Wu, Yuexiang (2019) *Exploring the commodity market: pricing Asian options with stochastic convenience yields and jump diffusions, and the study of the trading-date seasonality*. PhD thesis.

Exploring the Commodity Market: Pricing Asian Options with Stochastic Convenience Yields and Jump Diffusions, and the Study of the Trading-Date Seasonality

Yuexiang Wu

Master of Science in Quantitative Finance

Submitted in fulfilment of the requirements of the degree of
PhD in Economics
2019

Adam Smith Business School, College of Social Science
University of Glasgow

Abstract

The main underlying theme of this PhD thesis is the study of the commodity market. We first begin by pricing Asian options based on the Schwartz (1997) model. Asian options have been widely used in the global commodity market for its unique feature of using the average price instead of the price at maturity to determine the payoff function. We attempt to price Asian options written on commodity related future contracts under the model of three stochastic factors, namely, the spot price, the convenience yield, and the interest rate. We obtain closed-form solutions of geometric average Asian options, which will serve as control variates to price arithmetic average Asian options by Monte Carlo simulation. Our results show significant improvements in terms of simulation accuracy. We also manipulate the parameters of the model to see how the options prices behave accordingly. Next, a jump diffusion process is introduced to the model. Although analytical solution is unobtainable, a new numerical method is found to price arithmetic average Asian options with jumps, which lead to observable accuracy improvements.

During our journey to further explore the behaviour of the commodity futures prices, we found a new seasonality pattern. The traditional idea of seasonality in the future market relates to the maturity date of a future contract. However, we find a new seasonal pattern in the futures prices that relates to the trading dates. We decide to explore such phenomenon in three energy commodity markets, namely, natural gas, gasoline, and crude oil. To conduct our initial empirical research, we design the so-called backward curve, as opposite to the forward curve, to visually illustrate the pattern of the trading-date seasonality. We find that when the prices of a collection of future contracts with the same maturity month can be averaged over the different years, the seasonality of trading dates is obvious to observe. We also find an interesting change of behaviour in the natural gas futures prices. Then, we conduct multiple statistical tests to further confirm our findings, which include the Kruskal-Wallis test, the autocorrelation test, and the power spectrum test. The results show strong evidence to support the existence of the trading-date seasonality.

In light of what we find in the second chapter, we decide to look further into the new seasonality that relates to the trading dates, by constructing a trading strategy that is designed specifically to profit from the new seasonal pattern in three commodity markets. The results show promising profit over the long run for all three commodities, with relatively low risks. Then, we establish a model based on the Sorensen (2002) model, with the introduction of an arbitrage factor to capture the trading-date seasonality. We calibrate the

model using the Kalman filter in the state space form, and the results suggest that the vast majority of the parameters are highly statistically significant in explaining the movement of the futures prices in the three commodity markets.

Pricing Asian Options with Stochastic Convenience Yield and Jump Diffusions

Abstract

We attempt to price Asian options written on commodity related future contracts under the model of three stochastic factors, namely, the spot price, the convenience yield, and the interest rate. We obtain closed-form solutions of geometric average Asian options, which will serve as control variates to price arithmetic average Asian options by Monte Carlo simulation. Our results show significant improvements in terms of simulation accuracy. We also manipulate the parameters of the model to see how the options prices behave accordingly. Next, a jump diffusion process is introduced to the model. Although analytical solution is unobtainable for geometric average Asian options, a new numerical method is found to price them with jumps, which lead to observable accuracy improvements to price the arithmetic counterparts.

1 Introduction

Option pricing of commodity related financial instruments has always been a popular topic among both academics and industry. A large amount of efforts has been devoted by academics to this realm, and yet the true behaviour of commodity prices remains puzzling and unpredictable. A number of methods have been developed to price the financial products with commodities as their underlying assets. For example, one important family of such methods starts from Heath *et al* (1992), or the HJM model in the following content, where the authors take as given the prices of the zero-coupon bond, and then attempt to price contingent claims that are interest rate sensitive. In other words, it intends to directly price the entire forward curve of the interest rate. A significant amount of researches has been developed under the HJM framework, such as Inui and Kijima (1998), Jong and Clara (1999) and Agca (2005), to name a few. Another group of methods to price contingent claims is with the help of a finite number of parameters that carry verifiable economic meanings, or the so called state variables. Our paper falls into the second group, where we attempt to price Asian options written on commodity futures with the state variables of the spot price, the convenience yield and the interest rate.

