commission (CFTC) futures-only Commitments of Traders (COT) report. As defined in the report, hedgers (commercials) are those investors who have direct exposure to the underlying crude oil commodity and use futures for hedging purposes; and speculators (non-commercials) are those investors who are not directly engaged in crude oil business activities and use derivatives markets for the purpose of financial profits. Before September 30, 1992, the CFTC publishes reports twice every month; then reports are available every week. The top panel displays the aggregate long (short) futures positions taken by large traders; the mid panel shows the ratio of speculator positions to hedger positions; the bottom panel reports the closest-to-maturity futures price averaged within the CFTC reporting week.

Figure 5: Crude Oil Options Contracts, Open Interests, Trading Volume, and Implied Volatility



Notes: We plot the number of out-of-the-money (OTM) crude oil option contracts, their open interests, and trading volume aggregated within a week. OTM calls have moneyness (K/F) > 1.05 and OTM puts have moneyness (K/F) < 0.95. We only report options with the maturity below 180 calendar days. In the bottom panel we report 30-day option-implied oil volatility averaged within each week. The sample covers January 2, 1990 to December 31, 2012.

Figure 6: Measures of Investor Beliefs



Notes: This figure shows different measures of investor beliefs in the crude oil market, the equity market and the overall economy. The speculation index measures the extent by which speculative positions exceed the necessary level to offset hedging positions as per the CFTC commitment of trader reports. Futures volume is based on contracts with a maturity of 30 days. OTM put/call open interest ratio divides OTM put by OTM call open interest. Open interest and trading volume of options are both measured with a maturity of 30 days. The bull-bear spread is defined as the percentage of bullish investors minus the percentage of bearish investors based on Investors Intelligence surveys. The consumer sentiment index is from the Surveys of Consumers from the University of Michigan. VIX is CBOE’s Volatility Index.

Table 1: Crude Oil Futures and Options Data

Panel A: Futures Data by Maturity (months)

	1	2	3	6	12	24	All
Number of Contracts	5759	5778	5780	5781	5774	4414	33286
Average Price	42.673	42.773	42.833	42.827	42.527	44.669	42.984
Average OI	141954	120005	62132	28914	15209	5701	64596
Average Volume	118041	64838	22997	4879	1334	323	36793
Average Volume/OI Ratio	83.15%	54.03%	37.01%	16.88%	8.77%	5.67%	56.96%

Panel B: Option Data by Maturity (days)

	(0,30]	(30,60]	(60,90]	(90,120]	(120,180]	(180,∞)	All
Number of Contracts	57032	69285	70475	65641	100265	398810	761508
Average IV	0.459	0.415	0.395	0.380	0.360	0.303	0.348
Average Price	5.799	5.532	5.690	6.127	6.334	10.768	8.466
Average OI	4207	3579	2919	2585	2197	1672	2299
Average Volume	401	213	109	73	51	21	84
Average Volume/OI Ratio	9.54%	5.94%	3.72%	2.83%	2.32%	1.28%	3.64%

Panel C: Option Data by Moneyness (K/F)

	(0,0.85]	(0.85,0.90]	(0.90,0.95]	(0.95,1.05]	(1.05,1.10]	(1.10,1.15]	(1.15,∞)	All
Number of Contracts	151885	62582	76275	165759	56497	43542	204968	761508
Average IV	0.369	0.329	0.316	0.307	0.317	0.327	0.395	0.348
Average Price	6.300	7.045	6.625	6.316	6.426	6.724	13.861	8.466
Average OI	2685	2409	2356	2227	2234	2261	2041	2299
Average Volume	52	87	100	152	103	87	39	84
Average Volume/OI Ratio	1.93%	3.59%	4.25%	6.81%	4.59%	3.87%	1.93%	3.64%

Notes: This table reports descriptive statistics on crude oil futures and options data from January 2, 1990 to December 31, 2012. Panel A presents daily futures contracts across maturities. Panels B and C report options data across maturities and moneyness on each Wednesday, with moneyness defined as strike price *K* divided by underlying futures price *F*.

