Spot Convenience Yield Models for Energy Assets

Paper by
Michael Ludkovski
Rene Carmona

Presentation by
Thomas Nedunthally

Outline

- What is convenience yield?
- Gibson-Schwartz model and its estimation
- Miltener extension
- Enlarging the observation equation with a stochastic market price of risk
- Empirical Implementation
- Results
- Conclusion

Convenience Yield

- A factor implied by the futures or other derivative prices of commodities
- Net benefit minus the cost of holding energy
- Unlike financial derivatives, storage of energy products is costly.
- Physical ownership of commodity, carries an associated flow of services
- Agent has the option of flexibility with regards to consumption, but this decision of postpone consumption implies storage expenses.
- δ = benefit of direct access - cost of carry

Forward model

Commodity pricing models are obtained via various assumptions on the behavior of δ_t. The implicit assumption is that S_t the spot price process of the commodity in fact exists. This is not true for some commodities, such as electricity. Even for mature markets like crude oil where spot prices are quoted daily, the exact meaning of the spot is difficult to pin down. Nevertheless, we will maintain the industry-standard assumption of traded spot asset.

By a basic no-arbitrage argument it follows that the price of a forward contract $F(t,T)$ which has payoff S_T at future time T must equal

$$F(t,T) = S_t \mathbb{E}_Q \left[e^{\int_t^T (r_s - \delta_s) ds} \right]. \tag{1}$$

Essence of the paper

- δ_t is defined indirectly as a "correction" to the drift of the spot price process.
- The next model, Gibson-Schwartz, takes the spot price and δ_t as factors.

  The forward curves given for 2 different maturities have their convenience yield term structure estimated and the results are not satisfactory.
- The Miltener extension basically tries to estimate δ_t based on an HJM type framework.
- Finally a third factor, market price of risk is added in to get satisfactory results.

Gibson Schwartz model

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ be a filtered probability space. As our state processes we consider the spot commodity asset S_t and the spot instantaneous convenience yield δ_t. According to Gibson and Schwartz, under the risk-neutral measure $\mathbb{Q}$,

$$dS_t = (r_t - \delta_t)S_t\,dt + \sigma S_t\,dW_t^1, \tag{2a}$$

$$d\delta_t = \kappa(\theta - \delta_t)dt + \gamma\,dW_t^2, \tag{2b}$$

with W^1, W^2 1-dimensional Wiener processes satisfying $d\langle W^1, W^2\rangle_t = \rho\,dt$. Note that unlike interest rates, depending on market conditions convenience yields can be either positive or negative, and so the choice of Ornstein-Uhlenbeck process for δ_t in (2b) makes sense. From now on we shall also assume that

Assumption 1 *The interest rate r_t is deterministic.*

In practice, one observes that the volatility of the convenience yield is an order of magnitude higher than the volatility of interest rates. Consequently, letting r_t be stochastic as in Schwartz [19] does not give much qualitative improvement to the model.

Mean reversion in spot price

If $\rho > 0$ then the stochastic convenience yield induces weak mean-reversion in the spot price. Observe that when S_t is increasing due to increment dW^1, thanks to positive correlation ρ, δ_t is also likely to increase. In turn, this reduces the drift of the spot. Note that this is a second-order effect. Empirically, mean-reversion of the spot has been widely documented [8], with values of $\rho \sim 0.3 - 0.7$. The theory of storage developed back in the 1950s [5] shows that the endogeneous economic link is through inventory levels: when inventories are low, shortages are likely, causing high prices, as well as valuable optionality of holding the physical asset.

Forward curve Model

Let $X_t = \log S_t$. Then the model (2) is linear in the state vector $Z_t = [X_t,\ \delta_t]$. In fact, it is exponential affine and hence admits analytical futures and options prices. As shown by Bjork and Landen [2],

$$F(t,T) = F(t,T;Z_t) = S_t e^{\int_t^T r_s ds} e^{B(t,T)\delta_t + A(t,T)} \quad \text{where} \tag{3}$$

$$B(t,T) = \frac{e^{-\kappa T} - 1}{\kappa}, \tag{4}$$

$$A(t,T) = \frac{\kappa\theta + \rho\sigma_s\gamma}{\kappa^2}(1 - e^{-\kappa(T-t)} - \kappa(T-t)) + \tag{5}$$

$$+ \frac{\gamma^2}{\kappa^3}(2\kappa(T-t) - 3 + 4e^{-\kappa(T-t)} - e^{-2\kappa(T-t)}).$$

