



Pricing Option on Commodity Futures under String Shock

Deepak Bisht
Arnab K. Laha

W. P. No. 2017-07-02
July 2017

The main objective of the working paper series of the IIMA is to help faculty members, research staff and doctoral students to speedily share their research findings with professional colleagues and test their research findings at the pre-publication stage. IIMA is committed to maintain academic freedom. The opinion(s), view(s) and conclusion(s) expressed in the working paper are those of the authors and not that of IIMA.



Pricing Option on Commodity Futures under String Shock

Deepak Bisht* and Arnab Kumar Laha**
Indian Institute of Management Ahmedabad

Abstract

Forward curve movements, particularly of industrial and energy commodities, suggests that futures price do not move in tandem with the spot price, and not all futures contracts move in the same direction. We incorporate these subtleties into our model with parsimony. This article offers a new approach to value commodity derivatives by using *string shock*. We use it to perturb the term structure of future convenience yield as if every futures contract has its source of risk. The no-arbitrage condition on the drift of future convenience yield and closed-form formula for the European call option written on a futures contract is derived. Our model has separate volatility and correlation functions that ensure easier parameterization and calibration to market data. We compare absolute and relative option pricing errors of our model with the two factor Schwartz (1997) model for 440 trading days. It is found that the new string shock based model has better performance than the Schwarz's model regarding having lesser pricing errors.

JEL Classification: C52, C61, G13

Keywords: Contingent Pricing, Crude Oil, Derivatives, Ornstein-Uhlenbeck sheet, Semimartingale, Wiener Process

* E-mail : deepakbisht@iima.ac.in Alternate E-mail: deepak_bisht@aol.in

**Corresponding Author. E-mail : arnab@iima.ac.in ; Phone : +91 79 66324947

Alternate E-mail: arnab.laha@gmail.com

1. Introduction

Santa-Clara and Sornette (SS) (2001) used string shocks for modeling the term structure of forward interest rate. Earlier implementation of a similar concept in the context of forward rate modeling can be found in Kennedy (1994, 1997), Goldstein (2000) and Kimmel (2004). Sornette (1998) takes the entire forward interest rate curve as a stochastic string and gives a general equation based on which interest rate derivatives can be priced and hedged. Longstaff et al. (2001) used the string-market model to study caps and swaptions, which combines the notion of string shock framework of SS (2001) and the market-model framework of Jamshidian (1997) and Brace et al. (1997). Han (2007) discusses a string-market model of interest rates incorporating both stochastic volatility and correlation.

This article presents a new arbitrage-free model to value commodity contingent claims when the future convenience yield is a random field driven by a string shock. In continuous time setting, we visualize the entire term structure of continuously compounded instantaneous future net convenience yield (hereafter, future convenience yield) as a piece of string under the influence of an infinite number of shocks acting across its length such that entire term structure remains continuous. These shocks are correlated with one another and are of different intensity. Collectively these shocks are referred as a *String Shock*. To the best of our knowledge, this is the first attempt to use string shock to model the dynamics of future convenience yield.

The seemingly less correlated movement between the short-dated and long-dated contracts is not uncommon and motivates us to use string shock which permits realistic correlation between futures contracts of different maturities. We decomposed the forward curve movement into *level shock* and *higher order shocks*. The former is due to the fluctuations in the spot price which forces forward curve to shift parallel while the latter is due to the

fluctuations in future convenience yields and distorts the shape of the forward curve. Another way to achieve realistic correlations and complex shapes of forward curve is to use a large number of state variables along with multidimensional Wiener process. Unfortunately, such a model lacks parsimony for which parameters are hard to estimate. A key feature of our approach from the practitioner's perspective is the separation of volatility and correlation functions. This allows independent parametrization of these functions and thus, calibration of the model to the observed prices becomes easier and yields realistic parameter values. This feature is absent from the earlier term structure models of future convenience yield such as Miltersen and Schwarz (MS) (1998) where authors used a multidimensional Wiener process to model the dynamics of spot price, future convenience yield, and forward rate. In such a case, the possible correlation among these three processes can be achieved only by specifying volatility functions in terms of correlation.

In commodities literature, term structure modeling has become more sophisticated during the last three decades. Spot price models like Gabillon (1991), Schwartz (1997), Cortazar and Schwartz (2003) to name a few, endogenously determine the futures price under partial equilibrium. The model determined futures prices are not in line with the observed forward curve. The whole process is contingent on the delicate task of estimating market price of state variable's risk which is model dependent. Even increasing the number of state variables may not yield a desirable fit to the observed forward curve. It is not very clear whether the introduction of a third factor other than spot price and convenience yield improves the performance of model or merely causes over-parameterization (Paschke and Prokopczuk, 2010).

The difficulties associated with spot price models can be avoided by using forward curve models which take the given initial forward curve and model its evolution under equivalent martingale measure, such as Cortazar and Schwartz (1994), Clewlow and Strickland (1999a,

1999b), and Crosby (2008a). This approach is based on no-arbitrage principle and motivated by the work of Heath et al. (1992; hereafter HJM) in the context of forward rate modeling. There is no need to estimate market price of non-tradable state variable's risk from the historical data as it is included in the observed futures prices. This kind of modeling is more amenable to derivative pricing. Various researchers identify a number of risk factors which dictate the majority of term structure movements. The suitable number of risk factors for a model depends on the commodity. For example, by applying Principal Component Analysis (PCA) to the historical futures price data, Cortazar and Schwartz (1994) found that only three factors are needed to explain 98% of daily return variance in copper futures while Koekebakker and Ollmar (2005) found more than ten factors are required to explain 95% of the movements in the term structure of electricity futures. The inclusion of more risk factors in a model increases its performance but also makes it more complex and non-parsimonious. Hence a trade-off is required. Also, fitting the initial term structure restricts the forward curve from taking any shape in the future, which is not desirable.

In the context of crude oil, by using Gibson and Schwartz (1990) model, Carmona and Ludkovski (2004) found sharp spikes in the time series plot of implied spot convenience yield obtained from futures contracts of different maturities. They also conclude that the implied spot convenience yields obtained by using 3-month and 12-month futures contracts are very different from each other. These authors suggested that each futures contract seems to have its source of risk and observed inconsistencies could be tackled by using the model based on the term structure of future convenience yield, pioneered by MS (1998). Their work is the hybrid of traditional spot price models and forward curve models as they specify the dynamics of spot price process and model a stochastic term structure of forward rates and convenience yields based on HJM framework. With no-arbitrage assumption, authors obtained the dynamics of forwards and futures prices and provide a general framework for

pricing options on them. MS (1998) introduced the notion of forward *convenience yield* and *future convenience yield*. The former is the forward value of the flow of benefits (per unit commodity per unit time) that accrues to the owner of the physical commodity at the contract maturity but not to the owner of the forward contract whereas the latter has no straightforward economic interpretation due to the continuous resettlement inflows and outflows of the futures contract. However, once interest rates are assumed to be deterministic, there remains no difference between the two, and the notion of the forward convenience yield applies to the future convenience yield (Miltersen, 2003). Björk and Landen (2002) extended the work of MS (1998) to the point process. The literature is sparse on the term structure modeling of convenience yield. Our model belongs to this category of models and is close in spirit to MS (1998). Like other forward curve models, our model does not require the estimation of the market price of any risk as it is embedded in observed derivative prices.

Commodity spot prices seem to exhibit mean reversion (Bessembinder et al., 1995; Casassus and Collin-Dufresne, 2005). Under the risk-neutral probability measure, standard no-arbitrage argument determines the drift of spot price and futures convenience yield processes. The drift of spot price process depends (negatively) on the level of spot convenience yield. If one makes sure that spot price and spot convenience yield are always positively correlated, then the spot price becomes a mean reverting process. For option pricing, seasonal effects can be captured by allowing the spot price and future convenience yield volatilities to be parametrized as sinusoidal functions.

The paper is organized as follows. After the introduction, Section 2 reviews the string shock and its qualifying conditions given by SS (2001). We will also discuss a specific string shock, Ornstein-Uhlenbeck (O-U) Sheet, which we will use later in this paper. In Section 3 we present the model and derive the general no-arbitrage condition for future convenience yield dynamics driven by the string shock. In Section 4 we obtain a closed-form solution for a call

option on commodity futures. In Section 5, parametrization of volatility and correlation functions is discussed. In Section 6, we discuss the crude oil implied volatility data, and calibrate the proposed model and the benchmark two-factor model of Schwartz (1997) (hereafter referred as 2FS-97 model). In Section 7, we obtain in-sample, and out-of-sample mean absolute pricing errors and relative root mean square errors with respect to fair call prices (based on Bloomberg volatility surface) and conduct nonparametric hypothesis tests for the individual parts (short, middle, long segment) and the entire forward curve. Section 8 concludes. Three appendices contain proofs.