Table 2: Average Return of Crude Oil Futures Options

Panel A: Put Option Returns	1	2	3	4	5	OTM - ITM		
Moneyness	(0.5, 0.85]	(0.85, 0.95]	(0.95, 1.05]	(1.05, 1.15]	(1.15, 1.5)	(1-5)	(1-4)	(2-5)
1990-2012								
Return	-0.8808***	-0.3091**	-0.1303	-0.1111**	-0.0565	-0.7406***	-0.6909***	-0.1709
t-stat	(-11.19)	(-2.04)	(-1.48)	(-2.11)	(-1.47)	(-5.02)	(-7.08)	(-0.71)
10^{th} Percentile	-1.0000	-1.0000	-1.0000	-1.0000	-0.5796	-1.3442	-1.8509	-1.1504
50^{th} Percentile	-1.0000	-1.0000	-0.9158	-0.2356	-0.0360	-0.9148	-0.7116	-0.7091
90^{th} Percentile	-1.0000	0.6479	1.8261	1.0348	0.5340	-0.3212	0.0000	0.2909
1990-2004								
Return	-0.9590***	-0.4476***	-0.1792	-0.1488**	-0.1090**	-0.8462***	-0.6638***	-0.4841***
t-stat	(-42.24)	(-3.08)	(-1.65)	(-2.35)	(-2.23)	(-13.41)	(-8.32)	(-3.71)
2005-2012								
Return	-0.8120***	-0.0709	-0.0386	-0.0454	0.0173	-0.6625***	-0.7125***	0.2507
t-stat	(-5.54)	(-0.22)	(-0.26)	(-0.49)	(0.28)	(-2.62)	(-4.35)	(0.47)

Panel B: Call Option Returns	1	2	3	4	5	OTM - ITM		
Moneyness	(0.5, 0.85]	(0.85, 0.95]	(0.95, 1.05]	(1.05, 1.15]	(1.15, 1.5)	(5-1)	(5-2)	(4-1)
1990-2012								
Return	0.0041	0.1470**	0.2737**	0.0985	-0.7740***	-0.8850***	-0.8904***	-0.4040**
t-stat	(0.07)	(2.38)	(2.41)	(0.37)	(-6.92)	(-12.59)	(-8.36)	(-2.54)
10^{th} Percentile	-0.7589	-1.0000	-1.0000	-1.0000	-1.0000	-1.5367	-2.0506	-1.2320
50^{th} Percentile	0.0001	0.0267	-0.6946	-1.0000	-1.0000	-0.9568	-0.9230	-0.8352
90^{th} Percentile	0.6544	1.2286	2.8071	2.2012	-1.0000	-0.2020	0.0000	-0.0143
1990-2004								
Return	0.0898	0.2437***	0.4292***	0.4108	-0.6184***	-0.9097***	-0.8376***	-0.2960
t-stat	(1.09)	(2.85)	(2.71)	(1.02)	(-2.86)	(-7.01)	(-4.37)	(-1.13)
2005-2012								
Return	-0.0781	-0.0046	-0.0177	-0.4447***	-0.9347***	-0.8623***	-0.9473***	-0.5076***
t-stat	(-1.06)	(-0.06)	(-0.13)	(-2.64)	(-22.92)	(-13.33)	(-11.78)	(-2.72)

Notes: This table reports the average return on options across moneyness. Moneyness is defined as strike price *K* divided by futures price *F*. On the third Wednesday of each month, we calculate hold-to-maturity returns of available options with the maturity as close to 30 calendar days as possible. We allocate available options to bins by moneyness, and returns are averaged within each bin. For each moneyness interval, we show the average return, its t-statistics, the 10th, 50th and 90th percentiles from 1990 to 2012, as well as average returns for two subsample periods: 1990 to 2004 and 2005 to 2012. We also report return differences between out-of-the-money (OTM) and in-the-moneyness (ITM) options. Returns on calls and puts are calculated as $r_{t,T}^{call} = \max(F_{t,T}e^{R} - K, 0)/c_{t,T} - 1$ and $r_{t,T}^{put} = \max(K - F_{t,T}e^{R}, 0)/p_{t,T} - 1$, where $F_{t,T}$ is the price of underlying futures, $R_{t,T} = \log(F_T/F_{t,T})$ is the return on the underlying futures contract during the period $T - t$, K is the strike price, $c_{t,T}$ and $p_{t,T}$ are prices of European style call and put options. T-statistics are reported in parentheses. ***, **, and * represent significant levels of 1%, 5% and 10%, respectively.