Hence, under $\mathbb{Q}$ the forward contract follows

$$dF(t,T) = F(t,T)\left[r_t dt + \sigma\, dW_t^1 + \gamma \frac{e^{-\kappa T} - 1}{\kappa}\, dW_t^2\right]. \tag{6}$$

Also, recall that the convenience yield follows an Ornstein-Uhlenbeck (OU) process. Conditional on $\mathcal{F}_s$, δ_t is Gaussian with

$$\delta_t | \mathcal{F}_s \sim \mathcal{N}\left((1 - e^{-\kappa(t-s)})\delta_s + e^{-\kappa(t-s)}\theta, \frac{\gamma^2}{2\kappa}(1 - e^{-2\kappa(t-s)})\right).$$

Model Estimation

Because the model is conditionally Gaussian (i.e. linear), we can estimate all the parameters empirically using the standard Kalman filter method. The filtering of the Gibson-Schwartz model (2) is very easy since everything is 1-dimensional. Of course, historical estimation requires us to make an assumption on the market prices of risk λ for S_t, and λ_ϵ for the unobserved convenience yield (each random source must have its own market price of risk). The simplest choice is

Assumption 2 *λ and λ_ϵ are constant.*

Denote by $(\widetilde{W}^1, \widetilde{W}^2)$ the Brownian motions (W^1, W^2) under $\mathbb{P}$. Then the historical dynamics of δ_t are

$$d\delta_t = [\kappa(\theta - \delta_t) - \lambda_\epsilon]dt + \gamma\, d\widetilde{W}_t^2 \tag{7}$$

and we just need to adjust $\widehat{\theta} = \theta - \lambda_\epsilon/\kappa$. Consequently, our measurement and transition equations are

$$X_n = X_{n-1} + \left(r_n - \delta_n - \frac{\sigma_S^2}{2} + \lambda\right)\Delta t + \xi_n, \qquad \xi_n \sim \mathcal{N}(0, \sigma_S^2 \Delta t), \tag{8a}$$

$$\delta_n = \mathrm{e}^{-\kappa\Delta t}\delta_{n-1} + (1 - \mathrm{e}^{-\kappa\Delta t})\hat{\theta} + \eta_n, \qquad \eta_n \sim \mathcal{N}\left(0, \frac{\gamma^2}{2\kappa}(1 - \mathrm{e}^{-2\kappa\Delta t})\right). \tag{8b}$$