What lays the foundation of our paper is the theory of storage, which attempts to deconstruct the difference between the instantaneous spot price and the price of future contracts of the same underlying asset. Early work on theory of storage constructed a solid foundation on identifying the three factors that constitute the difference between the spot price and the futures price of the same underlying, namely, the interests forgone for keeping the commodity, the storage cost of the commodity, and the convenience yield[1].

In commodity market, the convenience yield has been argued to play one of the major roles in interpreting the movements of futures prices, and has thus inspired a large number of studies to reveal its true nature. To name a few, Fama and French (1987) find statistical evidence to support the theory of storage, and that seasonality plays a significant role in explaining futures prices. Fama and French (1988) test and confirm the hypothesis from the theory of storage that the marginal

[1] The convenience yield represents the advantage of physically owning and storing the commodity asset, over holding some financial contract of the same asset as the underlying. An example could be that the owner of gold could store his asset as inventory, making it into some jewellery when the price is too low to make a profit from selling it, or selling it when the price is high. However, the owner of a financial contract of gold only has the option to keep it or sell it based on its price. See Kaldor (1939); Working (1948); Brennan (1958); Telser (1958).

convenience yield falls at a decreasing rate with the increase of inventory, by studying the variability differences between the spot and futures prices. Heinkel *et al* (1990) confirm the negative relationship between the level of aggregate inventory and the convenience yield, and also find two additional determinants of the convenience yield, namely, marginal production costs, and spot prices of the commodity. Routledge *et al* (2000) find a positive relationship between the convenience yield and the spot prices, and a time-varying correlation between them. Casassus and Dufresne (2005) construct a three-factor model where the convenience yield depends on both the spot prices and the interest rate, and a time-varying risk premia is embedded. Researchers also treat the convenience yield as a resemblance of plain vanilla call option, since the convenience yield mainly represents the value of the option that the owner of an physical asset carries over that of a financial contract (see Milonas and Thomadakis (1997)).

We conduct our study following a specific research route where the convenience yield are treated to carry the characteristics of dividend to a stock, and modelled by a stochastic process (as in Gibson and Schwartz, 1990; Schwartz, 1997; Miltersen and Schwartz, 1998). We base our model largely on Schwartz (1997), where three stochastic factors are embedded, including the spot price, the convenience yield, and the interest rate. The stochastic process of the convenience yield and the interest rate are assumed to follow a mean-reverting processes, in light of the study by Gibson and Schwartz (1990) and Vasicek (1977).

We further extend the Schwartz (1997) model to price Asian options, which is commonly traded in the commodity market all over the world. Its unusual feature of using the average of the underlying prices over the option period instead of the spot price at the maturity to calculate the option price gives it unique advantages over plain vanilla options, such as preventing some hostile manipulation of the spot price when close to maturity date. However, it is also well known that most commonly traded Asian options use arithmetic average of the spot price to calculate the price, which leads to no analytical solution. This is because although the spot prices are assumed to follow the log-normal distribution, its arithmetic average does not. As a result, researchers seek for numerical solutions such as Monte Carlo simulation to price arithmetic average Asian options. Kemna and Vorst (1990) suggest using geometric average Asian option that adopts the geometric average to calculate the option's price as a control variate to reduce the simulation error to yield more accurate results. This is feasible because the geometric average of the spot prices is still log-normally distributed, and is arguably closely related to the arithmetic counterpart. Accordingly, we follow the same procedure of Kemna and Vorst (1990) to first find the analytical solution to the geometric average Asian option, then employ it as a control variate to price arithmetic average Asian option using Monte Carlo simulation. Our results show significant improvement of accuracy, in comparison with standard Monte Carlo simulation with no variance reduction method, and with antithetic method (see Boyle *et al* (1997) for the references of variance reduction methods in Monte Carlo simulation, including the antithetic and control variate methods). Then, we manipulate the parameters of the model in pairs to discover how the Asian option prices react to changes of the parameters.

Our next step is to include jump diffusions in our model. A jump could be referred to as a spike in the price movement. It is usually triggered by the sudden arrival of some unexpected news that has an immediate and profound effect on the underlying price. Merton (1976) was arguably the first to introduce jumps into options pricing, where a closed-form solution is derived for plain vanilla European options. Hilliard and Reis (1998) add jump diffusions to the Schwartz (1997) model, and also obtain analytical solution to European options. Given the important role that commodities play in the development of the modern world, the supply and demand of many commodities are prone to geopolitical conflicts between different countries, which could trigger some huge movement of the prices in a blink of an eye. Hence, we aim to extend the model of three stochastic factors to include a jump diffusion in the spot price to price Asian options written on commodity futures. The difficulty lies in that there appears to be no closed-form solutions for either arithmetic average or geometric average Asian options. This is because Asian options are path-dependent, while plain vanilla options are not. As a result, the price of plain vanilla options only depends on the jump size and the accumulated effect of all the jumps on the spot price at maturity, but not the time of each jump occurrence. Nevertheless, Asian options are path-dependent and, hence, the time of each jump matters. When the average of the spot prices is calculated to obtain the price of an Asian option, a jump that occurs at the beginning of an option's period obviously affects the price differently than if it occurs near the end.