Table 3: Summary Statistics of Regression Variables

Variable	Mean	Std Dev	Percentile			P value (ADF test)
			25%	50%	75%	
Risk-Neutral Skewness	-0.1732	0.1932	-0.2888	-0.1824	-0.0782	0.0010
Speculation Index	1.0660	0.0437	1.0286	1.0490	1.1043	0.6714
Speculation Index (Adjusted)	0	0.0138	-0.0090	-0.0010	0.0085	0.0010
Futures Volume	109846	88097	44489	74400	158649	0.0010
Futures Volume (Adjusted)	0	31580	-12825	-1415	11935	0.0010
OTM Put/Call OI	1.5149	1.3863	0.7737	1.1617	1.7859	0.0010
OTM Call Volume	6710	7017	2562	4582	8177	0.0010
OTM Call Volume (Adjusted)	0	4718	-1994	-608	1135	0.0010
OTM Put Volume	7687	7423	2773	5241	9781	0.0010
OTM Put Volume (Adjusted)	0	4629	-2155	-733	1390	0.0010
Bull-Bear Spread	14.2516	15.2359	4.7000	16.2000	25.3000	0.0010
Consumer Sentiment	86.3806	13.2932	76.2000	88.4500	94.7000	0.4058
Consumer Sentiment (Adjusted)	-0.0688	4.2139	-2.5000	-0.2000	2.4000	0.0010
VIX (%)	20.9520	8.9127	14.4000	19.2650	24.8650	0.0120
Basis	0.0162	0.1998	-0.0845	0.0199	0.0851	0.0010
Storage Level	325329	25734	305639	327121	343221	0.6515
Storage Level (Adjusted)	0	10667	-7495	-631	7880	0.0010
IVOil (%)	33.3345	13.1565	26.2556	31.7878	37.9442	0.0456
Historical Return (%)	0.0296	1.0030	-0.5686	0.0791	0.6241	0.0010

Notes: This table reports summary statistics of variables used in the regression analysis (13). Variables are measured as closely as possible to each Wednesday from January 1990 to December 2012. Skewness of one-month futures returns is calculated from the risk-neutral densities implied by option prices, as defined by equation (10). The speculation index measures the extent by which speculative positions exceed the level necessary to offset hedging positions and is adjusted to remove the time trend. OTM Put/Call OI is the ratio of total open interests of OTM puts to OTM calls with a maturity of 30 days. Futures Volume, OTM Call Volume and OTM Put Volume are the trading volume of crude oil futures and OTM options with a maturity of 30 days adjusted for the time trend. The bull-bear spread is the proportion of bullish investors minus bearish investors based on the survey conducted by Investors Intelligence. Consumer sentiment is the Index of Consumer Sentiment from the University of Michigan and adjusted as its first differences. VIX is CBOE's Volatility Index. Basis is the discounted spread between futures and spot prices. Storage level is U.S. total stock of crude oil, excluding SPR, based on reports from the Energy Information Administration (EIA) and adjusted for the time trend. IVOil is the 30-day oil volatility of futures returns implied by option prices. Historical return is the moving average of daily returns of holding a long position of one-month futures contract during the previous week.