Convenience yield estimation for spot price

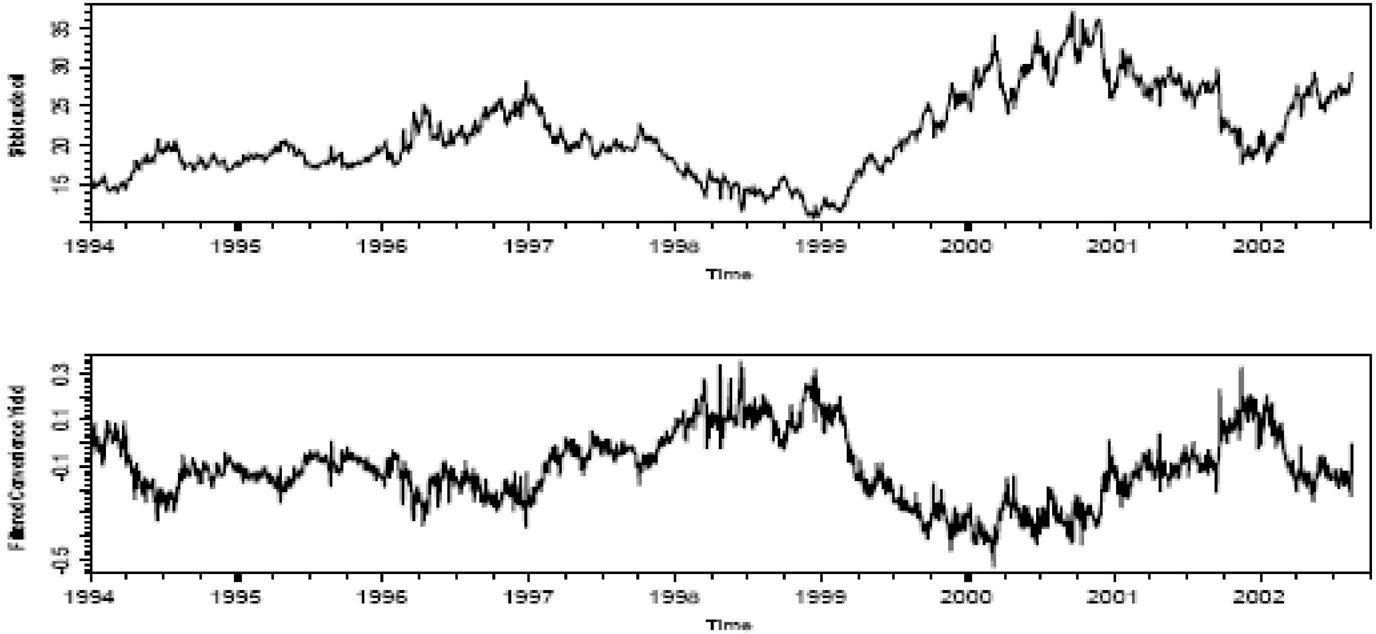

Figure 1: Filtered convenience yield for crude oil, 1994-2002. $\kappa = 0.2, \gamma = 0.5, \theta = -0.15, \rho = 0.7, \lambda = 0$.

Estimation for Forward prices with two different maturities

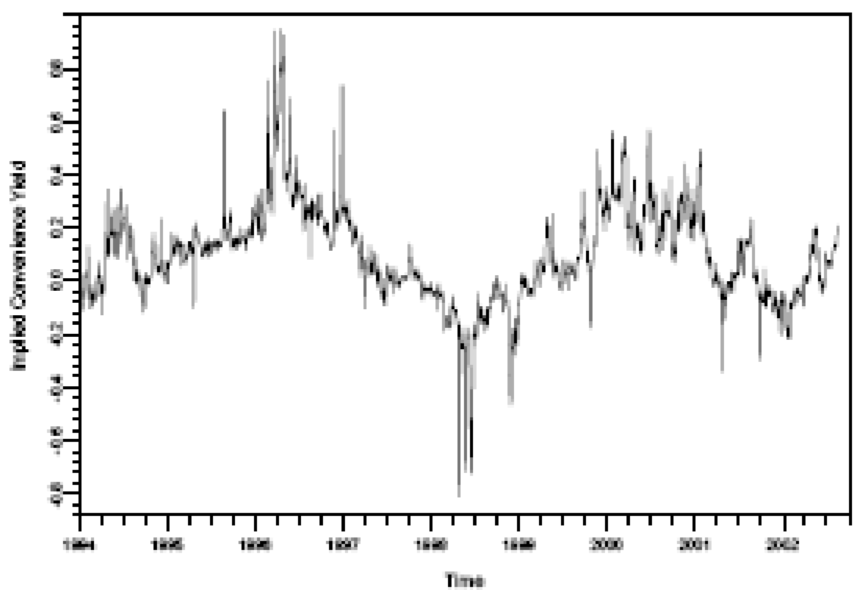

Figure 2: Implied convenience yield using a 3-month futures and same parameters as in Figure 1.

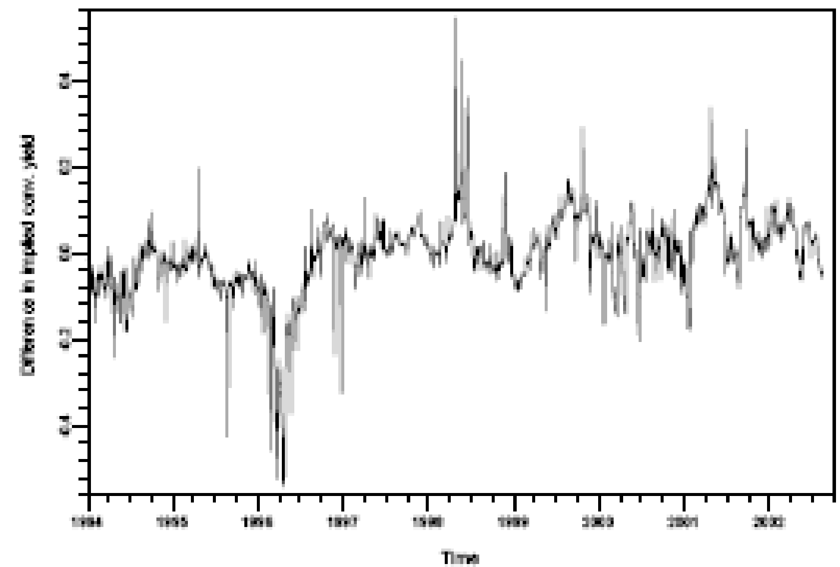

Figure 3: Difference in implied convenience yields between 3- and 12-month futures.