Provided such challenges, we argue that conditional on knowing when each jump occurs during the option period, there is a unique analytical solution to a geometric average Asian option. The reason is that if the timing of each jump could be assumed as known, then the only unknown random variable of the jump process is the jump size, which follows log-normal distribution, and so are the spot price and the geometric average of it. Therefore, a closed-form solution is feasible for such geometric average Asian option, which will serve as a control variate to price the corresponding arithmetic average Asian option using Monte Carlo simulation. Nonetheless, the result is conditional upon a specific set of jumping times over the option period that we assumed as known. Hence, the next step would be to simulate a large number of different sets of jumping times, each of which lead to a unique solution to an arithmetic average Asian option. The unbiased final result emerges by taking the mean of all these solutions. We compare our method of control variate with standard Monte Carlo with no variance reduction method and with antithetic method. The results show observable improvement in terms of simulation accuracy.

The remainder of this paper is organized as follows. Section 2 introduces the Schwartz (1997) model that we adopt in this paper. We derive analytical solutions of geometric average Asian options as well as the results of pricing arithmetic average Asian options using Monte Carlo simulations. In Section 3, we manipulate the parameters of the model to see how the Asian option prices behave under different parameters. The model with jump diffusions is discussed in Section 4. Conclusions are drawn in Section 5.

2 The Model

We follow Schwartz (1997) to construct the two-factor and three-factor models for Asian options written on future contracts, respectively. General conditions of option pricing apply here, including log-normal property for the underlying price, continuous time framework, and no transaction costs, taxes, or any limitation on short sale.

2.1 Two-Factor Model

In the two-factor model, both the underlying spot price and instantaneous convenience yield are driven by the following stochastic processes, namely,

$$dS = (\mu - \delta)Sdt + \sigma_1 SdZ_1 \ , \tag{1}$$

$$d\delta = \kappa(\alpha - \delta)dt + \sigma_2 dZ_2 \ , \tag{2}$$

$$dZ_1 dZ_2 = \rho dt \ . \tag{3}$$

Equation (1) shows the stochastic process for the underlying price, where the drift indicates a theoretically negative effect of instantaneous convenience yield, δ, on its long-term growth rate, μ. This is consistent with the role convenience yield plays in the theory of storage. The instantaneous convenience yield follows an Ornstein-Uhlenbeck stochastic process as shown in equation (2), which is in line with Gibson and Schwartz (1990). κ and α denote the speed of adjustment and the long-term mean of the convenience yield, respectively. σ_1^2 and σ_2^2 represent the variances of the underlying prices and the convenience yield, respectively. Both dZ_1 and dZ_2 are the increments of standard Brownian motion, with ρ being the correlation coefficient.

Under the two-factor model, we assume a constant interest rate, r. It is also worth noting that the convenience yield is not a tradable asset, or it cannot be hedged. As a result, it must carry a market price of risk after adjusting for risk, λ, which we assume as a constant. The resulted stochastic processes under the risk-neutral measure, $\mathbb{Q}^*$, are expressed as,

$$dS = (r - \delta)Sdt + \sigma_1 S dZ_1^* \ , \tag{4}$$

$$d\delta = \kappa(\hat{\alpha} - \delta)dt + \sigma_2 dZ_2^* \ , \tag{5}$$

$$dZ_1^* dZ_2^* = \rho dt \ . \tag{6}$$

where

$$\hat{\alpha} = \alpha - \frac{\lambda}{\kappa} \ .$$

It is easy to observe the change in the drift term of the spot underlying process, where μ has been replaced by the risk-free rate, r, under the risk-neutral measure. The effect of λ on the long term mean of the instantaneous convenience yield, α, has also been absorbed and addressed when the risk-adjusted long term mean, $\hat{\alpha}$, takes the place of α.

