Assessment of Density Forecast for Energy Commodities in Post-Financialization Era

Deepak Bisht
Arnab K. Laha

W. P. No. 2017-07-01
July 2017

INDIAN INSTITUTE OF MANAGEMENT

AHMEDABAD-380 015

INDIA

Assessment of Density Forecast for Energy Commodities in Post-Financialization Era

Deepak Bisht and Arnab K. Laha
Indian Institute of Management, Ahmedabad

Abstract

Probability density for the future price of an asset can be estimated from historical asset prices or exchange-traded derivatives. In this paper, prices of futures and options contracts that embed the forward-looking information are used to obtain the density forecast of the underlying asset under $\boldsymbol{Q}$- measure. Along with Probability Integral Transform (PIT), various statistical testes are conducted to determine whether the option-implied density forecast is unbiased under the real world measure, $\boldsymbol{P}$. We have worked with the settlement prices of NYMEX traded futures and options contracts for WTI crude oil and Henry Hub natural gas during the post-financialization period of 2006 to 2013. Statistical analysis of the PIT values indicate that the option-implied density forecast is unbiased under the real world measure, $\boldsymbol{P}$.

1 Introduction

Option prices are used to generate forecasts of underlying asset price often referred as risk neutral density (RND) forecast with the horizon equal to the time to expiration of the options. The density forecast is valuable for a variety of reasons in economics and finance. Central banks base their monetary policy decision on RND forecasts retrieved from the interest rate, and FX-options (Clews et al., 2000). In finance, RND plays a major role in risk management and regulatory reporting as it is used to calculate a variety of risk measures such as VaR, CVaR. Kostakis et al. (2011) used risk-adjusted implied RND for asset allocation. Options-implied RND reflects participants' expectations, and thus, is forward looking. Chiras and Manaster (1978) found that implied variance of future stock returns does better than the variance of historical stock price in forecasting the volatility of the future stock return. Since then several authors have studied the “forward-looking” aspects of options. A majority of them focused on the second moment of the option-implied distribution (see Poon and Granger (2003) for review). The predictive capability of FX option-implied correlation to forecast realized the correlation between two currency pairs had been studied by Campa and Chang (1998), and Castrén and Mazzotta (2005). Chang et al. (2012) studied implied beta and Kempf et al. (2015) used implied covariance estimates from option prices for portfolio optimization. All these studies suggest that use of option-implied measures can result in better forecasting performance than using historical data based measures.

Despite the fact that option-implied measures are forward-looking, they are usually contaminated by the risk premium, which can lead to inaccurate forecasts. For real world applications like risk management, it is beneficial to approximate the risk adjusted forecast from the option-implied RND. The well-calibrated approximation[1] of true density forecast is referred as real world density (RWD). Assuming market participants are rational, RWD forecast must coincide with their subjective density forecast (Bliss and Panigirtzoglou, 2004).

[1]Price evolution of a derivative can be described by a Partial Differential Equation (PDE) having risk-neutral drift. Theoretically, PDE's drift is composed of actual drift and a term which depends on the risk aversion of the representative agent. Cross-section of options prices are the solutions of this PDE and are silent on the decomposition. Getting accurate, true density from the option prices is impossible.

Unlike equity markets, there is a significant debate on the question of whether commodity futures price is a biased predictor of future expected price. The empirical work dates back to Houthakker (1957), as the author, suggests that small speculators have no price forecasting skill, perform quite well when they take a long side in futures. In line with the theory of normal backwardation, they capture the risk premium offered by the commercial hedgers. On the other hand, for cotton and wheat Rockwell (1967) reports close to zero returns for the strategy of going long (short) in futures markets when hedgers are net short (long). However, these authors did not include statistical tests to back their claim. Using statistical framework, Chang (1985) re-evaluates Rockwell's strategy and found that positive returns earned by speculators were not significantly different from those earned by naïve speculators. In the works of Dusak (1973) and Carter et al. (1983), risk premium embeds in futures price is related to the systematic risk. On the theoretical side, Stoll (1979) and Hirshleifer (1988, 1989) allow both systematic risk and hedging pressure to affect the futures prices. For agricultural commodities, Carter et al. (1983) and Bessembinder (1992) gave empirical evidence that systematic risk and hedging pressure are the determinants of premium in futures. De Roon et al. (2000) posit that cross-commodity hedging pressure also affects the risk premium in the individual commodity futures prices.

Several studies were done to find out the bias present in the crude oil futures price. Deaves and Krinsky (1992) found some evidence of positive risk premium in short-dated crude oil futures contracts during 1983 to 1990. In their study, Moosa and Al-Loughani (1994) concluded that price of NYMEX traded WTI crude oil futures are inefficient and biased forecasters of spot prices while Gulen (1998) found that short-dated WTI crude oil futures are an unbiased and efficient predictor of the spot price. Kolos and Ronn (2008) determined the sign and magnitude of the market price of risk and found crude oil and natural gas futures are biased downward. Bhar and Lee (2011) found time-varying risk premium in the crude oil market, which was influenced by the same risk factors that affect equity and bond markets. Hamilton and Wu (2014) reported significant variation in the risk premium due to the increased participation by long-only financial investors in crude oil futures markets in 2005. Authors claim compensation to speculators was positive on average before 2005, which is in line with the theory of normal backwardation. They posit that expected premium in having a long position in the futures market has gone down over time, even became significantly negative during contango.

For natural gas, the literature on risk premium in futures markets is quite thin. Buchanan et al. (2001) investigated the market-timing ability of large speculators for one-month forecast horizons. They reported the outperformance of large speculators relative to the large hedgers. However, after considering other traders' positions in a multivariate regression framework, authors found that the reason for this outperformance is the premium captured by the speculators instead of their services to the hedgers. Movassagh and Modjtahedi (2005) regress the spot prices against lagged futures prices after correcting for the correlation between the futures prices and error terms and found that US natural gas futures (with 3 to 12 months times-to-expiration) sell at a discount to the expected future spot price. On the contrary, Cartea and Williams (2008) and Haff et al. (2008) reported positive risk premiums in UK natural gas futures.

A majority of studies show implied volatility overestimates the realized volatility in commodity markets. Simon (2002) simulated a trading strategy of selling corn option straddles when implied volatility is higher than the out-of-sample volatility forecasts. He could not find profitable trading rules and concluded that the options are not necessarily overpriced. Doran and Ronn (2008) collected the daily prices of NYMEX traded futures and options data for heating oil, natural gas, and crude oil contracts from 1995 to 2005, and found significantly negative volatility risk premium for all three commodities. Wang et al. (2011) reported negative and time-varying variance risk premiums for corn from 1987 to 2009. Trolle and Schwartz (2010) found significant negative variance risk premia for crude oil and natural gas. However, the annualized Sharpe ratios of the strategy of shorting the energy variance are not high as that of shorting S&P 500 Index variance. Prokopczuk and Simen (2014), who studied 21 commodity markets from 1989 to 2011, further corroborate the claim of negative variance risk premium.

Evaluation of RND forecasts is primarily based on Probability Integral Transform (PIT) approach. Earlier PIT was used in the studies of Fackler and King (1990), Diebold (1998), Anagnou et al. (2002), Bliss and Panigirtzoglou (2004), Christoffersen and Mazzotta (2005), Liu et al. (2007), and Høg and Tsiaras (2011). In the context of commodities, an insufficient amount of work has been done. Fackler and King (1990) evaluated density forecasts for the corn, soybean, cattle, and hog markets during 1985-1988. They found density assessments for corn and live cattle futures are reliable while those for soybean and hog futures overstate volatility and understate location respectively. For the period 1994-2006, Høg and Tsiaras (2011) modeled the crude oil RNDs as GB2 distribution and found biasedness[2] in the density forecasts. They used parametric and nonparametric statistical calibration to obtain improved forecasts.

Commodity markets have undergone a profound transformation during the first decade of the 21st century. The risk premia in commodity derivative markets have been affected due to the three major developments. First, the unprecedented pouring of institutional money into the commodity futures often referred to as financialization of commodities. Historically negative correlation between the commodity futures returns, and stocks and bond returns (see Erb and Harvey, 2006; Gorton and Rouwenhorst, 2006) encouraged financial investors to lessen their portfolio risk by taking a long position in commodity futures. This has changed[3] the composition of participants over time (see Rouwenhorst and Tang, 2012; Cheng and Xiong, 2014) as the participation of financial investors has increased significantly over time. Several studies suggest that this has caused an increase in correlation between commodities and financial markets (see Daskalaki and Skiadopoulos, 2011; Silvennoinen and Thorp, 2013). Commodities such as crude oil bear high systematic risk and are now integrated with the global financial system (Ji and Fan, 2016). The effect of financialization on the risk premium is not straightforward[4]. As increased integration of commodity markets will result in more risk premium due to systematic risk while the greater participation of financial investors will weaken the Keynesian-Hicks hypothesis of normal backwardation and thus reduce the risk premium, or even push it to the other direction (see Bouchouev, 2012).