Table 4: Correlation Matrix of Regression Variables

	Speculation Index	Futures Volume	OTM Put/Call OI	OTM Call Volume	OTM Put Volume	Bull-Bear Spread	Consumer Sentiment	VIX	Basis	Storage Level	IVOil
Futures Volume	4.93%										
OTM Put/Call OI	7.05%	6.62%									
OTM Call Volume	-3.65%	33.43%	-12.13%								
OTM Put Volume	-0.25%	44.63%	11.73%	52.35%							
Bull-Bear Spread	-11.96%	2.51%	7.91%	-1.31%	0.34%						
Consumer Sentiment	-3.60%	2.45%	12.69%	1.32%	1.16%	22.67%					
VIX	10.72%	-1.28%	-4.70%	3.69%	-2.22%	-26.85%	-19.11%				
Basis	-5.33%	0.10%	-41.04%	5.15%	-5.59%	0.53%	-19.12%	4.65%			
Storage Level	-9.91%	5.57%	-1.19%	-3.54%	7.61%	1.37%	4.40%	-6.33%	28.82%		
IVOil	3.89%	1.61%	-0.28%	11.04%	2.61%	-8.54%	-29.34%	65.33%	12.76%	-9.04%	
Historical Return	-2.13%	-3.91%	35.47%	-13.31%	-0.91%	8.17%	4.18%	-5.74%	-13.49%	6.46%	-2.81%

Notes: This table reports the correlation matrix of variables used in the regression (13). Variables are measured as closely as possible to each Wednesday from January 1990 to December 2012. The speculation index measures the extent by which speculative positions exceed the level necessary to offset hedging positions and is adjusted to remove the time trend. OTM Put/Call OI is the ratio of total open interests of OTM puts to OTM calls with a maturity of 30 days. Futures Volume, OTM Call Volume and OTM Put Volume are the trading volume of crude oil futures and OTM options with a maturity of 30 days adjusted for the time trend. The bull-bear spread is the proportion of bullish investors minus bearish investors based on the survey conducted by Investors Intelligence. Consumer sentiment is the Index of Consumer Sentiment from the University of Michigan and adjusted as its first differences. VIX is CBOE's Volatility Index. Basis is the discounted spread between futures and spot prices. Storage level is U.S. total stock of crude oil, excluding SPR, based on reports from the Energy Information Administration (EIA) and adjusted for the time trend. IVOil is the 30-day oil volatility of futures returns implied by option prices. Historical return is the moving average of daily returns of holding a long position of one-month futures contract during the previous week.

Table 5: Investor Beliefs and the Slope of the SPDs in the Crude Oil Market

	Decreasing-Region Slope (a)		Decreasing-Region Slope (b)		Increasing-Region Slope (a)		Increasing-Region Slope (b)	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Speculation Index	0.0987**	0.103***	0.106**	0.119***	-0.0422	-0.0365	-0.0350	-0.0303
	(2.26)	(2.63)	(2.50)	(3.11)	(-0.97)	(-0.84)	(-0.80)	(-0.70)
Futures Volume	0.0289	0.0305	0.0217	0.0186	-0.00550	-0.00880	-0.0162	-0.0196
	(0.88)	(0.96)	(0.68)	(0.59)	(-0.19)	(-0.31)	(-0.53)	(-0.64)
OTM Put/Call OI	0.0815*	0.0810**	0.118**	0.116***	-0.116**	-0.122**	-0.0989*	-0.102*
	(1.92)	(2.00)	(2.55)	(2.75)	(-2.26)	(-2.34)	(-1.90)	(-1.94)
OTM Call Volume	0.0363	0.0404	0.0311	0.0371	0.0693*	0.0746**	0.0651*	0.0674*
	(0.92)	(1.11)	(0.78)	(1.00)	(1.84)	(2.04)	(1.69)	(1.79)
OTM Put Volume	0.0128	0.00872	0.0214	0.0189	0.00519	0.00236	0.00592	0.00587
	(0.36)	(0.25)	(0.61)	(0.55)	(0.14)	(0.07)	(0.15)	(0.15)
Bull-Bear Spread		-0.0584		0.0533		-0.00205		0.0468
		(-1.44)		(1.31)		(-0.05)		(1.05)
Consumer Sentiment		0.0281		0.00730		-0.0252		-0.0236
		(0.70)		(0.18)		(-0.62)		(-0.59)
VIX		-0.0822		-0.0514		-0.0720		-0.00838
		(-1.50)		(-0.90)		(-1.37)		(-0.15)
Adjusted R^2	0.017	0.023	0.026	0.030	0.018	0.021	0.013	0.013