Unsatisfactory?

Miltener extension

- Time dependent parameter for mean reversion term Θ for δ_t
- Analogous to Hull-White extension of Vasicek model for interest rates.

Miltersen extension

The simplest choice is to make the mean-reversion level θ time-dependent, $\theta = \vartheta(t)$. This allows to fit the initial futures prices directly, in the same way that the Hull-White model can fit the initial term structure of bond prices [6]. The calibration is performed by letting the spot yield be "dragged around" its changing mean $\vartheta(t)$.

Integrating the equation for δ_t in (2) we obtain

$$\delta_t = \delta_s \mathrm{e}^{-\kappa(t-s)} + \int_s^t \mathrm{e}^{-\kappa(t-u)} \vartheta(u) du + \gamma \int_s^t \mathrm{e}^{-\kappa(t-u)} dW_u^2. \qquad (9)$$

If we define

$$\alpha_t := \int_\cdot^t \mathrm{e}^{-\kappa(t-u)} \vartheta(u) du, \qquad (10)$$

it follows that we have the deterministic shift decomposition $\delta_t = a_t + \alpha_t$, where a_t follows the mean-zero OU process

$$da_t = -\kappa a_t \, dt + \gamma \, dW_t^2.$$

To complete the calibration we must take $\vartheta(t)$ to match a chosen set of observed futures prices $F(0, T_i), i = 1, 2, \ldots, n$. For this purpose define $\epsilon(0, t)$ via

$$F(0, T_i) = S_0 e^{\int_0^{T_i} (r_s - \epsilon(0,s)) ds}. \qquad (11)$$

Miltersen extension

In HJM-type models [17], $\epsilon(0,t)$ is called the initial term structure of futures convenience yields. Then it can be shown that

$$\vartheta(t) = \frac{\epsilon_T(0,t)}{\kappa} + \epsilon(0,t) + \frac{\gamma^2}{2\kappa^2}(1 - \mathrm{e}^{-2\kappa t}) - \frac{\rho\sigma s\gamma}{\kappa} \qquad (12)$$

or alternatively $\alpha_t = \epsilon_T(0,t) + \frac{\gamma^2}{\kappa^2}(1 - \mathrm{e}^{-\kappa t})^2$.

Solving for $\epsilon(0,t)$ in (11) we obtain

$$\epsilon(0,t) = r_t - \frac{\partial \log(F(0,t))}{\partial t}. \qquad (13)$$

Miltersen extension

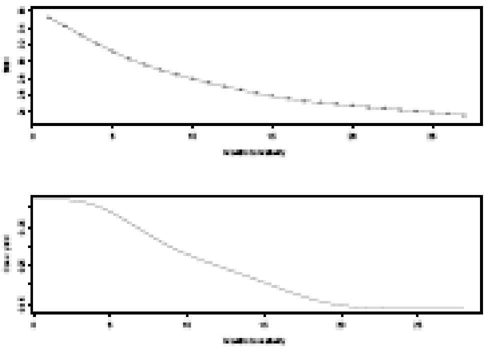

Figure 4: Crude oil forward curve and the interpolated term structure of convenience yields. Backwardation of 11/21/2000.

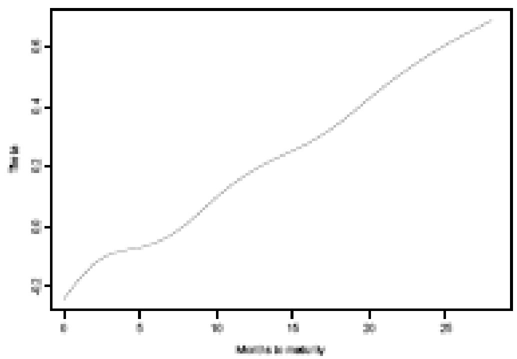

Figure 5: Term structure of mean reversion level ϑ_t, 11/21/2000.