Second, commodity options markets, in particular, energy complex have witnessed a dramatic rise in the open interest and liquidity. A commercial entity may trade both the futures and options to hedge its exposure to stochastic spot price and volatility. As options are written on

[2]Null hypothesis for a joint test of normality and independence is rejected at 10% level of significance.

[3]Earlier commodity producers dominate the futures market. Financial investors filled the gap between producer and consumer hedging in lieu of a risk premium, i.e., positive expected return from holding a future contract until maturity (Keynes, 1923; Hicks 1939).

[4]As initiating a position in futures contract requires no investment; ideally, its expected return must be zero. If expected return differs significantly from zero, then it is due to the non-diversifiable risks of the underlying. In the context of commodities, two sources of risk premium are systematic risk and hedging pressure (see Hirshleifer (1989), de Roon et. al. (2000)).

the only one burdened by the hedging pressure, option market comes for its rescue[5]. Third, as a direct response to the subprime crisis in 2010, Dodd-Frank Wall Street Reform and Consumer futures, both the markets move in sync. Unlike futures, options have a visible premium that is paid by the buyer to the seller as the price for buying insurance. Thus, the futures market is not Protection Act was signed into law which places stricter capital and margin rules on dealers. Also, regulatory reporting requirement has been increased significantly. Several big banks like JP Morgan Chase and Barclays exited from the physical commodity trading businesses. The increase in the open interest of exchange-traded futures and options suggest that commercial hedging is moving to the public trading venues amid a decline in OTC activity (see Table1). The diminishing risk-bearing capacity of OTC dealers has significant implications for the risk premia in commodity derivatives markets (Etula, 2013).

Notional Amount (in billions of USD)	2011	2012	2013	2014	2015	2016
H1	3197	2993	2458	2206	1671	1392
H2	3091	2587	2204	1869	1320	-

Table 1: Declining OTC activity for commodity derivatives (Source: Bank for International Settlements)

It will be interesting to evaluate the forecasting ability of option-implied RNDs in the post-financialization era. To our knowledge, there has been no post-financialization study for the empirical performance of the option-implied density forecasts for commodity futures. In this study, we use PIT approach for density forecast evaluation. We consider NYMEX traded WTI crude oil (CL) and Henry Hub Natural Gas (NG) contracts[6] and the options written on them during 2006-2013. Most of the commodity indices give maximum weight to the energy commodities, making them more prone to the effects of financialization.

[5]According to a Bloomberg report, Mexico hedged its $9.5 billion of revenue from oil exports by buying put options that give it the right to sell oil for $38 a barrel for 2017 (Martin, 2016).

[6]CL and NG are the ticker symbols for the WTI crude oil and Henry Hub natural gas futures whereas LO and ON are the symbols for the options written on CL and NG contracts respectively. In this study, crude oil and natural gas options are referred as CL and NG options respectively.

2 Option Implied Risk Neutral Density

2.1 Non-parametric RND

With a complete set of options, one can recover the underlying asset's entire RND (Breeden and Litzenberger, 1978). The RND is then, $q(F_T) = e^{rT} \frac{\partial^2 C}{\partial K^2}\Big|_{K=F_T}$, where F_T is the price of the underlying asset at future date T, r is the non-stochastic interest rate, q is the implied risk neutral probability density and C denotes the price of the European style call option with strike price K. It is well known that, $C(K) = e^{-rT} \int_{x=K}^{\infty} (x - K) q(x) dx$. In practice, options with continuum of strikes do not exist. For example, CL options are listed at $0.50 per barrel strike increment. RND extraction, thus, is possible only over a range of traded strikes. The obtained option prices must be smoothed in order to have a well-behaved RND. One can use spline interpolation of implied volatilities in either strike-space (see Figlewski, 2008) or delta-space (see Malz, 1997). Tails of RND can be extended by using a well-known density function. For example, Shimko (1993) used lognormal distribution, and Figlewski (2008) appended tails from a Generalized Extreme Value (GEV) distribution.

2.2 Parametric RND

Estimating parameters when a well-known density function is used for modeling implied RND is much easier than obtaining the nonparametric density. The parametric approach usually takes up to six variables while non-parametric approach requires fitting of RND either pointwise or constructing it from linear or nonlinear segments. The number of parameters has no theoretical upper bound and can grow with the number of strikes available for RND estimation under non-parametric approach. By minimizing a loss function, which depends on the difference between market option price and model price, one can obtain parametric RND. We use GB2 and 2-MLN distributions for modeling the option-implied RNDs for crude oil and natural gas.

2.2.1 The GB2 Distribution

Bookstaber and McDonald (1987) proposed GB2 distribution for describing asset returns. GB2 distribution nests many popular distributions, like exponential, lognormal, log-logistic, t distribution. Four parameters describe GB2 density function, $\theta \equiv (a, b, c, d)$, as

$$q_{GB2}(x|\theta) = \frac{a}{b^{ac}B(c,d)} \frac{x^{ac-1}}{\left[1+\left(\frac{x}{b}\right)^{a}\right]^{c+d}}, \ x > 0, \tag{1}$$

where $B(c,d) = \frac{\Gamma(c)\Gamma(d)}{\Gamma(c+d)}$ is the well known Beta function, $\Gamma(.)$ is the gamma function, b is the scale parameter, and $a, c,$ and d are the shape parameters. When F_T follows the GB2(θ) distribution, the n[th] moment of F_T is given by, $E[{F_T}^n|\theta] = \frac{b^n B\left(c+\frac{n}{a}, d-\frac{n}{a}\right)}{B(c,d)}$, $n < ad$. Under $\boldsymbol{Q}$ −measure, current futures price is the expected futures price at the expiration date of options, T. Hence,

$$F_t = \mathbb{E}[F_T|\theta] = \frac{bB\left(c+\frac{1}{a},\ d-\frac{1}{a}\right)}{B(c,d)}. \tag{2}$$

McDonald and Bookstaber (1991) obtained an analytic expression for call option when F_T follows the GB2(θ) distribution. The GB2 call option price (with strike price K_i) at time t is,

$$C_{i|\theta} = F_t e^{-r(T-t)}\left(1 - Q_{GB2}\left(K_i|a, b, c + \frac{1}{a}, d - \frac{1}{a}\right)\right) - K_i e^{-r(T-t)}\left(1 - Q_{GB2}(K_i|a, b, c, d)\right), \tag{3}$$

where Q_{GB2} is the CDF of the GB2 distribution. Using Eq. (2) and (3) one can find expression for the call option price in terms of the beta distribution CDF (Q_β) given as,

$$C_{i|\theta} = Fe^{-r(T-t)}\left(1 - Q_\beta\left({K_i}'|c + \frac{1}{a}, d - \frac{1}{a}\right)\right) - K_i e^{-r(T-t)}\left(1 - Q_\beta({K_i}'|c, d)\right), \tag{4}$$

Where ${K_i}' = \frac{\left(\frac{K_i}{b}\right)^a}{1+\left(\frac{K_i}{b}\right)^a}$. Now one can obtain the values of the three parameters that provide the "best fit" to the observed option prices. The process of obtaining the parameter values is discussed in section 2.2.3.

2.2.2 The 2-MLN Distribution

Ritchey (1990) introduced the idea of using a mixture of lognormal (MLN) distributions in finance while obtaining the expression for call options. Guo (1998) has re-derived the Ritchey (1990)'s model under the assumption of the heterogeneous expectations of the investors about the lognormal process of the underlying asset price and proved that such a model would not admit any arbitrage. Melick and Thomas (1997) obtained the option-implied RND to infer the market's assessment of possible disruption in crude oil markets during the Gulf war. The 2-MLN density is the convex combination of the constituent lognormal densities, and can be described by five parameters, $\varphi = (w, \tau_1, \sigma_1, \tau_2, \sigma_2)$ as,

$$q_{MLN}(x|\varphi) = w\, q_{LN}(x|\mu_1, \sigma_1) + (1-w) q_{LN}(x|\mu_2, \sigma_2)\,, \tag{5}$$

where $\tau_i = e^{\mu_i + \frac{\sigma_i^2}{2}}$ for i = 1, 2. When F_T follows the 2-MLN$\left(w, \tau_1, \sigma_1\sqrt{T}, \tau_2, \sigma_2\sqrt{T}\right)$ distribution, the nth moment of F_T is given by,

$$E[{F_T}^n|\varphi] = w\, \tau_1^n \exp\left(\frac{1}{2}(n^2-n)\sigma_1^2 T\right) + (1-w)\tau_2^n \exp\left(\frac{1}{2}(n^2-n)\sigma_2^2 T\right).$$

Under $\boldsymbol{Q}$ $-$measure, futures price has zero drift and its expected price at the options expiration date (T) is,

$$F_t = \mathbb{E}[F_T|\varphi] = w\, \tau_1 + (1-w)\tau_2 \tag{6}$$

Under 2-MLN, an option price is the convex combination of two Black (1976) option prices,

$$C_{i|\varphi} = w\, C^{B76}(F_1, \sigma_1, r, T, K_i) + (1-w) C^{B76}(F_2, \sigma_2, r, T, K_i), \tag{7}$$

where $C^{B76}(F, \sigma, r, T, K) = e^{-r(T-t)}\left(FN\left(\frac{ln\left(\frac{F}{K}\right) + \frac{\sigma^2}{2}(T-t)}{\sigma\sqrt{T-t}}\right) - KN\left(\frac{ln\left(\frac{F}{K}\right) - \frac{\sigma^2}{2}(T-t)}{\sigma\sqrt{T-t}}\right)\right)$.