By constructing a no-arbitrage portfolio including two future contracts with different maturities and the spot underlying asset, allowing one non-traded variable, namely the convenience yield, Schwartz (1997) shows that the futures prices, $F(S, \delta, \tau)$, with time till maturity $\tau = T - t$, must follow the partial differential equation,

$$\frac{1}{2}\sigma_1^2 S^2 F_{SS} + \sigma_1\sigma_2\rho S F_{S\delta} + \frac{1}{2}\sigma_2^2 F_{\delta\delta} + (r - \delta)SF_S + (\kappa(\hat{\alpha} - \delta))F_\delta - F_\tau = 0 \ , \tag{7}$$

subject to the boundary condition,

$$F(S, \delta, 0) = S \ .$$

The solution of the futures price to the above equations can be given as follows,

$$F(S, \delta, \tau) = S \ exp(A(\tau)) \ , \tag{8}$$

where

$$A(\tau) = -\frac{\delta}{\kappa}(1 - e^{-\kappa\tau}) + (r - \hat{\alpha} + \frac{\sigma_2^2}{2\kappa^2} - \frac{\sigma_1\sigma_2\rho}{\kappa})\tau + \frac{\sigma_2^2(1 - e^{-2\kappa\tau})}{4\kappa^3} + (\hat{\alpha}\kappa + \sigma_1\sigma_2\rho - \frac{\sigma_2^2}{\kappa})\frac{1 - e^{-\kappa\tau}}{\kappa^2} \ .$$

Jamshidian and Fein (1990) and Bjerksund (1991) independently derived the above solution in their unpublished notes.

We now turn to pricing Asian options written on future contracts. It is well known that under most circumstances, arithmetic average Asian option does not yield closed-form solution, because the arithmetic average of the log-normal underlying prices does not follow log-normal distribution. However, it is not the same with geometric average Asian option. Equation (8) clearly shows that the price of a future contract is proportional to the underlying price, indicating that the futures price also follows log-normal distribution, and so is the geometric average of the futures price. Hence, the derivation of closed-form solution for geometric average Asian option is feasible. As a result, Kemna and Vorst (1990) suggest pricing arithmetic average Asian options using numerical methods such as the Monte Carlo simulation, with the analytical solution of geometric average Asian option as a control variate for variance reduction purpose. The method leads to unbiased results and is extremely effective, because the structural similarity between the two types of Asian options guarantees a high covariance between their prices and thus significantly reduces the variance of the simulated prices of arithmetic average Asian options.

Therefore, we seek for analytical solutions of geometric average Asian options written on future contracts, whose prices are driven by the stochastic processes described by equation (4) to (6). The price of the geometric average Asian option with maturity T written on future contracts with maturity time $\hat{T}$ at time t under the risk-neutral measure can be represented by the following equation,

$$GA(t, T, \hat{T}) = e^{-r(T-t)}\mathbb{E}^* \ max(G(t, T, \hat{T}) - K, 0) \ , \tag{9}$$

where

$$G(t, T, \hat{T}) = exp(\frac{1}{T}\int_0^T lnF(S(u), \delta(u), \hat{T} - u) \ du) \ . \tag{10}$$

It is assumed that $0 \leq t \leq T \leq \hat{T}$, or the underlying futures contract cannot expire before the Asian option does. It is also interesting to see how the current time may change the calculation of equation (10). Assuming t is strictly larger than 0, equation (10) could be decomposed into two parts,

$$G(t,T,\hat{T}) = exp\{\frac{1}{T}[\int_0^t lnF(S(u),\delta(u),\hat{T}-u)\,du + \int_t^T lnF(S(v),\delta(v),\hat{T}-v)\,dv]\}\ , \quad (11)$$

where the first part of the exponent in equation (11) is already known at time t and could easily be taken out from the expectation part of equation (9). As a result, the strike, K, and the discount factor in equation (9) will change its value correspondingly, while the technical problem remains to solve the second part of equation (11). For notational simplicity, it is assumed that the current time, t, coincide with the starting point of the Asian option. In other words, $t = 0$. The extension to any $t > 0$ is fairly straightforward.