2. String Shock

Sornette (1998) in the context of modeling forward interest rates proposed a stochastic string shock model where the forward rate curve is treated as a string where its length and transverse deformations are identified with the time-to-maturity x and the forward rate at a given time t . In this framework, the time increment of the forward rate curve is parameterized as $d_t f(t, x) = \alpha(t, x)dt + \sigma(t, x)d_t Z(t, x)$ where $Z(t, x)$ is a random field continuous in t and x , and both $\alpha(t, x)$ and $\sigma(t, x)$ are *a priori* arbitrary functions of $f(t, x)$. Under the no-arbitrage condition this leads to

$$d_t f(t, x) = (f_2(t, x) + A(t, x))dt + \sigma(t, x)d_t Z(t, x),$$

where $A(t, x) = \sigma(t, x) \left(\int_0^x \sigma(t, y) c_z(t, y, x) dy \right)$, $f_2(t, x) = \frac{\partial f(t, x)}{\partial x}$ and

$$c_z(t, y, y') = Cov[d_t Z(t, y), d_t Z(t, y')].$$

$Z(t, x)$ is taken to be the solution of a second order SPDE given as:

$$\begin{aligned} A_1(t, x)Z_{11}(t, x) + A_2(t, x)Z_{12}(t, x) + A_3(t, x)Z_{22}(t, x) + A_4(t, x)Z_1(t, x) \\ + A_5(t, x)Z_2(t, x) + A_6(t, x)Z(t, x) = \mathcal{N}(t, x), \end{aligned}$$

where $A_i(t, x)$ are functions, Z_i and Z_{ij} are the first and second order partial derivatives of Z respectively and \mathcal{N} is the random “source” term. In addition, $Z(t, x)$ must satisfy the following conditions in order to qualify as a string shock : (a) $Z(t, x)$ is continuous in t at all times x , and in x for all t , (b) time increments in string shocks has zero expected drift i.e. $E(d_t Z(t, x)) = 0$ for all x , (c) variance of time increments in string shocks is equal to the time elapsed i.e. $\text{var}(d_t Z(t, x)) = dt$ for all x , and (d) correlation of increments is not dependent on t i.e. $\text{corr}(d_t Z(t, x), d_t Z(t, y))$ is independent of t . The last three conditions ensure that string shocks are Markovian. Intuitively, $Z(t, x)$ can be thought of as a two dimensional analogue of the standard Wiener process $W(t)$ the difference being that $Z(t, x)$ depends both on t and x while $W(t)$ depends only on t . SS (2001) mentioned several kinds of string shocks capable of producing varied correlation structures across the forward rate curve parsimoniously. One can choose a string shock based on the desired correlation structure and the complexity one can manage. In this paper, we work with O-U sheet which meets the above conditions and is thus suitable for modeling the dynamics of future convenience yield.

O–U sheet (see Khoshnevisan, 2009, p. 9) is an extension of the O–U process (see Karlin and Taylor, 1981, p. 170). In our case, we consider Z to be a two parameter O–U random field that is constructed from the two parameter Brownian sheet W as:

$$Z(t, x) = e^{-\beta x} W(t, e^{2\beta x}).$$

In general, a n –parameter Brownian sheet W with parameter vector \mathbf{p} of n dimensions has the given covariance function,

$$\text{Cov}[W(\mathbf{p}_1), W(\mathbf{p}_2)] = \prod_{i=1}^n p_{1i} \wedge p_{2i},$$

where $\mathbf{p}_i = (p_{i1}, \dots, p_{iN})$. Hence, \mathbf{W} is a Gaussian process which generalizes standard Brownian motion to a two parameter random field. SS (2001) provides the following alternative representation of the O-U sheet:

$$Z(t, x) = e^{-\beta x} \int_{y=0}^{2\beta x} \int_{v=0}^t \eta(v, y) dv dy,$$

where η is the two dimensional white noise characterized by the covariance function,

$$Cov[\eta(t, x), \eta(s, y)] = \delta(s - t)\delta(y - x),$$

where δ is the Dirac Delta function. The correlation function of the string shock process Z is given as,

$$c(x, y) \equiv Cor[d_t Z(t, x), d_t Z(t, y)] = e^{-\beta|x-y|}.$$

It is important to note that correlation is a function of times of maturity only as t does not enter in it. Correlation is strong when times to maturity x and y are close to each other and decreases exponentially as the gap between them increases.

3. The Model

We made the following assumptions about the model: (1) Spot price, and future convenience yield follows diffusion process (2) Interest rates are non-random¹ (3) Spot price volatility, and future convenience yield volatility is deterministic (4) Correlation between the innovations of spot price, and future convenience yield is deterministic (5) Markets are arbitrage-free, and (6) Markets are frictionless, complete, and continuous trading is possible

¹For crude oil and copper, Schwartz (1997) found that parameters associated with spot price and spot convenience yield processes are robust to the specification of interest rate process. Miltersen (2003) states that the cost of complexity due to the use of stochastic interest rates in the model is much bigger than the gain associated with it.

with zero bid–ask spreads. We also assume a filtered probability space, $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_t, \mathbf{Q})$, where \mathbf{Q} is the risk neutral probability measure. We consider two progressively measurable stochastic processes- spot price, and future convenience yield. The dynamics of the spot price S , under \mathbf{Q} is taken as:

$$S(t) = S(0) + \int_{v=0}^t \mu_S(v) S(v) dv + \int_{v=0}^t \sigma_S(v) S(v) dW(v), \quad (1)$$

moreover, the dynamics of continuously compounded future convenience yield, ϵ , under \mathbf{Q} is:

$$\epsilon(t, x) = \epsilon(0, x) + \int_{v=0}^t \mu_\epsilon(v, x) dv + \int_{v=0}^t \sigma_\epsilon(v, x) d_v Z(v, x). \quad (2)$$

Let for each fixed x , $Z(t, x)$ be denoted as $Z^x(t)$. The cross variation between the stochastic processes W and Z^x , $\langle W, Z^x \rangle$, is assumed to satisfy:

$$d\langle W, Z^x \rangle_t = \rho(t, x) dt \text{ where } \rho(t, x) \in [-1, 1] \quad \forall x \geq 0 \quad \forall t \geq 0. \quad (3)$$

Further, cross variation of the string shocks satisfy (condition (d) for qualified strings):

$$d\langle Z^x, Z^y \rangle_t = c(x, y) dt \text{ where } c(x, y) \in [-1, 1] \quad \forall x, y \geq 0. \quad (4)$$

The diffusion terms, $\sigma_S(\cdot)$ and $\sigma_\epsilon(\cdot, \cdot)$, are deterministic real functions and satisfy regularity conditions so that SDEs have a strong solution. Let $B(t, s)$ be the time t price of a risk-free zero coupon bond with face value of one and maturity date s which is also the maturity date of futures the contract. Mathematically, $B(t, s) = e^{-\int_{v=t}^s r(v) dv}$, where r is non-stochastic short interest rate. The futures price is related to the futures convenience yield, and the price of the zero-coupon bond by the relation (MS, 1998):

$$F(t, s) = \frac{S(t)}{B(t, s)} e^{-\int_{y=0}^{s-t} \epsilon(t, y) dy}. \quad (5)$$

No-arbitrage conditions on the stated SDEs specifies the drift terms $\mu_S(\cdot)$ and $\mu_\epsilon(\cdot, \cdot)$. The drift of spot price process under \mathbf{Q} is:

$$\mu_S(t) = r(t) - \epsilon(t, 0), \quad (6)$$

where $\epsilon(t, 0)$ is spot convenience yield. Eq. (6) indicates that mean reversion in spot price is induced by the spot convenience yield. The more positive the value of $\rho(t, x)$, faster the mean reversion is (Miltersen, 2003). Let us introduce a process Y such that:

$$Y(t, s) = - \int_{y=0}^{s-t} \epsilon(t, y) dy. \quad (7)$$

Process Y can be represented by the SDE (see Appendix A for details):

$$\begin{aligned} Y(t, s) = Y(0, s) + \int_{v=0}^t \left(\epsilon(v, s - v) - \int_{y=0}^{s-v} \mu_{\epsilon}(v, y) dy \right) dv \\ - \int_{v=0}^t \int_{y=0}^{s-v} \sigma_{\epsilon}(v, y) dy d_v Z(v, y). \end{aligned} \quad (8)$$

Y is a continuous semimartingale (see Appendix B for proof).

Define $G(t, s) = e^{Y(t, s)}$. Since G is a \mathcal{C}^2 -function, applying Ito's lemma for continuous semimartingale (see Protter, 2004, p. 81) gives:

$$G(t, s) = G(0, s) + \int_0^t G(v, s) d_v Y(v, s) + \frac{1}{2} \int_0^t G(v, s) d\langle Y \rangle_v, \quad (9)$$

where $\langle Y \rangle_t = \int_{y=0}^{s-t} \sigma_{\epsilon}(t, y) \int_{w=0}^{s-t} \sigma_{\epsilon}(t, w) c(w, y) dw dy dt$, is the quadratic variation of Y . As

Y is a semimartingale, it forms the largest class of stochastic processes with respect to which Ito integral can be defined. Writing Eq. (8) and Eq. (9) in differential form and a little algebra yields:

$$\begin{aligned} \frac{d_t G(t, s)}{G(t, s)} = \left(\epsilon(t, s - t) - \int_{y=0}^{s-t} \mu_{\epsilon}(t, y) dy + \frac{1}{2} \int_{y=0}^{s-t} \sigma_{\epsilon}(t, y) \int_{w=0}^{s-t} \sigma_{\epsilon}(t, w) c(w, y) dx dy \right) dt \\ - \int_{y=0}^{s-t} \sigma_{\epsilon}(t, y) dy d_t Z(t, y). \end{aligned} \quad (10)$$

As $F(t, s) = \frac{S(t)}{B(t, s)} G(t, s)$, using Ito's product rule and Theorem 29 (see Protter, 2004, p. 75)

we get:

$$\begin{aligned} \frac{d_t F(t, s)}{F(t, s)} = & \left(\epsilon(t, s - t) - \int_{y=0}^{s-t} \mu_\epsilon(t, y) dy + \frac{1}{2} \int_{y=0}^{s-t} \sigma_\epsilon(t, y) \int_{w=0}^{s-t} \sigma_\epsilon(t, w) c(w, y) dw dy - \right. \\ & \left. \epsilon(t, 0) - \sigma_S(t) \int_{y=0}^{s-t} \sigma_\epsilon(t, y) \rho(t, x) dy \right) dt + \sigma_S(t) dW(t) - \\ & \int_{y=0}^{s-t} \sigma_\epsilon(t, y) dy d_t Z(t, y). \end{aligned} \quad (11)$$

As the futures price process F is a continuous local martingale with zero drift² under \mathbf{Q} -measure, we have:

$$\begin{aligned} \epsilon(t, s - t) - \int_{y=0}^{s-t} \mu_\epsilon(t, y) dy + \frac{1}{2} \int_{y=0}^{s-t} \sigma_\epsilon(t, y) \int_{w=0}^{s-t} \sigma_\epsilon(t, w) c(w, y) dw dy - \\ \epsilon(t, 0) - \sigma_S(t) \int_{y=0}^{s-t} \sigma_\epsilon(t, y) \rho(t, x) dy = 0. \end{aligned} \quad (12)$$

Differentiating Eq. (12) with respect to $x = s - t$ gives:

$$\mu_\epsilon(t, x) = \frac{\partial \epsilon(t, x)}{\partial x} + \sigma_\epsilon(t, x) \left(\int_{y=0}^x \sigma_\epsilon(t, y) c(x, y) dy - \sigma_S(t) \rho(t, x) \right). \quad (13)$$

Eq. (13) gives the drift of the future convenience yield process under no-arbitrage condition.