Noting that $\tau_i = F_i$ for $i = 1, 2$ and using Eq. (6) one can rewrite Eq. (7) as,

$$C_{i|\varphi} = w\, C^{B76}(F_1, \sigma_1, r, T, K_i) + (1-w) C^{B76}\left(\frac{F_t - w\, F_1}{1-w}, \sigma_2, r, T, K_i\right). \tag{8}$$

Now the values of the four parameters that provide the “best fit” to the observed option prices can be obtained following the method discussed in section 2.2.3 below.

2.2.3 Estimation of Parameters

The parameters for the parametric (GB2 and 2-MLN) RNDs are estimated by minimizing the root mean square error (RMSE) over N liquid strike prices defined i.e.

$$\hat{\xi} = \arg\min_{\xi}\left(\sqrt{\frac{\sum_{j=1}^{N}\left(C_i - C_{i|\xi}\right)^2}{N}}\right),$$

where C_i is the observed price of the European call option with strike price K_i, and $C_{i|\xi}$ is the fitted value when the parametric RND with parameter ξ is used. Differential evolution (DE) approach is used to perform the above minimization. The DE method uses an iterative approach to perform the minimization. Unlike classic optimization methods, DE does not require differentiable loss function (in our case RMSEs) and therefore can also be used when the loss function is not smooth or even discontinuous. However, it does not guarantee that the global optimal solution will be obtained in a finite number of iterations. More details regarding this method can be found in Storn and Price (1997). *DEoptim* package[7] in R is used to find the best parameters for GB2 and 2-MLN RNDs.

2.2.4 Berkowitz test

Our analysis is focused on evaluating the closeness of the option-implied density forecasts, q_t, of underlying futures, and the unknown data generating process p_t. This can be evaluated by using PIT of the ex-post futures price $\overline{F}_T$ at the date of expiration of the option taken with respect to the density forecast q_t. Mathematically, the PIT value is defined as,

$$y_t = \int_{f=0}^{\overline{F}_T} q_t(f)\, df = Q_T\left(\overline{F}_T\right). \tag{9}$$

If an RND forecast coincides with RWD then, y_t will be independent Uniform(0,1) distributed random variates. Since most goodness-of-fit tests such as Anderson-Darling, Kolmogorov-

[7]Details about *DEoptim* package is documented in Mullen et al. (2011).

Smirnov, and chi-squared tests assume that the observations are independent, one needs to establish the independence of $\{y_t\}$ before proceeding with these tests. Hence, for correct inference joint test for uniformity and independence is required. In this direction, Berkowitz (2001) suggested a joint parametric test to evaluate density forecasts. First, the PIT values undergo an inverse normal transformation given as, $z_t = \Phi^{-1}(y_t)$ where Φ is the CDF of Normal(0,1) distribution. Second, the hypothesis of independence, zero mean, and unit variance of transformed PIT values is jointly tested under the model,

$$z_t - \mu = \rho(z_{t-1} - \mu) + \varepsilon_t, \tag{10}$$

where ε is the innovation term with $E(\varepsilon_t) = 0$. The log-likelihood function for this model is well known and is given in Hamilton (1994) as,

$$L(\mu, \sigma_\varepsilon^2, \rho) = -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln\left(\frac{\sigma_\varepsilon^2}{1-\rho^2}\right) - \frac{(z_t - \mu/(1-\rho))^2}{2\sigma_\varepsilon^2/(1-\rho^2)} - \frac{T-1}{2}\ln 2\pi - \frac{T-1}{2}\ln \sigma_\varepsilon^2 - \sum_{t=2}^{T}\left(\frac{(z_t - \mu - \rho z_{t-1})^2}{2\sigma_\varepsilon^2}\right).$$

The null hypothesis of Berkowitz test is $\mu = 0, \rho = 0, and\ var(\varepsilon_t) = 1$. This hypothesis can be tested by using the likelihood ratio test statistic, $LR3 = -2(L(0,1,0) - L(\hat{\mu}, \hat{\sigma}_\varepsilon^2, \hat{\rho}))$. Under the null hypothesis, LR3 is distributed as $\chi^2(3)$. Eq. (10) is an AR (1) model which captures only a specific kind of serial dependence in the PIT values. It is possible to expand the model to AR (p) in a straight forward way but that increases the number of model parameters and leads to a decrease in the power (Bliss and Panigirtzoglou, 2004).

3 Data and RND Estimation

The forecasting capability of option-implied RNDs for one, two, and three months of horizon is investigated. Non-overlapping data[8] for CL and NG futures and options contracts traded on NYMEX is collected from April 2006 to December 2013 and January 2006 to December 2013 respectively. The settlement[9] prices on the option expiration days are retrieved from Bloomberg. Like Anagnou et al. (2002), only out-of-the-money (OTM) options are considered to obtain the RND. The deep OTM options that are out-of-sync with the options at neighboring strikes indicate the presence of arbitrage opportunities and hence, are ignored for this study. For CL and NG options, strikes are separated by 0.5 and 0.05 points respectively.

Futures are the underlying for the commodity options, and these contracts have different expiration dates. CL futures contracts expire on the third trading day prior to the 25th of the month preceding the delivery month whereas CL options contracts expire three trading days before the expiration of the underlying CL futures contract. Similarly, NG futures contracts will expire three trading days prior to the first day of the delivery month whereas NG options contracts expire one trading day before the expiration of the underlying NG futures contract. Due to this, one cannot take the observed price of an underlying futures contract as the expected price at the expiration of options under $\boldsymbol{Q}$ −measure. Thus, one needs to calculate synthetic futures price (F_T) which has same expiry date as of options contracts. It can be seen from the derivation in Appendix A that F_T is given as,

$$F_T = S_t \left(\frac{F_{T_F}}{S_t}\right)^{\frac{T-t}{T_F-t}}, \tag{11}$$

where S_t is the unobservable spot price which can be approximated by the price of nearby futures contract, and T_F is the date on which underlying futures contract will expire. A synthetic futures contract is a hypothetical (non-traded) security whose price represents the mean of the option-implied RND. Constant cost of carry up to the expiration date of options, which includes net convenience yield and financial cost is considered.

[8]Though ensuring that the date at which derivative prices are recorded is never before the previous options expiration date reduces the sample size; it makes autocorrelation negligible among the PIT values. For 1-month horizon, data for all months are considered. For 2-month horizon, data for only even months is considered. For 3-month horizon, data for January, April, July, and October is recorded.

[9]As intra-day call and put prices may not be synchronous across exercise prices, only settlement prices are reliable for empirical studies (Bahra, 1997).

Estimation of option-implied RND requires prices of European options. As CL and NG options are of American style, it is necessary to "de-Americanize"[10] them by approximating the residual value after excluding the early exercise premium. Like Trolle and Schwartz (2009) and Høg and Tsiaras (2011), we obtain lognormal implied volatilities by inverting the Barone-Adesi and Whaley (1987) formula. Afterward, the Black-76 formula is used to get European call prices from the implied volatilities. These equivalent call options are used to optimize the parameters for the RND forecast. As GB2 and 2-MLN RND functions must extrapolate the tails due to the unavailability of traded strikes, their tail shapes may differ. Figure 1 shows typical shapes of GB2 and 2-MLN option-implied RNDs. However, if there are a reasonably good number of liquid options across a wide range of strikes, the choice of a particular method does not matter much (Jackwerth, 2004).

Tables 2 and 3 provide the information about the median and interquartile range (IQR) of the number of strikes for CL and NG options considered with time to expiration of one, two, and three months respectively for several (Black-76 call) delta intervals. Tails are sufficiently well covered. Tables 4 and 5 summarize the RMSE (in $) for call options obtained under GB2 and 2-MLN parametrization of CL and NG option-implied RNDs respectively. For crude oil, median RMSE is up to three times the tick size ($0.01) for CL contract while for natural gas it is up to twice the tick size ($0.001) for NG contract. The magnitude of error is well below the typical bid-ask spread in the respective futures markets. 2-MLN distribution takes one parameter more than GB2 distribution and hence possibly yields lower pricing errors. The small pricing errors indicate both the parametric assumptions are well suited for modeling the crude oil and natural gas RNDs.