Notes: This table presents the weekly regression of the slope of the SPDs on investor beliefs. The dependent variables in the left panel are the decreasing region (or negative return region) slopes (a) and (b), measured from return intervals (-0.27, -0.21) and (-0.25, -0.19), respectively; the dependent variables in the right panel are the increasing region (or positive return region) slopes (a) and (b), measured from return intervals (0.11, 0.17) and (0.12, 0.18), respectively. We do not report the intercept term in regressions to save the space. Newey-West t-stats with 3 lags are reported in parentheses. ***, **, and * represent the significant level of 1%, 5% and 10%, respectively. The sample covers January 2, 1990 to December 31, 2012.

Table 6: Investor Beliefs and Risk-Neutral Skewness

Panel A: Skewness						
	(1)	(2)	(3)	(4)	(5)	(6)
Speculation Index	0.0206					0.0340
	(0.93)					(1.49)
Speculation Index × Dummy	-0.0323					-0.0515*
	(-1.17)					(-1.86)
Futures Volume		0.0113				0.0134
		(0.78)				(0.85)
OTM Put/Call OI			-0.1110***			-0.0976***
			(-4.91)			(-4.54)
OTM Call Volume				0.0679***		0.0814***
				(4.10)		(4.11)
OTM Put Volume					-0.0171	-0.0547***
					(-1.29)	(-3.41)
Adjusted R^2	0.731	0.731	0.741	0.735	0.731	0.746

Panel B: AR(1) Residual Skewness						
	(1)	(2)	(3)	(4)	(5)	(6)
Speculation Index	-0.0091					0.0008
	(-0.29)					(0.02)
Speculation Index × Dummy	-0.0544					-0.0636
	(-0.80)					(-0.97)
Futures Volume		0.0129				0.0185
		(0.85)				(1.12)
OTM Put/Call OI			-0.1990***			-0.180***
			(-6.32)			(-5.92)
OTM Call Volume				0.0694***		0.0727***
				(4.33)		(3.98)
OTM Put Volume					-0.0241*	-0.0554***
					(-1.65)	(-3.44)
Adjusted R^2	0.017	0.016	0.091	0.033	0.018	0.105

Notes: This table presents the weekly regression of risk-neutral skewness on investor beliefs embedded in the crude oil futures and options markets. The dependent variable in Panel A is skewness based on the risk-neutral densities estimated from option prices, as defined in equation (10), and we include the lagged dependent variable in the regressions. The dependent variable and regressors in Panel B are the AR(1) residual in the regression of $X_t = a + b*X_{t-1} + \varepsilon_t$. The dummy variable is set to one during 2005 to 2012 and zero during 1990 to 2004. We do not report the intercept terms to save space. Newey-West t-stats with 12 lags are reported in parentheses. ***, **, and * represent significant levels of 1%, 5% and 10%, respectively. The sample covers January 2, 1990 to December 31, 2012.

Table 7: Investor Beliefs and Risk-Neutral Skewness.

Robustness to Control Variables

	(1)	(2)	(3)	(4)
Speculation Index	0.0334	0.0361*	0.0267	0.0294
	(1.53)	(1.72)	(1.16)	(1.31)
Speculation Index × Dummy	-0.0514*	-0.0631**	-0.0471*	-0.0575*
	(-1.93)	(-2.31)	(-1.65)	(-1.95)
Futures Volume	0.0134	0.0157	0.0098	0.0116
	(0.85)	(1.03)	(0.59)	(0.70)
OTM Put/Call OI	-0.0978***	-0.1090***	-0.0548***	-0.0666***
	(-4.45)	(-4.95)	(-2.64)	(-3.25)
OTM Call Volume	0.0823***	0.0810***	0.0657***	0.0652***
	(4.14)	(3.98)	(3.45)	(3.33)
OTM Put Volume	-0.0557***	-0.0542***	-0.0517***	-0.0519***
	(-3.46)	(-3.41)	(-3.39)	(-3.39)
Bull-Bear Spread	-0.0138	-0.0123	-0.0140	-0.0132
	(-0.94)	(-0.79)	(-0.91)	(-0.81)
Consumer Sentiment	-0.0073	-0.0091	-0.0060	-0.0060
	(-0.59)	(-0.68)	(-0.45)	(-0.44)
VIX	-0.0158	-0.0294	-0.0207	-0.0227
	(-1.09)	(-1.37)	(-1.42)	(-1.08)
Control Variables	No	Macro	Oil	Macro + Oil
Adjusted R^2	0.746	0.748	0.759	0.760