The dynamics of the futures convenience yield under no-arbitrage is given as:

$$\begin{aligned} d_t \epsilon(t, x) = \\ \left[\frac{\partial \epsilon(t, x)}{\partial x} + \sigma_\epsilon(t, x) \left(\int_{y=0}^x \sigma_\epsilon(t, y) c(x, y) dy - \rho(t, x) \sigma_S(t) \right) \right] dt + \sigma_\epsilon(t, x) d_t Z(t, x). \end{aligned} \quad (14)$$

Under \mathbf{Q} -measure, dynamics of futures price process reduces to:

$$\frac{d_t F(t, s)}{F(t, s)} = \sigma_S(t) dW(t) - \int_{y=0}^{s-t} \sigma_\epsilon(t, y) dy d_t Z(t, y). \quad (15)$$

²Futures price process may have non-zero drift under real world probability measure, \mathbf{P} , owing to the presence of risk premium. This has been confirmed by several authors- Fama and French (1987), Bessembinder (1992), Moosa and Al-Loughani (1994), Sadorsky (2002), Arouri et al. (2013), to name a few.

A key implication of Eq. (15) is that not all futures prices move in the same direction which is typical of energy commodities. The movement in a forward curve over the time can be attributed to two uncertainties³- *level shock* and *higher order shocks*. The former is due to spot price fluctuation and is responsible for inducing a parallel shift while the latter causes higher order distortions like steeping or flattening of the slope, and variations in curvature. *Higher order shocks* are the aggregation of string shock modulated by the future convenience yield volatility function taking different values at different times to maturity y . As maturity date, s , increases, and if second term starts dominating the first term then one observes the disconnect between short term futures contracts (or commodity spot price) and long term futures contracts.

The quadratic variation of futures price process is (see Appendix C for proof):

$$\begin{aligned} \langle F \rangle_t = & \int_{v=0}^t \sigma_S(v)^2 F(v, s)^2 dv - 2 \int_{v=0}^t \sigma_S(v) F(v, s)^2 \int_{y=0}^{s-v} \sigma_\epsilon(v, y) \rho(v, y) dy dv \\ & + \int_{v=0}^t F(v, s)^2 \int_{y=0}^{s-v} \sigma_\epsilon(v, y) \int_{w=0}^{s-v} \sigma_\epsilon(v, w) c(w, y) dw dy dv. \end{aligned} \quad (16)$$

By using Ito's lemma (see Protter, 2004, p. 81) on $\ln F(t, s)$ we get:

$$\begin{aligned} F(t, s) = & F(0, s) \exp \left[\int_{v=0}^t \sigma_S(v) dW(v) - \int_{v=0}^t \int_{y=0}^{s-v} \sigma_\epsilon(v, y) d_v Z(v, y) dy \right. \\ & - \frac{1}{2} \int_{v=0}^t \left(\sigma_S(v)^2 + \int_{y=0}^{s-v} \sigma_\epsilon(v, y) \int_{w=0}^{s-v} \sigma_\epsilon(v, w) c(w, y) dw dy \right. \\ & \left. \left. - 2\sigma_S(v) \int_{y=0}^{s-v} \sigma_\epsilon(v, y) \rho(v, y) dy \right) dv \right]. \end{aligned} \quad (17)$$

³Szymanowska et al. (2014) identify two risk premia responsible for expected futures returns under the real world probability measure, i.e., spot premium and term premia. Former is due to the spot price risk (or level shocks) while latter exists due to convenience yield risks (or higher order shocks).

4. Pricing European Future Options

In this section, we develop a formula for European call option with exercise price K , expiry date T written on a commodity futures contract with expiry date s ($s > T$) on date zero. Let Ψ be a stochastic process given as:

$$\Psi_T = \int_{v=0}^T \sigma_S(v) dW(v) - \int_{v=0}^T \int_{y=0}^{s-v} \sigma_\epsilon(v, y) dy d_v Z(v, y). \quad (18)$$

Note that Ψ is a martingale with $\Psi_0 = 0$. Further, variance of Ψ_T is:

$$\begin{aligned} \sigma_{\Psi_T}^2 = \mathbb{E}[\Psi_T^2] = & \int_{v=0}^T \left(\sigma_S(v)^2 - 2\sigma_S(v) \int_{y=0}^{s-v} \sigma_\epsilon(v, y) \rho(v, y) dy \right. \\ & \left. + \int_{y=0}^{s-v} \sigma_\epsilon(v, y) \int_{w=0}^{s-v} \sigma_\epsilon(v, w) c(y, w) dw dy \right) dv. \end{aligned} \quad (19)$$

σ_{Ψ_T} is the T -day volatility of the instantaneous return on underlying futures contract. The above result is obtained after straight forward calculations using the facts that (a) if $M(t)$ is a continuous, square integrable martingale such that $M(0) = 0$, then, $M^2 - \langle M \rangle$ is also a martingale and thus $\mathbb{E}[M^2] = \mathbb{E}[\langle M \rangle]$ and (b) if $N(t)$ is another continuous, square integrable martingale such that $N(0) = 0$, then, $MN - \langle M, N \rangle$ is also a martingale (see Grigoriu, 2002, p. 180). One can obtain call option price as:

$$C(0, T) = B(0, T)[F(0, s)N(d_1) - KN(d_2)],$$

$$\text{where } d_1 = \frac{\ln \frac{F(0, s)}{K} + \frac{1}{2} \sigma_{\Psi_T}^2}{\sigma_{\Psi_T}} \text{ and } d_2 = \frac{\ln \frac{F(0, s)}{K} - \frac{1}{2} \sigma_{\Psi_T}^2}{\sigma_{\Psi_T}}. \quad (20)$$

To see the result given in Eq. (20), we note that by using Eq. (18) one can write Eq. (17) concisely as:

$$F(T, s) = F(0, s) \exp\left(\Psi_T - \frac{1}{2} \sigma_{\Psi_T}^2\right), T \leq s. \quad (21)$$

Using Eq. (21) in the expression, $B(0, T) \mathbb{E}[(F(T, s) - K) 1_{F(T, s) > K}]$, where $1_{F(T, s) > K}$ is an indicator function that equals to 1 if condition given in subscript is true and zero otherwise, will yield,

$$\begin{aligned} C(0, T) &= B(0, T)F(0, s)\mathbb{E}\left[\exp\left\{\Psi_T - \frac{1}{2}\sigma_{\Psi_T}^2\right\} 1_{F(0, s) \exp\{\Psi_T - \frac{1}{2}\sigma_{\Psi_T}^2\} > K}\right] \\ &\quad - B(0, T)K \mathbb{E}\left[1_{F(0, s) \exp\{\Psi_T - \frac{1}{2}\sigma_{\Psi_T}^2\} > K}\right] \\ &= C_1(0, T) - C_2(0, T) \quad (\text{say}) \end{aligned}$$

Writing $\Psi_T = \sigma_{\Psi_T}\psi$, where $\psi \sim N(0, 1)$, we get:

$$\begin{aligned} C_1(0, T) &= B(0, T)F(0, s)e^{-\frac{1}{2}\sigma_{\Psi_T}^2}\mathbb{E}\left[e^{\Psi_T} 1_{\psi > \frac{\ln\frac{K}{F(0, s)} + \frac{1}{2}\sigma_{\Psi_T}^2}{\sigma_{\Psi_T}}}\right] \\ &= B(0, T)F(0, s)N\left(\frac{\ln\frac{F(0, s)}{K} + \frac{1}{2}\sigma_{\Psi_T}^2}{\sigma_{\Psi_T}}\right) \end{aligned}$$

and

$$C_2(0, T) = B(0, T)KN\left(\frac{\ln\frac{F(0, s)}{K} - \frac{1}{2}\sigma_{\Psi_T}^2}{\sigma_{\Psi_T}}\right)$$

The value of $C(0, T)$ is thus obtained.