[10]Due to low interest rates (from Jan 2010 onwards), the market traded OTM American options are used as an approximation for European options and by using Black-76 model, implied volatility is retrieved.

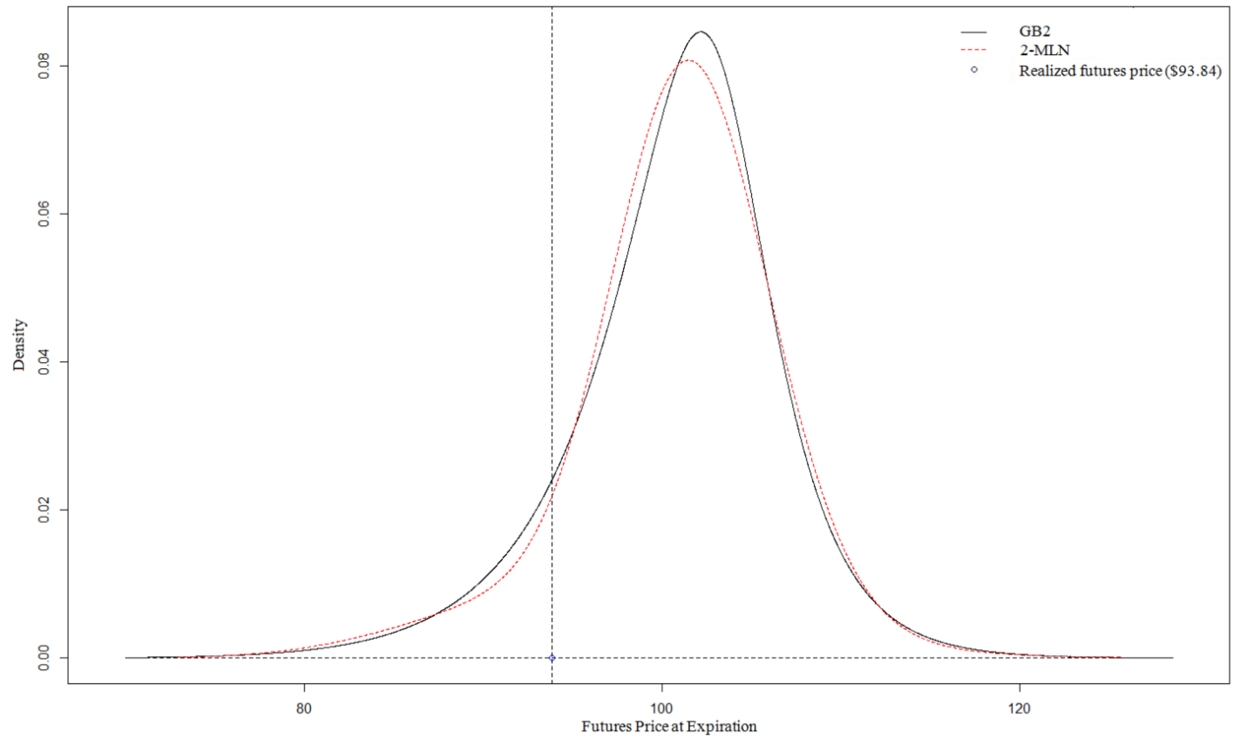

Figure 1: One-month option-implied RND forecasts for crude oil on October 17, 2013.

Delta Range	1-month	2-month	3-month
NOBS	93	47	31
(0.95,1)	14 (6)	14 (9.5)	13 (9)
(0.80, 0.95]	14 (5)	19 (7.5)	21 (8.5)
(0.60, 0.80]	9 (3)	13 (5)	15 (5)
(0.40, 0.60]	8 (2)	11 (4.5)	13 (4)
(0.20, 0.40]	9 (3)	13 (4.5)	14 (6.5)
(0.05, 0.20]	16 (8)	21 (10)	18 (14.5)
(0,0.05]	13 (5)	14 (7)	11 (11.5)
Total Strikes	84 (31)	107 (39.5)	108 (42)

Table 2: Summary statistics for the dataset of CL option prices (April 2006- December 2013).
*(IQR in parenthesis)
NOBS = number of observations.

Delta Range	1-month	2-month	3-month
NOBS	96	48	32
(0.95,1)	5 (4)	6 (3.25)	6 (4)
(0.80, 0.95]	7 (5.25)	8 (5)	10 (4.5)
(0.60, 0.80]	7 (7)	8 (6.5)	10 (6.75)
(0.40, 0.60]	6 (6.25)	8 (8.25)	10 (7.75)
(0.20, 0.40]	8.5 (8.25)	10.5 (10)	11 (7)
(0.05, 0.20]	11 (10)	13 (10)	15.5 (16.5)
(0,0.05]	7 (5)	9 (11.5)	10.5 (14.25)
Total Strikes	56 (48.5)	69 (41.5)	74 (30)

Table 3: Summary statistics for the dataset of NG option prices (January 2006- December 2013).
*(IQR in parenthesis)
NOBS = number of observations.

RMSE ($)	1-month		2-month		3-month	
	GB2	2-MLN	GB2	2-MLN	GB2	2-MLN
Minimum	0.009	0.004	0.009	0.009	0.008	0.007
Q1	0.016	0.014	0.020	0.019	0.024	0.024
Q2 (Median)	0.023	0.020	0.025	0.023	0.032	0.030
Q3	0.034	0.030	0.032	0.029	0.044	0.039
Maximum	0.138	0.113	0.122	0.066	0.108	0.105

Table 4: Summary statistics for the root mean square pricing error for crude oil call options with time to expiration of one, two, and three months. Tick size for CL contracts is $0.01.

RMSE ($)	1-month		2-month		3-month	
	GB2	2-MLN	GB2	2-MLN	GB2	2-MLN
Minimum	0.0003	0.0002	0.0002	0.0003	0.0004	0.0004
Q1	0.0006	0.0006	0.0008	0.0008	0.0009	0.0007
Q2(Median)	0.0011	0.0012	0.0016	0.0015	0.0019	0.0016
Q3	0.0026	0.0024	0.0031	0.0027	0.0031	0.0027
Maximum	0.0125	0.0089	0.0193	0.0106	0.0123	0.0117

Table 5: Summary statistics for the root mean square pricing error for natural gas call options with time to expiration of one, two, and three months. Tick size for NG contracts is $0.001.

4 Empirical Results

4.1 Graphical Analysis

Non-overlapping option-implied RND forecasts for both the commodities are generated on the option expiration dates for the horizon of one, two, and three months. With these, three time series (per commodity) of PIT values, $\{y_t\}$, are obtained. For each time series, the period between the dates of RND construction and option expiration do not overlap. As a result, PIT values can be viewed as random variates with no dependency with each other. This is evident in the middle panels ((b) and (e)) of Figures 2 to 7 as autocorrelations are not significantly different from zero for all the time-lags. In addition to this, correlograms of $\{y_t^2\}, \{y_t^3\}$, and $\{y_t^4\}$ (not shown in this paper) are evaluated, and no significant autocorrelation is found. This indicates that parametrized option-implied RND forecasts capture the volatility and skewness dynamics operative in the underlying CL and NG futures processes.

A PIT histogram is a useful tool in detecting biases. For example, a U-shaped histogram suggests underestimation of variance or kurtosis whereas a J-shaped histogram signals underestimation in the mean of true density forecast (Høg and Tsiaras, 2011). Another graphical tool that helps in detecting bias is the empirical CDF of PIT values. If PIT values are uniformly distributed between 0 and 1, then its CDF will coincide with the CDF of uniform distribution (represented by a straight line in panels (c) and (f) of Figures 2-7). As illustrated by Fackler and King (1990), a J-shaped pattern below the red line suggests underestimation of the mean, as most of the PIT value fall above 0.5. This usually happens during the normal backwardation when speculators capture the risk premium by going long in futures market against the commercial hedgers. An S-shaped (flat-steep-flat) CDF suggests that PIT values tend to fall in the center which results steepening of CDF in the center which indicates an overestimation of variability. This happens when option writers seek premium in excess of expected future volatility.