Notes: This table reports the weekly regression of risk-neutral skewness on investor beliefs and control variables. Risk-neutral skewness is estimated from option prices using equation (10). All regressions include the lagged dependent variable. The dummy variable is set to one during 2005 to 2012 and zero during 1990 to 2004. Column (2) includes macroeconomic control variables: the 1-month Treasury bill rate, the spread between yields on the 10-year Treasury bond and the 3-month Treasury bill, the difference in yields on Baa and Aaa corporate bonds, the Chicago Fed National Activity Index, and the log growth in aggregate industrial production over the last 12 months. Column (3) includes control variables in the crude oil market, which are the basis, storage level, historical returns, and oil option-implied volatility (IVOil). Column (4) includes both sets of control variables in columns (2) and (3). We do not report the intercept terms to save space. Newey-West t-stats with 12 lags are reported in parentheses. ***, **, and * represent significant levels of 1%, 5% and 10%, respectively. The sample covers January 2, 1990 to December 31, 2012.

Table 8: Investor Beliefs and Risk-Neutral Skewness.

Robustness to the Alternative Measure of Skewness

	(1)	(2)	(3)	(4)
Speculation Index	0.0238	0.0266	0.0237	0.0276
	(0.96)	(1.03)	(0.93)	(1.06)
Speculation Index × Dummy	-0.0668**	-0.0822***	-0.0797**	-0.1000***
	(-2.30)	(-2.60)	(-2.55)	(-2.86)
Futures Volume	0.0046	0.0082	0.0058	0.0095
	(0.29)	(0.52)	(0.35)	(0.57)
OTM Put/Call OI	-0.0365*	-0.0461**	-0.0517**	-0.0705***
	(-1.81)	(-2.47)	(-2.16)	(-3.00)
OTM Call Volume	0.0740***	0.0727***	0.0550**	0.0518**
	(3.72)	(3.52)	(2.49)	(2.24)
OTM Put Volume	-0.0761***	-0.0739***	-0.0733***	-0.0745***
	(-3.96)	(-3.88)	(-3.84)	(-3.87)
Bull-Bear Spread	0.0319	0.0314	0.0286	0.0266
	(1.37)	(1.34)	(1.18)	(1.11)
Consumer Sentiment	-0.0175	-0.0224	-0.0232	-0.0235
	(-1.01)	(-1.18)	(-1.35)	(-1.29)
VIX	0.0323	0.0206	0.0359*	0.0540*
	(1.58)	(0.87)	(1.70)	(1.85)
Control Variables	No	Macro	Oil	Macro + Oil
Adjusted R^2	0.598	0.603	0.610	0.616

Notes: This table reports the results of the weekly regression of risk-neutral skewness on investor beliefs using an alternative measure of risk-neutral skewness. Risk-neutral skewness is obtained by using the methodology of Bakshi, Kapadia, and Madan (2003). All regressions include the lagged dependent variable. Column (2) includes macroeconomic control variables: the 1-month Treasury bill rate, the spread between yields on the 10-year Treasury bond and the 3-month Treasury bill, the difference in yields on Baa and Aaa corporate bonds, the Chicago Fed National Activity Index, and the log growth in aggregate industrial production over the last 12 months. Column (3) includes control variables in the crude oil market, which are the basis, storage level, historical returns, and oil option-implied volatility (IVOil). Column (4) includes both sets of control variables in columns (2) and (3). We do not report the intercept terms in regressions to save the space. Newey-West t-stats with 12 lags are reported in parentheses. ***, **, and * represent significant levels of 1%, 5% and 10%, respectively. The sample covers January 2, 1990 to December 31, 2012.