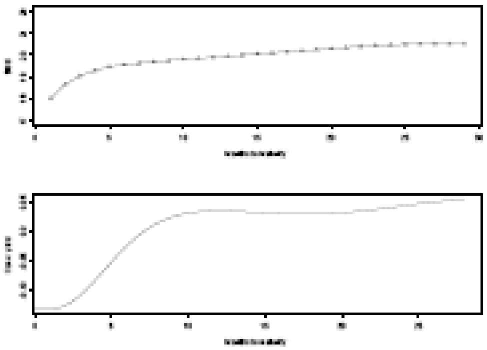

Figure 6: Crude oil forward curve and the interpolated term structure of convenience yields. Contango observed on 1/18/2002.

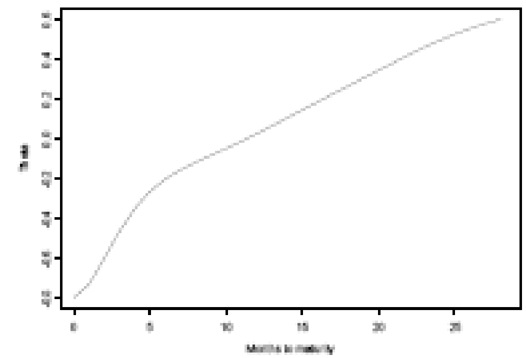

Figure 7: Term structure of mean reversion level ϑ_t, 1/18/2002.

Of course our data consists of just $\{F(0,T_i)\}$ so we must interpolate those smoothly and then take partial derivatives to infer the implied $\epsilon(0,t)$'s.

Miltersen extension

- Apparently it was the first actual implementation of the model in a paper, and it seemed to give consistent values!

Enlarging the Observation equation

- Assume a whole term structure for convenience yields, one for each maturity!

Enlarging the Observation equation

. After including the futures $F_s^i = F(s, T_i)$, $i = 1, 2, \ldots n$, the observable filtration is now $\mathcal{F}_t^O = \sigma\{S_s, F_s^i : 0 \leq s \leq t\}$.

By enlarging the set of observable instruments we are also able to address other limitations of the model. Recall that the behavior of the market price of risk λ for the spot was so far "swept under the rug". By (A2) λ is constant.

Runggaldier [18] suggests using another OU process for λ_t since mean reversion and the resulting stationarity is a desirable feature:

$$d\lambda_t = \kappa_\lambda(\bar{\lambda} - \lambda_t)dt + \sigma_\lambda dW_t^3. \tag{14}$$

It is reasonable to assume that the market price of risk carries its own Brownian motion and that W^3 is correlated with W^1, but not with W^2 the noise of the convenience yield. Intuitively, the market risk premia are independent from storage costs of the commodity. We also expect that there is negative correlation between the spot and the market price of risk $\rho_{S\lambda} < 0$. This is to strengthen the empirical mean-reversion in the spot [7].

Stochastic market price of risk

Working in an incomplete market we pick the minimal martingale measure $\mathbb{Q}$ so that the Girsanov transformation corresponding to λ_t affects only W^1 ($\widetilde{W}^1$ under $\mathbb{P}$):

$$dW_t^1 = d\widetilde{W}_t^1 - \lambda_t dt.$$

Our extended state is now $Z_t = [X_t, \ \delta_t, \ \lambda_t]$

Following the standard martingale method for pricing derivatives, we assume that the price of a futures contract is a function of the state vector

$$F_t^i = F^i(t, Z_t) = \mathbb{E}_Q[S_{T_i} | \mathcal{F}_t]\,. \qquad (15)$$

Then applying Itô's formula and using the fact that discounted traded asset prices are $\mathbb{Q}$-martingales we must have

$$dF_t^i = r_t F_t^i dt + \sigma_S S_t \frac{\partial F_t^i}{\partial S}\, dW_t^1 + \gamma \frac{\partial F_t^i}{\partial \delta}\, dW_t^2 + \sigma_\lambda \frac{\partial F_t^i}{\partial \lambda}\, dW_t^3. \qquad (16)$$