5. Parametrization

In this section, we will parameterize volatility and correlation functions. We observed that the term structure of implied volatility of WTI crude oil futures is backwarded during the observation period of Oct 30, 2014 to Jul 29, 2016. This implies future convenience yield for longer dated contract is less volatile than its shorter dated counterpart, assuming deterministic interest rates. This requirement is satisfied by using the time-homogeneous specification of

Miltersen (2003, p. 54-55) where spot price and future convenience yield volatilities are given as:

$$\sigma_S(t) = \hat{\sigma}_S g_S(t), \text{ where } g_S(t) = 1 + A_S \sin 2\pi(t + B_S) \quad (22)$$

$$\sigma_\epsilon(t, x) = \hat{\sigma}_\epsilon g_\epsilon(t) e^{-\gamma x}, \text{ where } g_\epsilon(t) = 1 + A_\epsilon \sin 2\pi(t + B_\epsilon) \quad (23)$$

The sinusoidal functions g_S and g_ϵ are used to model seasonality in the instantaneous volatility of spot price process and future convenience yield process respectively. Since no significant seasonal patterns⁴ are documented for crude oil, we assume $A_S = A_\epsilon = 0$. The correlation among future convenience yields corresponding to different maturities depends on the type of string shock used for modeling. We will operationalize string shock model by using the O-U sheet as the noise source for future convenience yield process. Correlation function when O-U sheet act as a noise source over future convenience yields is given as,

$$c(x, y) = e^{-\beta|x-y|}, \text{ where } \beta \geq 0. \quad (24)$$

Under \mathbf{Q} -measure, the drift of spot price process is solely determined by standard no-arbitrage argument. A possible way in which mean reversion can be induced in spot price process is through the positive correlation between spot price and spot convenience yield (Miltersen, 2003). The correlation function of time increments of spot price and future convenience yield processes can be parameterized as:

$$\rho(t, x) = \hat{\rho} e^{-\alpha x}, \text{ where } \alpha \geq 0 \text{ and } \hat{\rho} \in (0,1). \quad (25)$$

Note that $\rho(t, 0)$ is the correlation between spot price and spot convenience yield.

⁴Unlike agricultural and other energy commodities, crude oil show insignificant seasonality. Borovkova and Geman (2006) posit that seasonal premium in crude oil futures is negligible for all calendar months.

Parametrization in Eq. (25) takes care of less correlated movements in the long dated futures and the spot price. This parameterization is consistent with the framework of Routledge et. al. (2000) according to which correlation between the spot price and convenience yield falls to zero as the horizon lengthens. We take $\rho(t, 0) = \hat{\rho}$ as a constant for calibration purpose. Now Eq. (19) can be further solved for T -day volatility of the instantaneous return on futures contract, given as:

$$\sigma_{\Psi_T} = \left[\left(\hat{\sigma}_S^2 + \hat{\sigma}_\epsilon^2 \frac{\beta - 2\gamma}{\gamma(\beta^2 - \gamma^2)} \right) T - 2\hat{\sigma}_S \hat{\sigma}_\epsilon \hat{\rho} \left(\frac{T}{\alpha + \gamma} - \frac{e^{-(\alpha+\gamma)(s-T)} - e^{-(\alpha+\gamma)s}}{(\alpha + \gamma)^2} \right) - \hat{\sigma}_\epsilon^2 \frac{\beta e^{-2\gamma s}}{\gamma(\beta^2 - \gamma^2)} \left(\frac{e^{2\gamma T} - 1}{2\gamma} \right) + \hat{\sigma}_\epsilon^2 \frac{e^{-(\beta+\gamma)s}}{\beta^2 - \gamma^2} \left(\frac{e^{(\beta+\gamma)T} - 1}{\beta + \gamma} \right) + \hat{\sigma}_\epsilon^2 \frac{e^{-(\beta-\gamma)s}}{\beta^2 - \gamma^2} \left(\frac{e^{(\beta-\gamma)T} - 1}{\beta - \gamma} \right) \right]^{1/2}. \quad (26)$$

We will use Eq. (26) for calibration in the next section.

6. Data and Calibration

The contracts traded on NYMEX are one of the world's most liquid commodity futures options. As these options are of American style, one cannot use the market prices directly to calibrate the model⁵. Moreover, options at certain strikes and maturities are not liquid. To circumvent these difficulties we use implied volatility surface⁶ obtained from *Bloomberg*.

This volatility surface has many desirable features such as it is smooth, close to the quoted market data, admits fast interpolation, and has low market microstructure noise. Along with

⁵Other possible way is to invert an analytic approximation of American option to recover implied volatility. For crude oil, Trolle and Schwartz (2009) used Barone-Adesi and Whaley (1987) approximation to recover lognormal implied volatility and then priced European option by using the Black (1976) formula.

⁶Bloomberg follows data-driven methodology based on smoothing splines for the construction of implied volatility surface. Options are de-Americanized to infer the correct implied volatility. Only information from liquid options is considered. For illiquid strikes and maturities, other sources of information are used, like settlement prices of the previous day.

the futures prices, data for implied volatility and strike price corresponding to 50 delta at the money (ATM) region for various liquid maturities was obtained from Oct 30, 2014 to Jul 29, 2016 (440 trading days). By using Black-76 formula, we get the European call option prices. For future use, we refer to these prices as *fair values*. We calibrate string shock model and the 2FS-97 on 10th, 20th, and 30th of every month. In the case of a market holiday or when the month has less than 30 days, we calibrate using the data of the next trading day. Thus, for each month we have three non-overlapping windows of 10-11 calendar days. In this study, we have 63 such non-overlapping windows. Every window has one in-sample day (Day 0) and up to seven out-of-the-sample days (Days 1-7) depending on the number of trading holidays in a window. Depending on the option liquidity around 50 delta ATM region, on Day-0 a window can carry the information of 18 to 33 hypothetical options of maturities ranging from 22 days to 6.5 years.

6.1. String Shock model calibration

We estimate six parameters $\theta \equiv (\hat{\sigma}_S, \hat{\sigma}_\epsilon, \alpha, \beta, \gamma, \hat{\rho})$ for string shock model by minimizing the squared relative pricing error loss function,

$$L_\theta = \frac{1}{N} \sum_{i=1}^N \left(\frac{C_i(\theta) - C_i}{C_i} \right)^2.$$

Subject to the constraints: $\hat{\sigma}_S, \hat{\sigma}_\epsilon, \alpha, \beta, \gamma > 0$, and $0 < \hat{\rho} < 1$, where N is the number of option contracts of different maturities in a window, C_i is the fair price of the call option with the i^{th} maturity and $C_i(\theta)$ is the price given by string shock model for the same option. To minimize the above loss function we have used Differential Evolution⁷ (DE) algorithm (Storn

⁷DE is a heuristic approach that optimizes a problem by iteratively trying to improve a candidate solution. It can work with a loss function that is not differentiable, nonlinear, multidimensional, or have many local minima and constraints.

and Price, 1997) as implemented in *DEoptim* package in R software (Mullen et al. 2011). Figure 1 illustrates the fitting of term structure of volatility for string shock model on May 02, 2016, which is also a typical fit for other calibration days. Without loss of generality, we have considered zero interest rates, i.e., $B(0, T) = 1$ for the rest of this section.

6.2. 2FS-97 model calibration

MS (1998) showed how to value options under the three factor model of Schwartz (1997) with suitable parametrization (see pg. 44-46). One can obtain option price for 2FS-97 model by making the volatility of interest rates zero, i.e., $\sigma_f = 0$. As we have assumed deterministic interest rates, option prices obtained under 2FS-97 model will be the right benchmark for the evaluation of string shock model. The call option formula on futures prices stated in Eq. (16) of MS (1998) for 2FS-97 model using our notation is:

$$C_{MS} = B(0, t) \left(F(0, s) N \left(\frac{\ln \frac{F(0, s)}{K} + \frac{\sigma^2}{2}}{\sigma} \right) - KN \left(\frac{\ln \frac{F(0, s)}{K} - \frac{\sigma^2}{2}}{\sigma} \right) \right), \quad (27)$$

$$\text{where } \sigma^2 = \int_{v=0}^t \left\| \tilde{\sigma}_S(v) - \int_{z=v}^s \tilde{\sigma}_\epsilon(v, z) dz \right\|^2 dv. \quad (28)$$

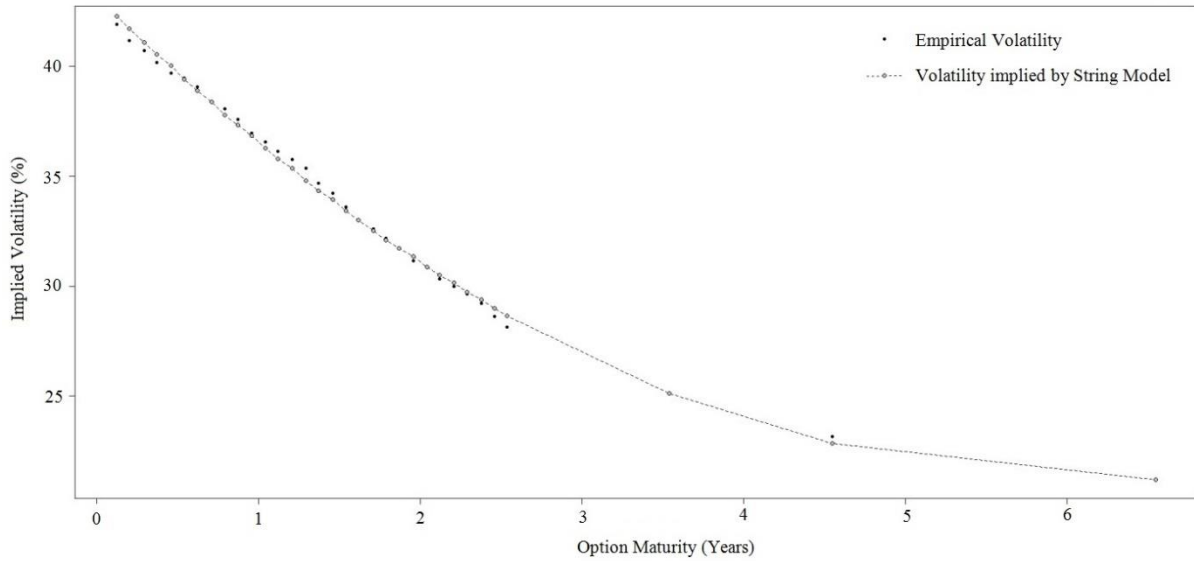


Figure 1: Fitting of the volatilities implied by string shock model to the empirical Black-76 volatility term structure on May 02, 2016

We calibrate the 2FS-97 model on Day 0 of 63 observation windows similarly as we did it for our proposed model. The four estimated 2FS-97 model parameters are obtained by minimizing $L_{\vec{\theta}}$ where $\vec{\theta} \equiv (\check{\sigma}_S, \check{\sigma}_\epsilon, k_\epsilon, \rho_{S\epsilon})$ subject to the constraints:

$$\check{\sigma}_S, \check{\sigma}_\epsilon, k_\epsilon > 0 \text{ and } 0 < \rho_{S\epsilon} < 1.$$

Like before the minimization is carried out using the *DEoptim* package in R software. A description of estimated parameters for both the models is given in Table 1(a) and (b).