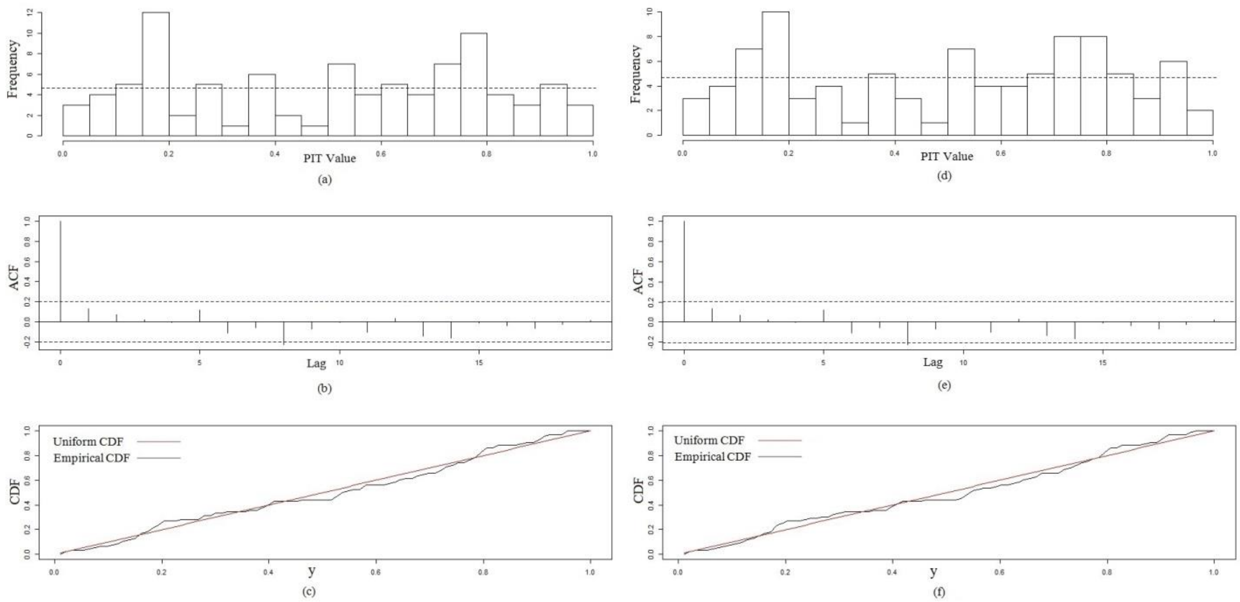

Figure 2: Histogram, correlogram, and CDF plots for PIT values obtained from the WTI crude oil option-implied RNDs of the one-month forecast horizon. Panel (a), (b), and (c) are obtained under the GB2 assumption while panel (d), (e), and (f) are obtained under 2-MLN parameterization of RNDs. It seems there is no major location and dispersion bias. (NOBS= 93)

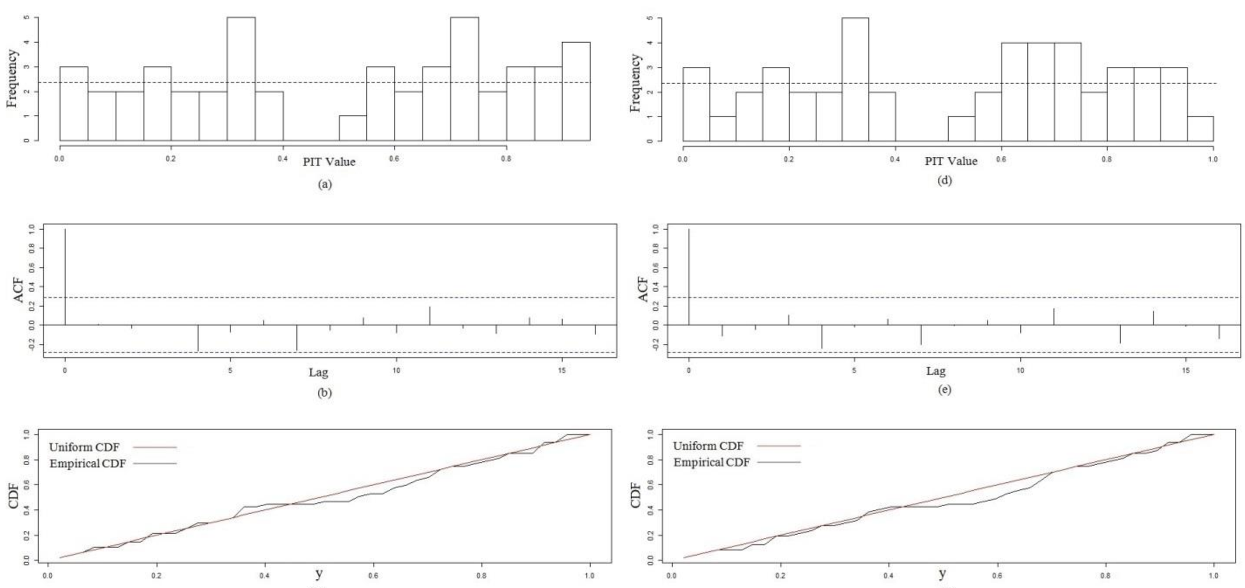

Figure 3: Histogram, correlogram, and CDF plots for PIT values obtained from the WTI crude oil option-implied RNDs of the two-month forecast horizon. Panel (a), (b), and (c) are obtained under the GB2 assumption while panel (d), (e), and (f) are obtained under 2-MLN parameterization of RNDs. It seems there is no major location bias, but one cannot rule out underestimation of dispersion. (NOBS=47)

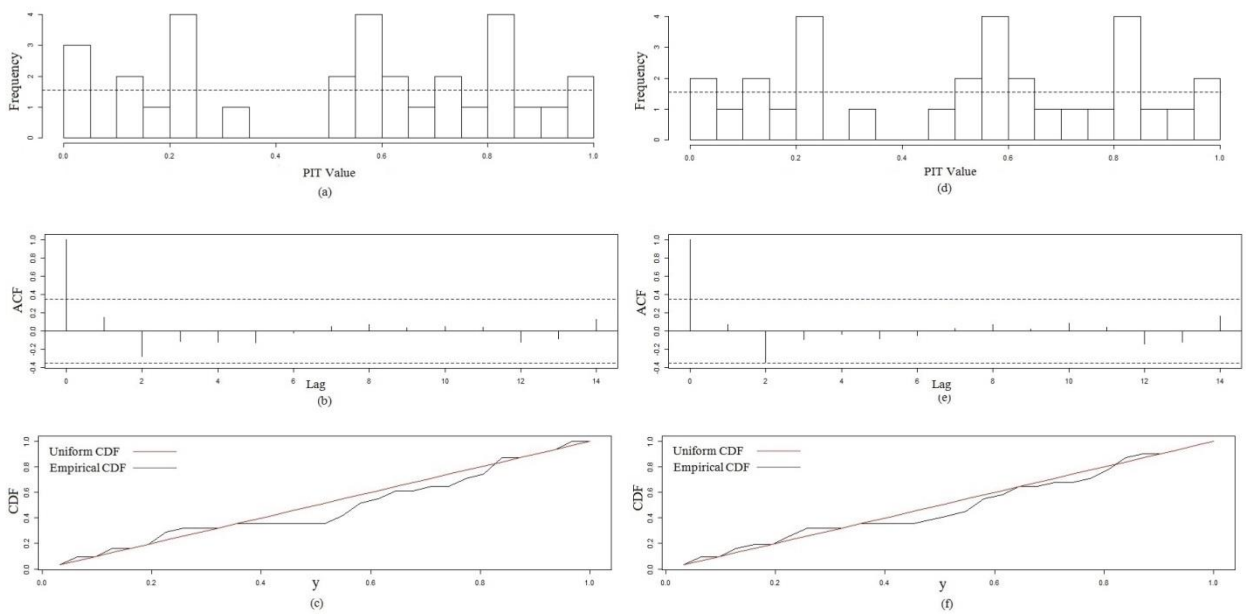

Figure 4: Histogram, correlogram, and CDF plots for PIT values obtained from the WTI crude oil option-implied RNDs of the three-month forecast horizon. Panel (a), (b), and (c) are obtained under the GB2 assumption while panel (d), (e), and (f) are obtained under 2-MLN parameterization of RNDs. It seems there is no major location bias but underestimation of dispersion is possible. (NOBS=31)

If option-implied RNDs are close to the RWDs, then the PIT histogram is of rectangular shape (see dashed lines in the panels (a) and (d) of Figures 2-7). For crude oil, PIT histograms for different horizons under GB2 and 2-MLN parametrization do not show any bias in location (see panels (a) and (d) in Figures 3-4). Also, there seems no dispersion bias in the density forecast of CL futures with one-month horizon. In Figures 3 and 4, PIT histograms for two and three-month horizons are relatively less populated in the middle due to which their empirical CDFs are flat in the center (panels (c) and (f)). This hints minor underestimation of dispersion.