Enlarging observation equation

In our case we already have the explicit expression (3) which held for the Gibson-Schwartz model. This was derived by a replication argument. Since λ only affects the distribution under $\mathbb{P}$, the replication argument still goes through in the extended model. In other words, $\frac{\partial F_t^t}{\partial \lambda} = 0$ and (16) simplifies to

$$\frac{dF_t}{F_t} = r_t dt + \left(\sigma_S + \rho\gamma \frac{e^{-\kappa T} - 1}{\kappa}\right) \lambda_t \, dt + \sigma_S d\widetilde{W}_t^1 + \gamma \frac{e^{-\kappa T} - 1}{\kappa} d\widetilde{W}_t^2 + \alpha \, dW_t^F. \tag{17}$$

The last term $\alpha \, dW_t^F$ is the idiosyncratic risk associated with F_t and used to smooth out the data. We expect α to be an order smaller than the other volatilities in (17). Summarizing, the complete filtering model under the real world probability $\mathbb{P}$ is given by

Enlarging observation equation

The last term $\alpha\, dW_t^F$ is the idiosyncratic risk associated with F_t and used to smooth out the data. We expect α to be an order smaller than the other volatilities in (17). Summarizing, the complete filtering model under the real world probability $\mathbb{P}$ is given by

$$d\lambda_t = \kappa_\lambda(\bar{\lambda} - \lambda_t)\, dt + \sigma_\lambda dW_t^3 \tag{18a}$$

$$d\delta_t = \kappa(\hat{\theta} - \delta_t)dt + \gamma\, d\tilde{W}_t^2 \tag{18b}$$

$$dS_t = (r_t - \delta_t + \sigma_S\lambda_t)S_t\, dt + \sigma_S S_t\, d\tilde{W}_t^1 \tag{18c}$$

$$dF_t^i = \left(r_t + \sigma_S\lambda_t + \rho\gamma\frac{e^{-\kappa T_i} - 1}{\kappa}\lambda_t\right)F_t^i\, dt + \sigma_S F_t^i\, d\tilde{W}_t^1 + \gamma F_t^i\, \frac{e^{-\kappa T_i} - 1}{\kappa}\, d\tilde{W}_t^2 + \alpha\, dW_t^{F^i}. \tag{18d}$$

According to the setup, δ_t and λ_t are not observable, S_t is fully observed and F_t^i is imperfectly observed in the market.

Notice that after taking logarithms of observed prices, the entire system is still linear and hence amenable to Kalman filtering. We note that this linearity is more an artifact of the model rather than its goal. Our choice of OU processes for λ and δ has been motivated by heuristic arguments, not by modelling convenience.