Parameter Estimates	σ_S	σ_ϵ	α	β	γ	ρ
Median (Q2)	0.434	0.509	0.419	0.446	0.219	0.641
IQR (Q3-Q1)	0.127	0.430	0.521	1.028	0.374	0.166

Table 1(a): Summary statistics of estimated String shock model parameters (Number of observations: 63)

Parameter Estimates	$\check{\sigma}_S$	$\check{\sigma}_\epsilon$	k_ϵ	$\rho_{S\epsilon}$
Median (Q2)	0.434	0.352	0.665	0.905
IQR (Q3-Q1)	0.127	0.260	0.580	0.140

Table 1(b): Summary statistics of estimated 2FS-97 model parameters (Number of observations: 63)

7. Results

We estimated 63 sets of parameters for our string shock model and 2FS-97 model and calculated the in-sample and out-of-sample call prices for 63 observation windows. Based on the fair values of call options and prices given by the two models we obtain the mean absolute pricing error for both the models. Summary statistics presented in Table 2, where we report in-sample and out-of-sample mean absolute pricing error averaged over the entire term structure for both the models, indicates that the string shock model is found to be more accurate than the 2FS-97 model over the study period. The plot of the errors from the two models over the observation windows (see Figure2) also suggests a superior performance of string shock model. In general, the margin of excess mean absolute price error fades as we move further away from calibration date (Day-0) which suggests a need for frequent calibration to back out the model parameters.

	Number of Observations	Mean Absolute Pricing Error (\$)									
		Minimum		1 st Quartile		Median		3 rd Quartile		Maximum	
		String	2FS-97	String	2FS-97	String	2FS-97	String	2FS-97	String	2FS-97
Day-0	63	0.0237	0.0255	0.0366	0.0505	0.0484	0.0800	0.0785	0.1035	0.1919	0.3339
Day-1	63	0.0367	0.0398	0.0750	0.0847	0.1182	0.1319	0.1752	0.1829	0.3306	0.4200
Day-2	63	0.0428	0.0439	0.0949	0.1090	0.1323	0.1584	0.2080	0.2204	0.6635	0.6625
Day-3	63	0.0439	0.0440	0.0982	0.1284	0.1698	0.2011	0.2761	0.2966	0.9762	0.9937
Day-4	63	0.0403	0.0453	0.1400	0.1479	0.2241	0.2631	0.3504	0.4271	1.2718	1.2909
Day-5	60	0.0553	0.0465	0.1616	0.1613	0.2548	0.2765	0.4237	0.4509	1.2982	1.3185
Day-6	44	0.0545	0.0564	0.1323	0.1318	0.2546	0.2676	0.4866	0.4998	1.2849	1.2850
Day-7	21	0.0505	0.0788	0.2171	0.2114	0.4377	0.4381	0.6162	0.6147	1.4149	1.4150

Table 2: Summary statistics of in-sample and out-of-sample mean absolute option pricing errors

The degree of accuracy is not very clear when option pricing error in absolute term is considered. For example, a pricing error of \$0.05 on an option with a fair price of \$2 is more significant than an error of \$0.10 if the fair price of the option is \$5. Hence, we use relative root mean square error (rRMSE) given as:

$$rRMSE_d = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(\frac{C_{i,d}(\theta) - C_{i,d}}{C_{i,d}} \right)^2},$$

where $d=0$ indicates in sample day (Day 0), and $d \in \{1, 2, \dots, 7\}$ indicates out-of-sample day. To scrutinize which segment of term structure of call option is priced more (or less) accurately we divide it into three parts, i.e., options having maturity of less than 6 months (short segment), options with maturity in between 6-18 months (middle segment), and options with maturity more than 18 months (long segment). We obtain in-sample and out-of-sample RMSEs for these three segments and for the entire term structure for both the models. Summary statistics of Tables 3(a)-(c) suggest that string shock model does less mispricing than the 2FS-97 model for all segments up to Day-2. For both the models, mispricing increases with the maturity of options. However, for entire term error dispersion for string shock model is lesser than 2FS-97 model for all the days (see Table 3(d)).

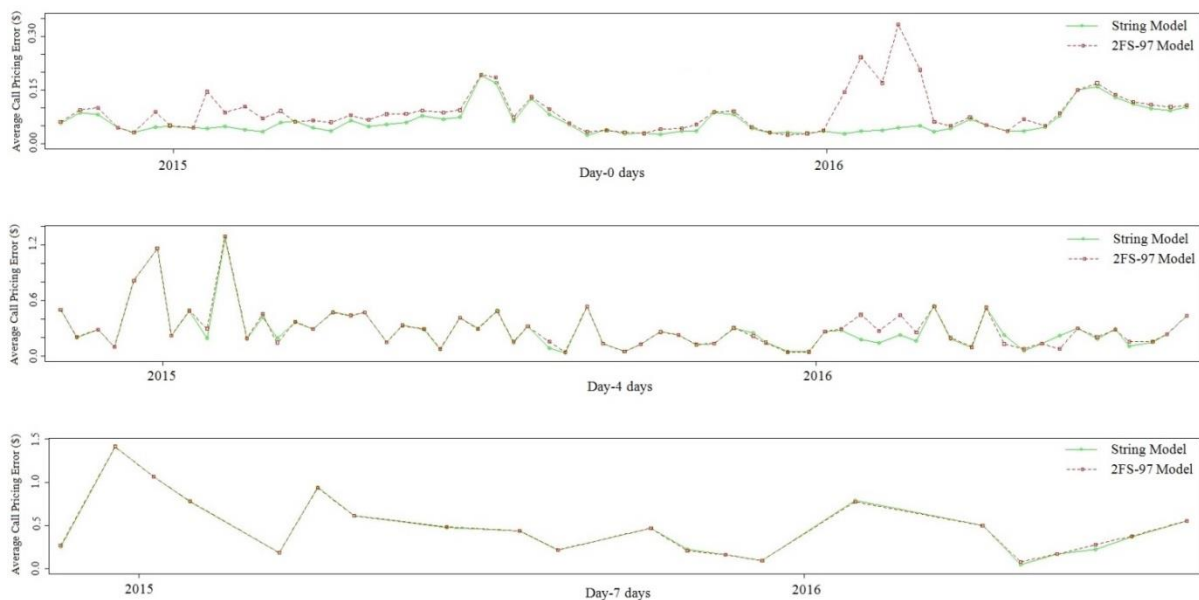


Figure 2: Topmost: Mean absolute in sample pricing errors (Oct 30, 2014 to Jul 20, 2016; 63 observations); Middle: Mean absolute out-of-sample pricing errors on Day-4 (Nov 05, 2014 to Jul 26, 2016; 63 observations); Bottom: Mean absolute out-of-sample pricing errors on Day-7 (Nov 19, 2014 to Jul 29, 2016; 21 observations)

Sample	Number of Observations	Relative Root Mean Square Pricing Error (%) - Short Segment									
		Minimum		1 st Quartile		Median		3 rd Quartile		Maximum	
		String	2FS-97	String	2FS-97	String	2FS-97	String	2FS-97	String	2FS-97
Day-0	63	0.44	0.49	0.89	1.06	1.25	2.30	2.61	3.60	7.15	17.80
Day-1	63	0.37	0.59	2.19	2.46	3.16	3.74	4.21	5.76	9.03	23.78
Day-2	63	0.73	0.80	2.34	2.81	3.81	4.24	5.74	7.13	13.54	24.51
Day-3	63	0.50	0.38	2.85	3.72	4.89	5.56	7.98	8.96	18.90	20.16
Day-4	63	0.42	0.60	3.13	3.99	6.07	6.53	10.13	10.55	19.89	21.91
Day-5	60	1.13	1.29	3.69	3.93	6.76	6.29	9.53	10.94	20.37	20.32
Day-6	44	1.81	2.21	4.13	4.30	6.29	5.83	9.56	9.70	25.64	25.59
Day-7	21	1.43	1.43	4.89	5.04	6.88	6.26	10.22	10.17	28.24	28.20

Table 3(a): Summary Statistics of in sample (Day-0) and out-of-sample (Days 1-7) pricing errors for short term options (up to 6 months)