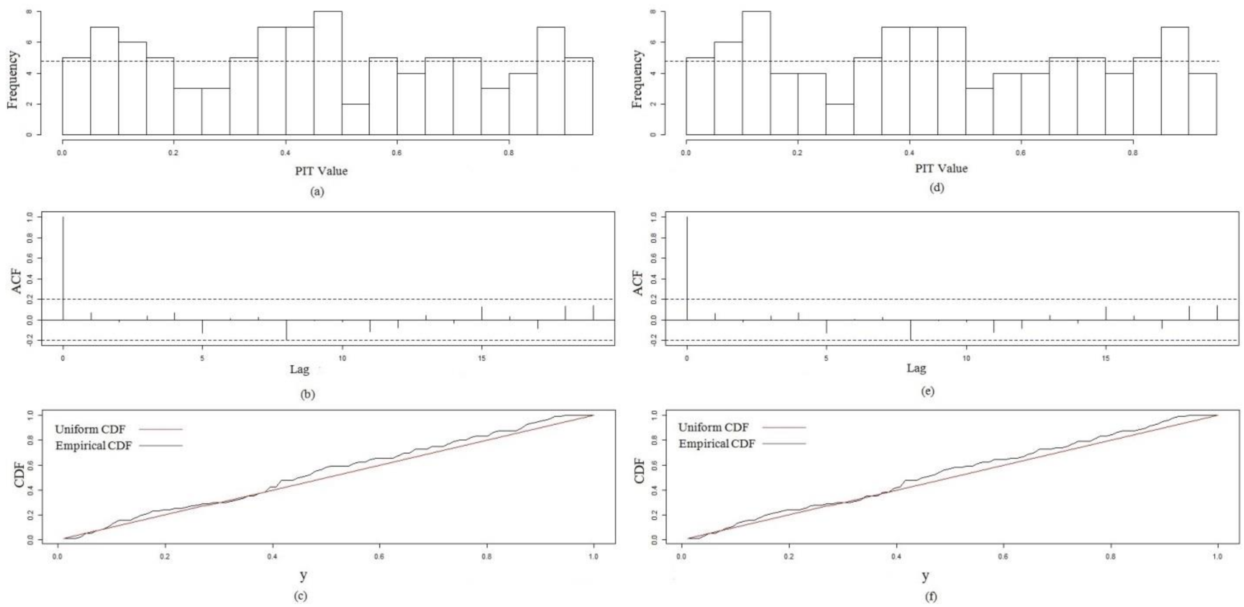

Figure 5: Histogram, correlogram, and CDF plots for PIT values obtained from the Henry Hub natural gas option-implied RNDs of the one-month forecast horizon. Panel (a), (b), and (c) are obtained under the GB2 assumption while panel (d), (e), and (f) are obtained under 2-MLN parameterization of RNDs. It seems there is a minor overestimation bias in location but no dispersion bias.

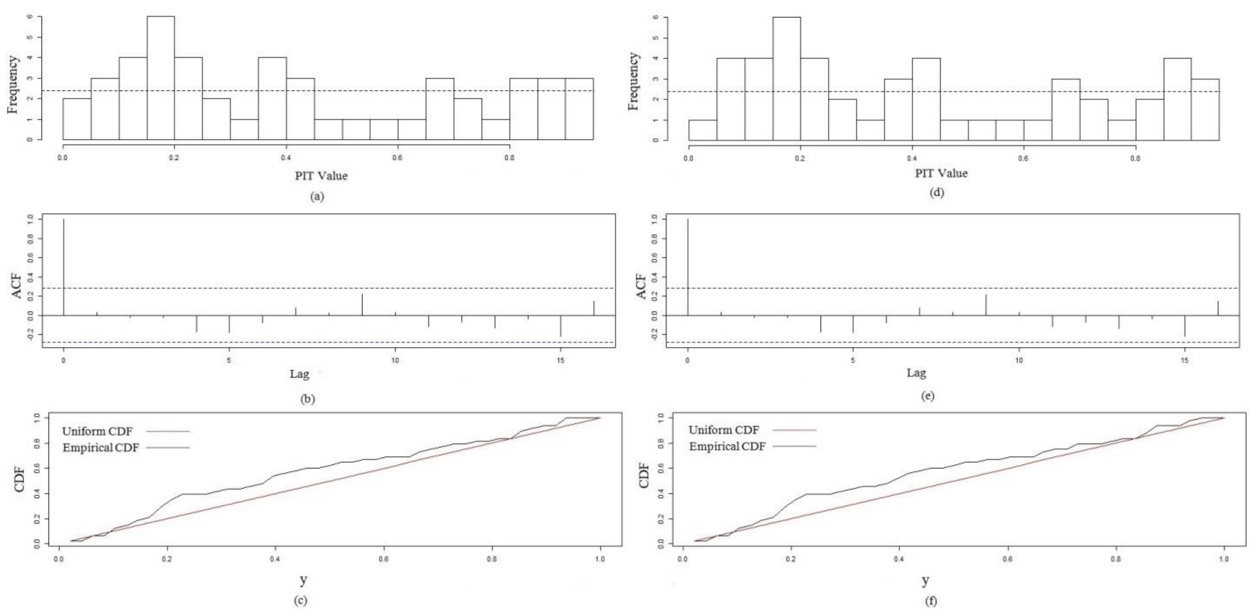

Figure 6: Histogram, correlogram, and CDF plots for PIT values obtained from the Henry Hub natural gas option-implied RNDs of the two-month forecast horizon. Panel (a), (b), and (c) are obtained under the GB2 assumption while panel (d), (e), and (f) are obtained under 2-MLN parameterization of RNDs. It seems location is overestimated with no major dispersion bias.

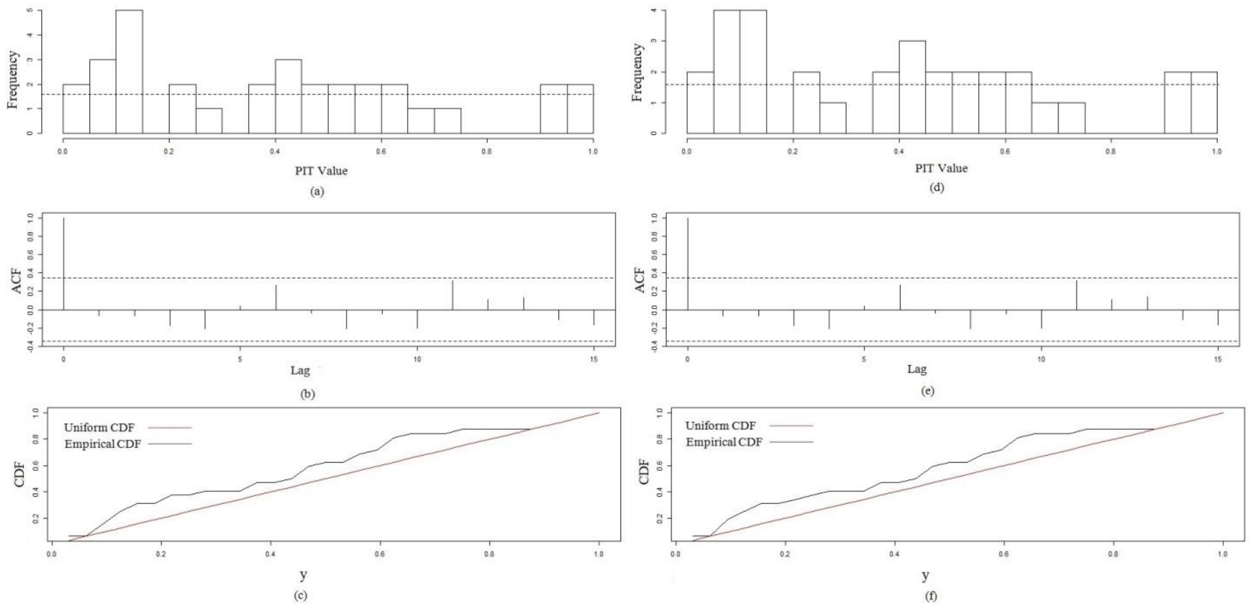

Figure 7: Histogram, correlogram, and CDF plots for PIT values obtained from the Henry Hub natural gas option-implied RNDs of the three-month forecast horizon. Panel (a), (b), and (c) are obtained under the GB2 assumption while panel (d), (e), and (f) are obtained under 2-MLN parameterization of RNDs. It seems location is overestimated with no major dispersion bias.

For natural gas, empirical CDFs are above the straight line which indicates the possibility overestimation in the mean of true distribution (see panels (d) and (f) in Figures 5-7). For the RNDs with the forecast horizon of two and three months, inverse J-shape is more evident than that of 1-month RND forecast. PIT histograms of two and three months are more concentrated towards the left (in Figures 6 and 7). However, there are no signs of bias in variability for all the three horizons for natural gas RNDs.

4.2 Statistical Tests

While graphical analysis focuses on the qualitative aspects and helps in understanding the data, formal tests are required to verify the suitability of a model. As noted earlier, if RND forecasts are unbiased then PIT values must be independent and distributed as Uniform(0, 1). Diebold et al. (1998) suggest that testing these assumptions separately may uncover potential weaknesses of the assumed model.