Implementation

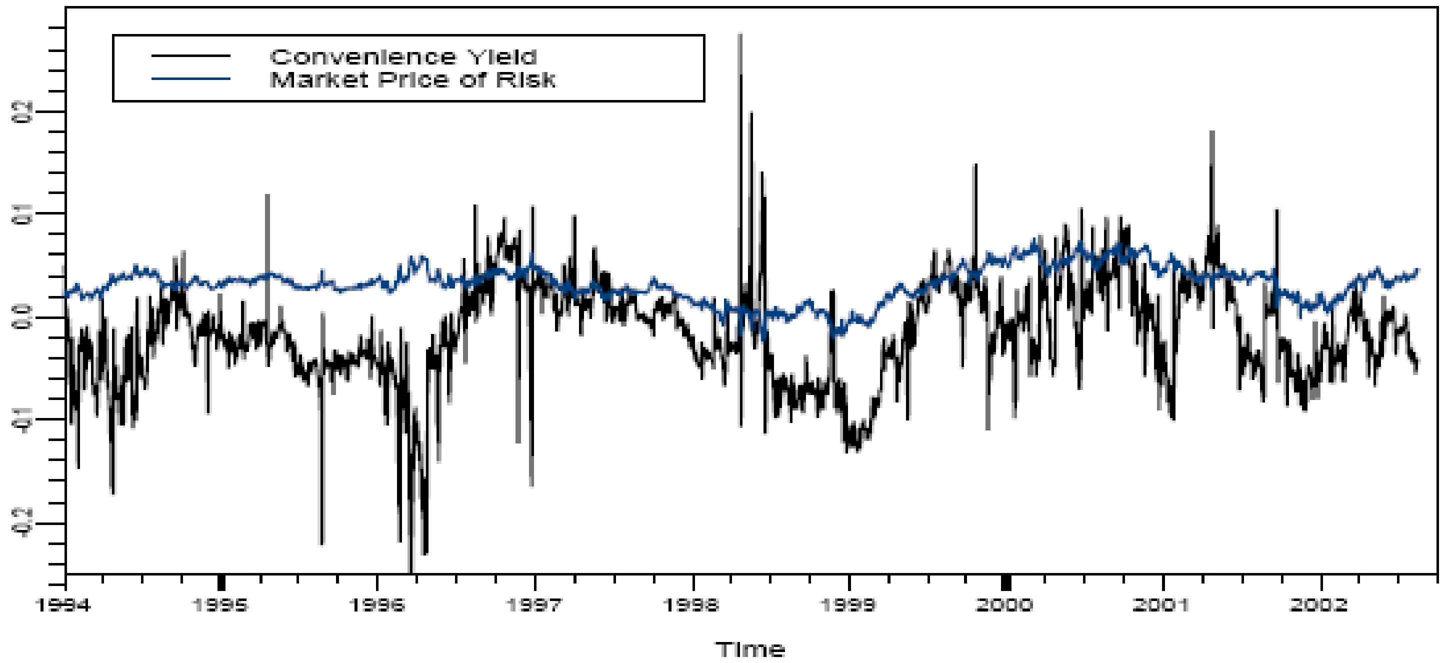

Figure 8: Filtered convenience yield for crude oil using the spot and 3-month futures. Parameter values are given in Table 1.

- After all this, the convenience yield still shows spikes which contradicts its assumption to be an OU process!

Empirical Results

First, we compare the consistency of the three-factor model (18) with respect to the forward curve. We find that our estimate of convenience yield is stable when using futures contracts with different maturities as inputs. Figure 10 shows the estimated δ_t using three different sets of two futures contracts as inputs. All three estimates are quite close to each other. Thus our model succeeds in removing the inconsistency of Figure 3.

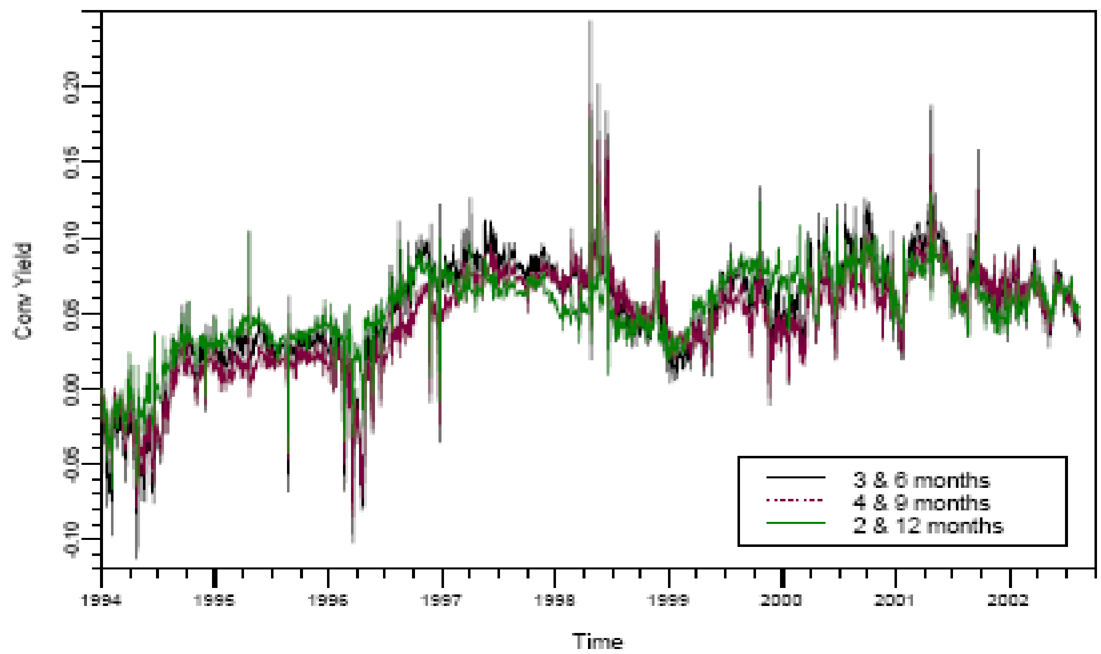

Figure 10: Estimates of δ_t using three different pairs of futures contracts.

Conclusions

- Convenience yield showing spikes goes against economic definition of an OU process for it.
- Work provides clues for future research and weaknesses of spot price models.
- Overall, believes that this study shows the need for more sophisticated term-structure models in order to explain both the spot and the forward curve.