Sample	Number of Observations	Relative Root Mean Square Pricing Error (%) - Middle Segment									
		Minimum		1 st Quartile		Median		3 rd Quartile		Maximum	
		String	2FS-97	String	2FS-97	String	2FS-97	String	2FS-97	String	2FS-97
Day-0	63	0.45	0.44	0.73	0.89	0.94	1.24	1.69	2.12	3.91	5.33
Day-1	63	0.56	0.55	1.33	1.60	2.13	2.31	3.29	3.51	6.64	9.78
Day-2	63	0.65	0.66	1.79	1.98	2.66	2.70	4.25	4.35	11.11	11.27
Day-3	63	0.54	0.55	1.65	1.97	3.10	3.28	5.40	5.45	14.50	13.64
Day-4	63	0.55	0.56	2.31	2.23	3.64	3.95	7.40	7.40	18.23	17.40
Day-5	60	0.40	0.40	2.67	2.73	5.22	5.01	8.49	8.18	18.07	17.23
Day-6	44	0.67	0.99	2.06	2.34	4.99	4.98	9.10	9.13	21.22	21.26
Day-7	21	0.67	1.30	4.28	4.26	8.27	8.14	13.06	13.13	22.38	22.41

Table 3(b): Summary statistics of in sample (Day-0) and out-of-sample (Days 1-7) pricing errors for medium term options (6-18 months)

Sample	Number of Observations	Relative Root Mean Square Pricing Error (%) - Long Segment									
		Minimum		1 st Quartile		Median		3 rd Quartile		Maximum	
		String	2FS-97	String	2FS-97	String	2FS-97	String	2FS-97	String	2FS-97
Day-0	63	0.38	0.38	0.83	1.11	1.15	1.54	1.61	2.12	3.46	7.50
Day-1	63	0.34	0.43	1.44	1.74	2.05	2.32	3.35	3.57	7.91	7.92
Day-2	63	0.64	0.67	1.58	2.01	2.67	2.63	3.66	3.91	12.25	11.35
Day-3	63	0.76	0.77	1.97	2.32	3.49	3.87	4.60	4.83	16.30	15.07
Day-4	63	0.78	0.77	2.52	2.63	4.44	4.49	5.68	6.35	21.37	20.18
Day-5	60	0.67	0.30	3.01	3.04	4.76	4.93	7.56	7.66	20.64	20.61
Day-6	44	0.97	1.04	2.80	2.82	4.58	4.96	8.33	8.15	20.67	19.16
Day-7	21	0.94	1.32	3.02	3.00	7.35	7.46	11.57	11.90	18.81	22.80

Table 3(c): Summary statistics of in sample (Day-0) and out-of-sample (Days 1-7) pricing errors for long term options (above 18 months)

Sample	Number of Observations	Relative Root Mean Square Pricing Error (%) - Entire Term									
		Minimum		1 st Quartile		Median		3 rd Quartile		Maximum	
		String	2FS-97	String	2FS-97	String	2FS-97	String	2FS-97	String	2FS-97
Day-0	63	0.60	0.69	0.88	1.13	1.20	1.78	1.89	2.44	4.82	9.67
Day-1	63	1.08	1.15	1.93	2.12	2.67	2.86	3.56	3.94	6.72	12.75
Day-2	63	1.06	1.08	2.25	2.46	3.29	3.57	4.32	4.92	11.84	13.93
Day-3	63	1.05	1.30	2.46	3.13	4.12	4.39	5.92	6.80	14.65	15.52
Day-4	63	0.95	1.07	3.10	3.29	5.21	5.79	7.72	7.94	18.12	18.86
Day-5	60	1.37	1.47	3.98	4.11	5.49	6.02	8.17	8.47	18.20	18.75
Day-6	44	1.57	1.98	3.72	3.79	5.31	5.66	9.73	9.78	21.35	21.35
Day-7	21	1.62	3.11	4.46	4.27	8.16	8.18	12.28	12.34	22.65	22.65

Table 3(d): Summary statistics of in sample (Day-0) and out-of-sample (Days 1-7) pricing errors for the entire term structure of 50 delta call prices

To gather statistical evidence that on average, pricing error under string shock model is less than 2FS-97 model a hypothesis test is conducted. As neither rRMSE samples nor their difference exhibit normality, we resort to nonparametric Wilcoxon signed rank test. Null and alternate hypothesis are given as:

$$H_0: \text{mean of } rRMSE_{\text{string}} \geq \text{mean of } rRMSE_{2FS-97}$$

$$H_A: \text{mean of } rRMSE_{\text{string}} < \text{mean of } rRMSE_{2FS-97}$$

There is strong indication that that string shock model outperforms the 2FS-97 model (see Table 4). Except for “Day-5” in middle segment and “Day-7” in short segment, one can reject the null hypothesis at 10% level of significance in all the other cases. For the long segment and the entire segment, the null hypothesis can be rejected at 5% level of significance.

P-value	Number of Observations	Short Segment	Middle Segment	Long Segment	Entire Segment
Day-0	63	0.00002	<10 ⁻⁵	<10 ⁻⁵	<10 ⁻⁵
Day-1	63	0.0029	0.0005	0.0004	<10 ⁻⁵
Day-2	63	0.0049	0.0012	0.0006	<10 ⁻⁵
Day-3	63	0.0777*	0.0009	0.0002	<10 ⁻⁵
Day-4	63	0.0002	0.0025	0.0005	<10 ⁻⁵
Day-5	60	0.0111	0.1684**	0.0021	<10 ⁻⁵
Day-6	44	0.0118	0.0629*	0.0020	<10 ⁻⁵
Day-7	21	0.3166**	0.0555*	0.0274	0.0351

Table 4: P-values of 32 Wilcoxon signed rank tests. Values without stars indicate rejection of null hypothesis at 5% level of significance.

*Null hypothesis can be rejected in favor of alternative hypothesis at 10% significance level

**Failed to reject the Null hypothesis in favor of alternative hypothesis even at 10% significance level

In general, option formula for two and three factor Schwartz (1997) model is a special case of MS (1998) in which authors use the same vector Wiener process to model the dynamics of the spot price, future convenience yield, and forward rates. To obtain a realistic correlation between these processes, it becomes necessary to make volatility a function of correlation coefficients (see MS, 1998, p. 44, Eq. (27) and (28)). Calibration of such a model may yield unrealistic⁸ parameter values, i.e., optimum solution may occur outside permitted parameter space.

8. Conclusion

In this article, we present a new approach to price options on commodity futures when a string shock perturbs the term structure of future convenience yield. Fluctuation at each point on the term structure is governed by a correlation structure which depends on the string shock used for modeling. We decomposed the forward curve dynamics into a level shift and higher order distortions. Former is due to the spot price movement while latter is the net effect of perturbation due to future convenience yields. In the back end of the forward curve, higher order distortions dominate the level shift. This is empirically consistent with the observed low correlation between commodity spot and long-term futures contracts. Using string shock instead of (multidimensional) Wiener process allows much-needed separation of volatility and correlation functions. This gives a modeler freedom to choose appropriate volatility and correlation functions. Model calibration is relatively easy and yields realistic values for the parameters. This aspect is missing in the earlier term structure models.

⁸To keep pricing error as low as possible, we accept the model parameters corresponding to $\rho_{S\epsilon} = 0.99$ during Jan 11- Mar 01, 2016 (six Day-0 days). Due to excessive crude oil inventory build-up around this time (Preciado, 2016), we feel such a high value is implausible during contango. Routledge et. al. (2000) posits correlation between spot price and spot convenience yield depends on inventory level, and is higher in backwardation than in contango.

We have obtained the no-arbitrage drift for the future convenience yield under the risk neutral measure. This is useful for pricing complex derivatives by simulation. We derived an explicit closed-form solution for a European style commodity futures option for a general qualified string stock. Finally, using O-U sheet as a noise source, we obtained option-based estimates for the proposed model parameters and showed that it leads to a lower pricing error when compared to benchmark 2FS-97 model. As calibration process is similar for both the models, the new proposed model is better specified than the 2FS-97 model.

References

1. Arouri, M.E.H., Hammoudeh, S., Lahiani, A., & Nguyen, D.K. (2013). On the short and long-run efficiency of energy and precious metal markets. *Energy Economics*, 40, 832–844.
2. Barone-Adesi, G., & Whaley, R. (1987). Efficient analytic approximation of American option values. *Journal of Finance*, 42, 301–320.
3. Bessembinder, H. (1992). Systematic risk, hedging pressure, and risk premiums in futures markets. *Review of Financial Studies*, 5, 637–667.
4. Bessembinder, H., Coughenour, J. F., Seguin, P. J., & Smoller, M. M. (1995). Mean reversion in equilibrium asset prices: Evidence from the futures term structure. *Journal of Finance*, 50, 361–375.
5. Björk, T., & Landen, C. (2002). On the term structure of forward and futures prices, In: H. Geman, D. Madan, S. R. Pliska, & T. Vorst, (Eds.), *Mathematical Finance-Bachelier Congress 2000*. Springer Verlag.
6. Black, F. (1976). The pricing of commodity contracts. *Journal of Financial Economics*, 3, 167–179.
7. Borovkova, S., Geman, H. (2006). Seasonal and stochastic effects in commodity forward curves. *Review of Derivative Research*, 9, 167–186.
8. Brace, A., Gatarek, D., & Musiela, M. (1997). The Market model of interest rate dynamics. *Mathematical Finance*, 7, 127-155.
9. Bueno-Guerrero, A., Moreno, M., & Navas, J. F. (2015). Stochastic string models with continuous semimartingales. *Physica A*, 433, 229–246.
10. Carmona, R., Ludkovski, M., 2005. Spot Convenience Yield Models for Energy Markets, unpublished manuscript, Princeton University.
11. Casassus, J., & Collin-Dufresne, P. (2005). Stochastic convenience yield implied from commodity futures and interest rates. *Journal of Finance*, 60, 2283–2331.
12. Clewlow, L., & Strickland C. (1999a). A Multi-Factor Model for Energy Derivatives. Working Paper, University of Technology, Sydney.
13. Clewlow, L., & Strickland, C. (1999b). Valuing energy options in a one factor model fitted to forward prices. Working paper, University of Technology, Sydney.
14. Cortazar, G., & Schwartz, E. S. (1994). The evaluation of commodity contingent claims. *Journal of Derivatives*, 1, 27–39.