Kolmogorov–Smirnov (K–S) test is conducted to test whether the PIT values follow Uniform(0,1) distribution under the assumption of no serial correlation. The test statistic in the K-S test is the maximum distance vertical between the reference CDF and the empirical CDF. As a result, K-S test is more sensitive to discrepancies in the central region of the CDF. For crude oil and natural gas, p-values for GB2 and 2-MLN parametrizations for all the three horizons are reported in Tables 6(a) and 6(b) respectively. Results suggest that for crude oil, one cannot reject the hypothesis of uniform distribution of PIT values for all the forecast horizons at 5% level of significance. The same applies to the natural gas PIT values for one-month forecast horizon.

p-value	1-month	2-month	3-month
GB2	0.6346	0.7705	0.3528
2-MLN	0.5412	0.4199	0.6762
NOBS	93	47	31

Table 6(a): Berkowitz test on transformed crude oil values.

p-value	1-month	2-month	3-month
GB2	0.3860	0.0941	0.1076
2-MLN	0.5405	0.0958	0.1136
NOBS	96	48	32

Table 6(b): Berkowitz test on transformed PIT natural gas PIT values.

Berkowitz test on the time series of transformed PIT values. The null hypothesis is $\mu = 0,\ \rho = 0, and\ var(\varepsilon_t) = 1$ for an AR (1) process described in Eq. (10). The test statistic, LR3, has chi-squared distribution with three degrees of freedom as the mean (μ), variance (σ_ϵ^2), and autoregressive coefficient (ρ) are not restricted under the alternative. The obtained p-values are given in Tables 7(a) and 7(b).

p-value	1-month	2-month	3-month
GB2	0.4645	0.9710	0.8106
2-MLN	0.4895	0.8403	0.9566
NOBS	93	47	31

Table 7(a): Berkowitz test on transformed crude oil values.
(NOBS=number of observations)

p-value	1-month	2-month	3-month
GB2	0.2788	0.3739	0.6043
2-MLN	0.2892	0.3946	0.6297
NOBS	96	48	32

Table 7(b): Berkowitz test on transformed PIT natural gas PIT values.

Based on the p-values mentioned above, the null hypothesis that the time series of Berkowitz-transformed PIT values are independent with mean zero and have standard deviation one for all the cases cannot be rejected. Since Berkowitz test alone is not enough to address the normality of transformed PIT values (z_t) or innovations (ε_t), it must be complemented by Jarque-Bera (J-B) test (Dowd, 2004). The null hypothesis of J-B test is of zero skewness and, excess kurtosis. The J-B test statistic is asymptotically chi-squared with two degrees of freedom. The innovations are subjected to the J-B test (Dowd, 2004) to determine whether our RND parametrization has correctly accounted for the excess kurtosis. For crude oil and natural gas, results of J-B tests are summarized in Tables 8(a) and 8(b) respectively. This determines whether our RND parametrization has correctly accounted for the excess kurtosis.

p-value	1-month	2-month	3-month
GB2	0.2466	0.4582	0.5655
2-MLN	0.2437	0.5093	0.6373
NOBS	93	47	31

Table 8(a): J-B test on crude oil innovations

p-value	1-month	2-month	3-month
GB2	0.4761	0.7874	0.7874
2-MLN	0.4463	0.6907	0.8244
NOBS	96	48	32

Table 8(b): J-B test on natural gas innovations

(NOBS=number of observations)

Above results suggest that for both the commodities, the hypothesis of zero skewness and excess kurtosis of the distribution of innovations cannot be rejected at 5% significance level for all the three horizons. When combined with the results of Berkowitz test, we conclude that the time series of transformed PIT values are iid $N(0,1)$. Hence, option-implied RND forecasts for crude oil and natural gas futures are unbiasedand not significantly different from the true density forecasts or RWDs.

5 Summary and Conclusion

Option prices are a rich source of information about the distribution of the underlying asset price at a future time instant. Typically, futures and options prices include risk premia which may cause bias in this forward-looking information. Under $\boldsymbol{P}$ −measure, risk premia exists due to the factors, such as systematic risk and hedging pressure, whose effect cannot be hedged. Risk premia affects the cost of hedging, and knowledge of expected premia will help a firm in taking informed decision about its hedging program. Three phenomena that have affected the risk bearing ability of the traders in recent times are the financialization of commodity futures markets that started in the early 2000s, emergence of the options markets (both public and private) as the new venue for hedging, and the decline in activity of OTC commodity linked products (mainly swaps) due to the Dodd-Frank Wall Street Reform and Consumer Protection Act (which was signed into law in 2010). These changes motivate us to evaluate the quality of density forecasts implied from the commodity futures and options prices.

In this paper, post-financialization data from 2006 to 2013 is used to identify whether the option-implied RNDs of crude oil and natural gas are contaminated by risk premia or in other words, whether the RND forecasts for these energy commodities are specified correctly over the entire support region. The non-overlapping option-implied RNDs are modeled as GB2 and 2-MLN density functions. Several statistical tests have failed to reject the hypothesis that the RND forecasts for these commodities are unbiased which indicates that there are no statistically significant risk premia in the derivative prices of these energy commodities. Our result is at variance with the finding of Høg and Tsiaras (2011) as these authors found biasedness in the RND density forecasts for the crude oil for the pre-financialization period of 1994-2006. This indicates that the risk capital deployed in energy markets is no longer skewed. We feel the three changes mentioned in the above paragraph have affected the risk premia in energy markets. However, to pinpoint actual reasons and their significance requires further investigation, which is beyond the scope of the current work.

References

1. Anagnou, I., Bedendo, M., Hodges, S., & Tompkins, R. (2002). The relation between implied and realised probability density functions. Working paper, University of Technology, Vienna.
2. Bahra, B. (1997). Implied risk-neutral probability density functions from option prices: Theory and application. Working paper, Bank of England.
3. Barone-Adesi, G., & Whaley, R. (1987). Efficient analytic approximation of American option values. Journal of Finance, 42, 301–320.
4. Berkowitz, J. (2001). Testing density forecasts, with applications to risk management. Journal of Business and Economic Statistics, 19, 465–474.
5. Bessembinder, H. (1992). Systematic risk, hedging pressure, and risk premiums in futures markets. Review of Financial Studies, 5 , 637–667.
6. Bhar, R., & Lee, D. (2011). Time-varying market price of risk in the crude oil futures market. Journal of Futures Markets, 31, 779–807.
7. Black, F. (1976). The pricing of commodity contracts. Journal of Financial Economics, 3, 167–179.
8. Bliss, R., & Panigirtzoglou, N. (2004). Option-implied risk aversion estimates. Journal of Finance, 59, 407–446.
9. Bookstaber, R., & MacDonald, J. (1987). A general distribution for describing security price returns. Journal of Business, 60, 401–424.
10. Bouchouev, I. (2012). The inconvenience yield, or the theory of normal contango. Quantitative Finance, 12, 1733–1777.
11. Breeden, D., & Litzenberger, R. (1978). Prices of state-contingent claims implicit in option prices. Journal of Business, 51, 621–651.
12. Buchanan, W. K., Hodges, P., & Theis, J. (2001). Which way the natural gas price: An attempt to predict the direction of natural gas spot price movements using trader positions. Energy Economics, 23, 279–293.
13. Campa, J., & Chang, P. (1998). The forecasting ability of correlations implied in foreign exchange options. Journal of International Money and Finance, 17, 855–880.

14. Castrén, O., & Mazzotta, S. (2005). Foreign exchange rate option and returns based correlation forecasts— evaluation and two applications. Working Paper, No. 447, European Central Bank.
15. Cartea, Á., & Williams, T. (2008). UK gas markets: The market price of risk and applications to multiple interruptible supply contracts. Energy Economics, 30, 829–846.
16. Carter, Colin A., Gordon C. Rausser, & Andrew Schmitz. (1983). Efficient asset portfolios and the theory of normal backwardation. Journal of Political Economy, 91, 319–331.
17. Chang, E. C. (1985). Returns to speculators and the theory of normal backwardation. Journal of Finance, 40, 193–208.
18. Chang, B.-Y., Christoffersen, P., Jacobs, K., & Vainberg, G. (2012). Option-implied measures of equity risk. Review of Finance, 16, 385–428.
19. Cheng, I.H., & Xiong, W. (2014). Financialization of commodity markets. Annual Review of Financial Economics, 6, 419–441.
20. Chiras, D., & Manaster, S. (1978). The information content of option prices and a test for market efficiency. Journal of Financial Economics, 6, 213–234.
21. Christoffersen, P., & Mazzotta, S. (2005). The accuracy of density forecasts from foreign exchange options. Journal of Finanical Econometrics, 3, 578–605.
22. Clews, R., Panigirtzoglou, N., & Proudman, J. (2000). Recent developments in extracting information from options markets. Bank of England Quarterly Bulletin, 40, 50–57.
23. Daskalaki, C., & Skiadopoulos, G. (2011). Should investors include commodities in their portfolios after all? New evidence. Journal of Banking and Finance, 35, 2606–2626.
24. Deaves, R., & Krinsky, I. (1992). Risk premiums and efficiency in the market for crude oil futures. Energy Journal, 13, 93–117.
25. De Roon, F., Nijman, T., & Veld, C. (2000). Hedging pressure effects in futures markets. Journal of Finance, 55, 1437–1456.