15. Cortazar, G., & Schwartz, E.S. (2003). Implementing a stochastic model for oil futures prices. *Energy Economics*, 25, 215–238.
16. Crosby, J. (2008). A multi-factor jump-diffusion model for commodities. *Quantitative Finance*, 8, 181–200.
17. Fama, E.F., & French, K.R. (1987). Commodity futures prices: some evidence on forecast power, premiums, and the theory of storage. *Journal of Business*, 60, 55–73.
18. Gabillon, J. (1991). The Term Structure of Oil Futures Prices. Working paper, Oxford Institute for Energy Studies.
19. Gibson, R., & Schwartz, E. S. (1990). Stochastic convenience yield and the pricing of oil contingent claims. *Journal of Finance*, 45, 959–976.
20. Goldstein, R. S. (2000). The term structure of interest rates as a random field. *Review of Financial Studies*, 13, 365–384.
21. Grigoriu M. (2002). *Stochastic Calculus—Applications in Science and Engineering*, first ed. Springer, New York.
22. Han, B. (2007). Stochastic volatilities and correlations of bond yields. *Journal of Finance*, 62, 1491–1524.
23. Heath, D., Jarrow, R., & Morton, A. (1992). Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. *Econometrica*, 60, 77–105.
24. Jamshidian, F. (1997). LIBOR and swap market models and measures. *Finance and Stochastics*, 1, 293-330.
25. Karlin, S., & Taylor, H. (1981). *A Second Course in Stochastic Processes*, Academic Press, New York.
26. Kennedy, D. P. (1994). The Term Structure of Interest Rates as a Gaussian Random Field. *Mathematical Finance*, 4, 247–258.
27. Kennedy, D. P. (1997). Characterizing Gaussian Models of the Term Structure of Interest Rates. *Mathematical Finance*, 7, 107–118.
28. Kimmel, R. (2004). Modeling the term structure of interest rates: A new approach. *Journal of Financial Economics*, 72, 143–183.
29. Khoshnevisan, D. (2009). A primer on stochastic partial differential equations, In D. Khoshnevisan, F. Rassoul-Agha, (Eds.), *A Minicourse on Stochastic Partial Differential Equations* (Chapter 1, pp. 1–38). Springer, Berlin.
30. Koekebakker, S., & Ollmar, F. (2005). Forward curve dynamics in the Nordic electricity market. *Managerial Finance*, 31, 73–94.

31. Longstaff, F. A., Santa-Clara, P., & Schwartz, E. S. (2001). The relative valuation of interest rate caps and swaption: Theory and empirical evidence. *Journal of Finance*, 56, 2067–2169.
32. Miltersen, K. R., & Schwartz, E. S. (1998). Pricing of options on commodity futures with stochastic term structures of convenience yields and interest rates. *Journal of Financial and Quantitative Analysis*, 33, 33–59.
33. Miltersen, K. R. (2003). Commodity price modeling that matches current observables: a new approach. *Quantitative Finance*, 3, 51–58.
34. 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.
35. 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.
36. Paschke, R., & Prokopczuk, M. (2010). Commodity derivatives valuation with autoregressive and moving average components in the price dynamics. *Journal of Banking and Finance*, 34, 2741–2752.
37. Preciado, J. (2016). High inventories help push crude oil prices to lowest levels in 13 years. U.S. Energy Information Administration. Retrieved from <http://www.eia.gov/todayinenergy/detail.php?id=24832>
38. Protter, P. (2004). *Stochastic Integration and Differential Equations*, second ed., Springer-Verlag, New York.
39. Routledge, B. R., Seppi, D. J., & Spatt, C. S. (2000). Equilibrium forward curves for commodities. *Journal of Finance*, 55, 1297–1338.
40. Sadorsky, P. (2002). Time-varying risk premiums in petroleum futures prices. *Energy Economics*, 24, 539–556.
41. Santa-Clara, P., & Sornette, D. (2001). The dynamics of the forward interest rate curve with stochastic string shocks. *Review of Financial Studies*, 14, 149–185.
42. Schwartz, E. S. (1997). The stochastic behavior of commodity prices: Implications for valuation and hedging. *Journal of Finance*, 52, 923–973.

43. Sornette, D. (1998). "String" formulation of the dynamics of the forward interest rate curve. *European Physical Journal B*, 3, 125–137.
44. 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.
45. Szymanowska, M., de Roon, F. A., Nijman, T.E., & Van den Goorbergh, R. (2014). An Anatomy of Commodity Futures Risk Premia. *Journal of Finance*, 69, 453–482.
46. Trolle, A. B., & Schwartz, E. S. (2009). Unspanned stochastic volatility and the pricing of commodity derivatives. *Review of Financial Studies*, 22, 4423–4461.

Appendix A. SDE for process Y

By definition,

$$Y(t, s) = - \int_{y=0}^{s-t} \epsilon(0, y) dy - \int_{y=0}^{s-t} \int_{v=0}^t \mu_{\epsilon}(v, y) dv dy - \int_{y=0}^{s-t} \int_{v=0}^t \sigma_{\epsilon}(v, y) d_v Z(v, y) dy$$

By interchanging the integrals we get,

$$Y(t, s) = - \int_{y=0}^{s-t} \epsilon(0, y) dy - \int_{v=0}^t \int_{y=0}^{s-t} \mu_{\epsilon}(v, y) dy dv - \int_{v=0}^t \int_{y=0}^{s-t} \sigma_{\epsilon}(v, y) dy d_v Z(v, y)$$

Note that both double integrals have deterministic integrands. As the first double integral does not have a stochastic integrator, integrals can be interchanged using the Fubini's theorem. One can show that $\int_{v=0}^t \int_{y=0}^{s-t} \sigma_{\epsilon}(v, y) dy d_v Z(v, y)$ and

$\int_{y=0}^{s-t} \int_{v=0}^t \sigma_{\epsilon}(v, y) d_v Z(v, y) dy$ are identical (Bueno-Guerrero et al., 2015, p. 234). Splitting the above integrals will yield

$$\begin{aligned} Y(t, s) = & - \int_{y=0}^s \epsilon(0, y) dy + \int_{y=s-t}^s \epsilon(0, y) dy - \int_{v=0}^t \int_{y=0}^{s-v} \mu_{\epsilon}(v, y) dy dv \\ & + \int_{v=0}^t \int_{y=s-t}^{s-v} \mu_{\epsilon}(v, y) dy dv - \int_{v=0}^t \int_{y=0}^{s-v} \sigma_{\epsilon}(v, y) dy d_v Z(v, y) \\ & + \int_{v=0}^t \int_{y=s-t}^{s-v} \sigma_{\epsilon}(v, y) dy d_v Z(v, y) \end{aligned}$$

Using $Y(0, s) = - \int_{y=0}^s \epsilon(0, y) dy$ we get:

$$Y(t, s) = Y(0, s)$$

$$\begin{aligned} & + \left(\int_{y=s-t}^s \epsilon(0, y) dy + \int_{v=0}^t \int_{y=s-t}^{s-v} \mu_\epsilon(v, y) dy dv \right. \\ & \left. + \int_{v=0}^t \int_{y=s-t}^{s-v} \sigma_\epsilon(v, y) dy d_v Z(v, y) \right) - \int_{v=0}^t \int_{y=0}^{s-v} \mu_\epsilon(v, y) dy dv \\ & - \int_{v=0}^t \int_{y=0}^{s-v} \sigma_\epsilon(v, y) dy d_v Z(v, y) \end{aligned}$$

Interchanging the order of integration of double integrals (inside parenthesis only)

$$\begin{aligned} Y(t, s) = Y(0, s) & + \int_{y=s-t}^s \left(\epsilon(0, y) + \int_{v=0}^t \mu_\epsilon(v, y) dv dy + \int_{v=0}^t \sigma_\epsilon(v, y) d_v Z(v, y) \right) dy \\ & - \int_{v=0}^t \int_{y=0}^{s-v} \mu_\epsilon(v, y) dy dv - \int_{v=0}^t \int_{y=0}^{s-v} \sigma_\epsilon(v, y) dy d_v Z(v, y) \end{aligned}$$

By using Eq. (2) the term in parenthesis can be reduced as

$$= Y(0, s) + \int_{y=s-t}^s \epsilon(t, y) dy - \int_{v=0}^t \int_{y=0}^{s-v} \mu_\epsilon(v, y) dy dv - \int_{v=0}^t \int_{y=0}^{s-v} \sigma_\epsilon(v, y) dy d_v Z(v, y)$$

By putting $y = s - v$ and $dy = -dv$ in $\int_{y=s-t}^s \epsilon(t, y) dy$ we get $\int_{v=0}^t \epsilon(v, s - v) dv$. This gives us SDE for Y in desired form

$$\begin{aligned} Y(t, s) = Y(0, s) & + \int_{v=0}^t \left(\epsilon(v, s - v) - \int_{y=0}^{s-v} \mu_\epsilon(v, y) dy \right) dv \\ & - \int_{v=0}^t \int_{y=0}^{s-v} \sigma_\epsilon(v, y) dy d_v Z(v, y) \end{aligned}$$

Appendix B. Y is a continuous semimartingale

Y is a continuous time real-valued process on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_t, \mathbf{Q})$.