26. Diebold, F., Gunther, T., & Tay, A. (1998). Evaluating density forecasts with applications to financial risk management. International Economic Review, 39, 863–883.
27. Doran, J. S., & Ronn, E. I. (2008). Computing the market price of volatility risk in the energy commodity markets. Journal of Banking and Finance, 32, 2541–2552.
28. Dowd, K. (2004). A modified Berkowitz back-test. Risk Magazine, 17, 86–87.
29. Dusak, K. (1973). Futures trading and investor returns: An investigation of commodity market risk premiums. Journal of Political Economy, 81, 1387–1406.
30. Erb, C.B., & Harvey, C.R. (2006). The strategic and tactical value of commodity futures. Financial Analysts Journal, 62, 69–97.
31. Etula, E. (2013). Broker-dealer risk appetite and commodity returns. Journal of Financial Econometrics, 11, 486–521.
32. Fackler, P. L., & King, R. P. (1990). Calibration of option-based probability assessments in agricultural commodity markets. American Journal of Agricultural Economics, 72, 73–83.
33. Figlewski, S. (2009). Estimating the implied risk-neutral density for the US market portfolio. In Tim Bollerslev, Jeffrey Russell, and Mark Watson (Eds.), Volatility and time series econometrics: Essays in Honor of Robert F. Engle. Oxford, UK: Oxford University Press.
34. Gorton, G., & Rouwenhorst, K.G. (2006). Facts and fantasies about commodity futures. Financial Analysts Journal, 62, 47–68.
35. Gulen, S. G. (1998). Efficiency in the crude oil futures market. Journal of Energy Finance and Development, 3, 13–21.
36. Guo, C. (1998). Option pricing with heterogeneous expectations. The Financial Review, 33, 81–92.
37. Haff, I. H., Lindqvist, O., & Loland, A. (2008). Risk premium in the UK natural gas forward market. Energy Economics, 30, 2420–2440.
38. Hamilton, J.D. (1994). Time Series Analysis. Princeton University Press, NJ.
39. Hamilton, J. D., & Wu, J. C. (2014). Risk premia in crude oil futures prices. Journal of International Money and Finance, 42, 9–37.
40. Hicks, J. R. (1939). Value and Capital. Oxford University Press, Cambridge.

41. Hirshleifer, D. (1988). Residual risk, trading costs, and commodity futures riskpremia. Review of Financial Studies, 1, 173–193.
42. Hirshleifer, D. (1989). Determinants of hedging and risk premia in commodity futures markets. Journal of Financial and Quantitative Analysis, 24, 313–331.
43. Høg, E., & Tsiaras, L. (2011). Density forecasts of crude-oil prices using option-implied and ARCH-type models. Journal of Futures Markets, 31, 727–754.
44. Houthakker, H. (1957). Can Speculators Forecast Prices? The Review of Economics and Statistics, 39, 143-151.
45. Jackwerth, J.C. (2004). Option-implied risk neutral distributions and risk aversion. Research Foundation of AIMR, Charlotteville, USA.
46. Ji, Q., & Fan, Y. (2016). Evolution of the world crude oil market integration: a graph theory analysis. Energy Economics, 53, 90–100.
47. Kempf, A., Korn, O., & Saßning, S. (2015). Portfolio optimization using forward-looking information. Review of Finance, 19, 467–490.
48. Keynes, J.M. (1923). Some aspects of commodity markets. Manchester Guardian Commercial, European Reconstruction Series Section 13, 784–786.
49. Kolos, S. P., & Ronn, E. I. (2008). Estimating the commodity market price of risk for energy prices. Energy Economics, 30, 621–641.
50. Kostakis, A., Panigirtzoglou, N., & Skiadopoulos, G. (2011). Market timing with option-implied distributions: a forward-looking approach. Management. Science, 57, 1231–1249.
51. Liu, X., Shackleton, M. B., Taylor, S. J., & Xu, X. (2007). Closed-form transformations from risk-neutral to real-world distributions. Journal of Banking and Finance, 31, 1501–1520.
52. Malz, A.M. (1997). Estimating the probability distribution of the future exchange rate from options prices. Journal of Derivatives, 5, 18–36.
53. Martin, E. (2016). Mexico's 2017 Oil Hedges Lock in $9.5 Billion Value: Takeaways. Bloomberg. Retrieved from https://www.bloomberg.com/news/articles/2016-08-29/mexico-s-2017-oil-hedges-lock-in-9-5-billion-value-takeaways
54. McDonald, J.B., & Bookstaber, R.M. (1991). Option pricing for generalized distributions. Communications in Statistics—Theory and Methods, 20, 4053–4068.

55. Melick, W., & Thomas, C. (1997). Recovering an asset's implied PDF from option prices: an application to crude oil during the gulf crisis. Journal of Financial and Quantitative Analysis, 32, 91–115.

56. Moosa, I.A., & Al-Loughani, N.E. (1994). Unbiasedness and time varying risk premia in the crude oil futures market. Energy Economics, 16, 99–105.

57. Movassagh, N., & Modjtahedi, B. (2005). Bias in backwardation in natural gas futures prices. Journal of Futures Markets, 25, 281–308.

58. Mullen, K., Ardia, D., Gil D. L., Windover, D., & Cline, J. (2011). DEoptim: An R package for global optimization by differential evolution. Journal of Statistical Software, 40, 1–26.

59. Poon, S.-H., & Granger, C. (2003). Forecasting volatility in financial markets: a review. Journal of Economic Literature, 41, 478–539.

60. Prokopczuk, M., & Wese Simen, C. (2013). Variance risk premia in commodity markets. WFA Meetings paper. Western Finance Association, US.

61. Rockwell, C. (1967). Normal Backwardation, Forecasting and the Returns to Commodity Futures Traders. Food Research Institute Studies, 7, 107-130.

62. Rouwenhorst, K.G., & Tang, K. (2012). Commodity investing. Annual Review of Financial Economics, 4, 447–467.

63. Shimko, D.C. (1993). Bounds of probability. RISK 6, 33–37.

64. Silvennoinen, A., & Thorp, S. (2013). Financialization, crisis and commodity correlation dynamics. Journal of International Financial Markets, Institutions & Money, 24, 42–65.

65. Simon, D. P. (2002). Implied volatility forecasts in the grains complex. Journal of Futures Markets, 22, 393–426.

66. Stoll, H. R. (1979). Commodity futures and spot price determination and hedging in capital market equilibrium. Journal of Financial and Quantitative Analysis, 14, 873–894.

67. Storn, R., & Price, K. (1997). Differential evolution- A simple and efficient heuristic for global optimization over continuous spaces. Journal of Global Optimization, 11, 341–359.

68. Trolle, A. B., & Schwartz, E. S. (2009). Unspanned stochastic volatility and the pricing of commodity derivatives. Review of Financial Studies, 22, 4423–4461.
69. Trolle, A. B., & Schwartz, E. S. (2010). Variance risk premia in energy commodities. Journal of Derivatives, 17, 15–32.
70. Wang, Z., Fausti, S., & Qasmi, B. (2011). Variance risk premiums and predictive power of alternative forward variances in the corn market. Journal of Futures Markets, 32, 587–608.

Appendix A: Synthetic Futures Prices

While working on the non-overlapping data, settlement prices of options contracts with times to expiration of one, two, and three months are recorded at the expiry date (t) of nearby option contracts. The underlying for expiring nearby options is the nearby futures contract, which will expire after three business days in case of WTI crude oil, and after one business day in case of Henry Hub Natural Gas. The price of nearby futures contract is used as a proxy for the unobservable spot price, S_t. The relation between the price of futures contract and underlying commodity spot price is given as,

$$F_{T_F} = S_t e^{c(T_F - t)}, \tag{A1}$$

where c is the annual cost of carry, i.e., the net of convenience benefits, financial cost, and storage cost incurred to buy and store a unit of physical commodity for one year. T_F is the expiration date of futures contract which is the underlying of options getting expired at date T. Rearranging the terms in Eq. (A1) will yield,

$$c = \frac{1}{T_F - t} \ln \frac{F_{T_F}}{S_t}. \tag{A2}$$

Due to the specification of contracts, $T_F > T$, one need to find the fair price of synthetic futures which is a hypothetical contract, expiring with the options, i.e., on date T. The no-arbitrage price of synthetic futures is,

By using Eq. (A2) in Eq. (A3) will yield,

$$F_T = S_t e^{\frac{T-t}{T_F - t} \ln \frac{F_{T_F}}{S_t}}$$

$$= S_t \left(\frac{F_{T_F}}{S_t} \right)^{\frac{T-t}{T_F - t}}$$