The SDE (8) can be written as

$$Y(t, s) = Y(0, s) + E_t - M_{1t}$$

where $E_t = \int_{v=0}^t \left(\varepsilon(v, s - v) - \int_{y=0}^{s-v} \mu_\varepsilon(v, y) dy \right) dv$, and

$$M_{1t} = \int_{v=0}^t \int_{y=0}^{s-v} \sigma_\varepsilon(v, y) dy d_v Z(v, y)$$

Bichteler-Dellacherie Theorem states that a semimartingale can be decomposed into the sum of a local martingale and a finite variation process (see Protter, 2004, p. 146). For Y to be a continuous semimartingale, we have to show that E is a continuous process with paths of finite variation and M_1 is a continuous local martingale.

$Y(0, s)$ is \mathcal{F}_0 -measurable. E_t is continuous and adapted with initial value zero. It is evident that E_t is differentiable in interval $[0, t]$ and its derivative is integrable in Riemann sense. The total variation of E is

$$\begin{aligned} TV_E &= \int_{v=0}^t |E'_v| dv \\ &= \int_{v=0}^t \left| \varepsilon(v, s - v) - \int_{y=0}^{s-v} \mu_\varepsilon(v, y) dy \right| dv \end{aligned}$$

Since E_t is continuous and differentiable in $[0, t]$ as a function of v , its slope will always be finite which makes the integrand a positive finite quantity. Thus, total variation will be bounded as $TV_E \in [0, \infty)$ and thus, E_t is a process of finite variation.

Now consider $M_{1t} = \int_{v=0}^t \int_{y=0}^{s-v} \sigma_\varepsilon(v, y) dy d_v Z(v, y)$. Note that,

$$\begin{aligned} \mathbb{E}(M_{1t}|\mathcal{F}_{t_0}) &= \mathbb{E} \left[\int_{v=0}^{t_0} \int_{y=0}^{s-v} \sigma_\varepsilon(v, y) dy d_v Z(v, y) | \mathcal{F}_{t_0} \right] \\ &+ \mathbb{E} \left[\int_{v=t_0}^t \int_{y=0}^{s-v} \sigma_\varepsilon(v, y) dy d_v Z(v, y) | \mathcal{F}_{t_0} \right] \end{aligned}$$

, for $0 \leq t_0 \leq t \leq s$.

Let $\pi_{\check{n}}$ and $\pi_{\check{m}}$ be a sequence of finite random partitions of time windows $[t_0, t]$ with $\lim_{\check{n} \rightarrow \infty} \text{mesh}(\pi_{\check{n}}) = 0$, and $[0, s - v_i]$ with $\lim_{\check{m} \rightarrow \infty} \text{mesh}(\pi_{\check{m}}) = 0$ respectively. Discretizing the

first and second integral we get,

$$\begin{aligned} \mathbb{E}(M_{1t}|\mathcal{F}_{t_0}) &= \int_{v=0}^{t_0} \int_{y=0}^{s-v} \sigma_\varepsilon(v, y) d_v Z(v, y) dy \\ &+ \mathbb{E} \left[\lim_{\check{n}, \check{m} \rightarrow \infty} \sum_{v_i, v_{i+1} \in \pi_{\check{n}}} \sum_{y_j, y_{j+1} \in \pi_{\check{m}}} \sigma_\varepsilon(v_i, y_j) (y_{j+1} - y_j) (Z(v_{i+1}, y_j) \right. \\ &\quad \left. - Z(v_i, y_j)) | \mathcal{F}_{t_0} \right] \end{aligned}$$

By Dominated Convergence Theorem,

$$\begin{aligned} &= M_{1t_0} + \lim_{\check{n}, \check{m} \rightarrow \infty} \sum_{v_i, v_{i+1} \in \pi_{\check{n}}} \mathbb{E} \left(\sum_{y_j, y_{j+1} \in \pi_{\check{m}}} \sigma_\varepsilon(v_i, y_j) (y_{j+1} - y_j) (Z(v_{i+1}, y_j) \right. \\ &\quad \left. - Z(v_i, y_j)) | \mathcal{F}_{t_0} \right) \\ &= M_{1t_0} + \lim_{\check{n}, \check{m} \rightarrow \infty} \sum_{v_i, v_{i+1} \in \pi_{\check{n}}} \mathbb{E} \left[\mathbb{E} \left(\sum_{y_j, y_{j+1} \in \pi_{\check{m}}} \sigma_\varepsilon(v_i, y_j) (y_{j+1} - y_j) (Z(v_{i+1}, y_j) \right. \right. \\ &\quad \left. \left. - Z(v_i, y_j)) | \mathcal{F}_{v_i} \right) | \mathcal{F}_{t_0} \right] \end{aligned}$$

Again using Dominated Convergence Theorem we have,

$$\begin{aligned}
&= M_{1_{t_0}} + \lim_{\check{n}, \check{m} \rightarrow \infty} \sum_{v_i, v_{i+1} \in \pi_{\check{n}}} \mathbb{E} \left[\sum_{y_j, y_{j+1} \in \pi_{\check{m}}} \mathbb{E} (\sigma_\varepsilon(v_i, y_j)(y_{j+1} - y_j)(Z(v_{i+1}, y_j) \right. \\
&\quad \left. - Z(v_i, y_j)) | \mathcal{F}_{v_i} \right) | \mathcal{F}_{t_0} \Big] \\
&= M_{1_{t_0}} + \lim_{\check{n}, \check{m} \rightarrow \infty} \sum_{v_i, v_{i+1} \in \pi_{\check{n}}} \mathbb{E} \left[\sum_{y_j, y_{j+1} \in \pi_{\check{m}}} (y_{j+1} - y_j) \sigma_\varepsilon(v_i, y_j) \mathbb{E} \left((Z(v_{i+1}, y_j) \right. \right. \\
&\quad \left. \left. - Z(v_i, y_j)) | \mathcal{F}_{v_i} \right) | \mathcal{F}_{t_0} \right] \\
&\quad , \mathbb{E}(M_{1_t} | \mathcal{F}_{t_0}) = M_{1_{t_0}}
\end{aligned}$$

Hence, M_{1_t} is a local martingale. This proves Y is a continuous semimartingale process.

Appendix C. Quadratic variation of futures price process, F

The SDE for futures price dynamics, Eq. (15) can be written in integral form as,

$$F(t, s) = F(0, s) + M_{2_t} - M_{3_t}$$

, where $M_{2_t} = \int_{v=0}^t \sigma_S(v) F(v, s) dW(v)$ and $M_{3_t} = \int_{v=0}^t F(v, s) \int_{y=0}^{s-v} \sigma_\varepsilon(v, y) dy d_v Z(v, y)$

Since $\sigma_S(v)$ is a finite valued deterministic function and futures price process F is assumed to be continuous, bounded and progressive we have that M_{2_t} is progressively measurable, such that $\int_{v=0}^t [\sigma_S(v) F(v, s)]^2 dv < \infty$, for all t . With Wiener process as integrator, M_{2_t} is a continuous local martingale starting at 0. Its quadratic variation is

$$\langle M_2 \rangle_t = \int_{v=0}^t \sigma_S(v)^2 F(v, s)^2 dv$$

M_{3_t} can be written as, $M_{3_t} = \int_{v=0}^t F(v, s) dM_{1_v}$

M_{1t} is a continuous local martingale (see Appendix B for proof) and the integrand F has all the required properties which makes M_{3t} also a continuous local martingale with zero initial value. Its quadratic variation is given as

$$\langle M_3 \rangle_t = \int_{v=0}^t F(v, s)^2 d\langle M_1 \rangle_v = \int_{v=0}^t F(v, s)^2 \int_{y=0}^{s-v} \sigma_\varepsilon(v, y) \int_{w=0}^{s-v} \sigma_\varepsilon(v, w) c(w, y) dw dy dv$$

The quadratic covariation of M_2 and M_3 is,

$$\begin{aligned} \langle M_2, M_3 \rangle_t &= \left\langle \int_{v=0}^t \sigma_S(v) F(v, s) dW(v), \int_{v=0}^t F(v, s) dM_{1v} \right\rangle \\ &= \int_{v=0}^t \sigma_S(v) F(v, s)^2 d\langle W, M_1 \rangle_v \end{aligned}$$

By using Eq. (3) we obtain,

$$\langle M_2, M_3 \rangle_t = \int_{v=0}^t \sigma_S(v) F(v, s)^2 \int_{y=0}^{s-v} \sigma_\varepsilon(v, y) \rho(v, y) dy dv$$

Finally by using $\langle F \rangle_t = \langle M_2 \rangle_t - 2\langle M_2, M_3 \rangle_t + \langle M_3 \rangle_t$, the quadratic variation of futures price process is,

$$\begin{aligned} \langle F \rangle_t &= \int_{v=0}^t \sigma_S(v)^2 F(v, s)^2 dv - 2 \int_{v=0}^t \sigma_S(v) F(v, s)^2 \int_{y=0}^{s-v} \sigma_\varepsilon(v, y) \rho(v, y) dy dv \\ &\quad + \int_{v=0}^t F(v, s)^2 \int_{y=0}^{s-v} \sigma_\varepsilon(v, y) \int_{w=0}^{s-v} \sigma_\varepsilon(v, w) c(w, y) dw dy dv \end{aligned}$$