Given that the futures prices follow log-normal distribution and a fixed interest rate, Kemna and Vorst (1990) show that equation (9) must lead to the following solution,

$$GA(0,T,\hat{T}) = e^{-rT}[e^{E+\frac{1}{2}V}N(\frac{E-ln(K)+V}{\sqrt{V}}) - KN(\frac{E-ln(K)}{\sqrt{V}})]\ , \quad (12)$$

where E and V denote the expectation and variance of the geometric average of the log-normal futures prices , and N represents the standard normal distribution function. The solution to E and V is, respectively,

$$\begin{aligned} E_{2-factor} = {} & ln(S_0) - \frac{\delta_0}{\kappa} + \frac{1}{2}(r - \frac{\sigma_1^2}{2} - \hat{\alpha})T + (\delta_0 T - \hat{\alpha}T - \frac{\hat{\alpha}}{\kappa})\frac{e^{-\kappa\hat{T}}}{\kappa T} \\ & + \hat{\alpha}\frac{e^{-\kappa(\hat{T}-T)}}{\kappa^2 T} + (r - \hat{\alpha} + \frac{\sigma_2^2}{2\kappa^2} - \frac{\sigma_1\sigma_2\rho}{\kappa})(\hat{T} - \frac{T}{2}) \\ & + \frac{\sigma_2^2}{4\kappa^3 T}(T - \frac{e^{-2\kappa(\hat{T}-T)} - e^{-2\kappa\hat{T}}}{2\kappa}) \\ & + \frac{\hat{\alpha}\kappa + \sigma_1\sigma_2\rho - \frac{\sigma_2^2}{\kappa}}{\kappa^2 T}(T - \frac{e^{-\kappa(\hat{T}-T)} - e^{-\kappa\hat{T}}}{\kappa})\ , \end{aligned} \quad (13)$$

$$\begin{aligned} V_{2-factor} = {} & \frac{1}{T^2}(\frac{1}{3}\sigma_1^2 T^3 + \frac{\sigma_2^2}{\kappa^2}(\frac{1}{4\kappa^3}e^{-2\kappa(\hat{T}-T)} - \frac{4}{\kappa^3}e^{-\kappa(\hat{T}-T)} \\ & - (\frac{T^2}{2\kappa} + \frac{T}{2\kappa^2} + \frac{1}{4\kappa^3})e^{-2\kappa\hat{T}} + (2\frac{T^2}{\kappa} + 4\frac{T}{\kappa^2} + \frac{4}{\kappa^3})e^{-\kappa\hat{T}} + \frac{T^3}{3}) \\ & + \frac{2\sigma_1\sigma_2\rho_1}{\kappa}(\frac{2}{\kappa^3}e^{-\kappa(\hat{T}-T)} - (\frac{T^2}{\kappa} + \frac{2T}{\kappa^2} + \frac{2}{\kappa^3})e^{-\kappa\hat{T}} - \frac{T^3}{3}))\ , \end{aligned} \quad (14)$$

where S_0 and δ_0 denote the spot price and the instantaneous convenience yield at the start of the Asian option period, when $t = 0$.

The analytical solution to geometric average Asian option will serve as a control variate to help reduce the error of Monte Carlo simulation for pricing arithmetic average Asian option numerically. To prove its effectiveness, we compare the results of pricing an arithmetic average Asian option by standard Monte Carlo simulation with no variance reduction method, with antithetic method, and with control variate method. The antithetic method provides another easy and effective way to reduce simulation error, by simply adding a negative sign in front of all the simulated Brownian motion increments to create a new stream of random variables with the same expectation and variance at minimal extra computational cost. Each stream of random variables produces one unbiased estimates of the arithmetic average Asian options price. The average of these two estimates is also unbiased, but with a much shrunken variance (see Boyle *et al* (1997) for references).

The results from the three Monte Carlo methods are shown in Table 1, where parameters are given with different values as indicated in the first two columns. Column 3 to 5 illustrates options prices and the standard deviations incurred by the simulation (listed in the brackets) using different Monte Carlo methods. It is apparent to see that control variate method consistently yields much shrunk standard deviation. Given 20,000 simulations in our demonstration, for an Asian option contract worth around \$2 - \$3, the standard errors incurred by standard Monte Carlo simulation could be as large as 2 to 3 cents, or 1% to the price. In the case of antithetic method, they tend to shrink to as low as 0.5%. However, when control variate method is applied, the standard errors are usually around 0.1 cent, or 0.05%. On average, the standard errors from simulations with control variate is more than 16 times smaller than those from the standard Monte Carlo method, and 8.5 times smaller than from the antithetic method.

2.2 Three-Factor Model

The assumption of constant interest rate is relaxed in the three-factor model, where we adopt a Ornstein-Uhlenbeck stochastic process for the instantaneous interest rate, inspired by Vasicek (1977) when the mean-reverting feature of the interest rate was first captured. The stochastic process is very similar to that of the convenience yield. Under the risk-neutral measure, $\mathbb{Q}^*$, the stochastic processes for the underlying spot price, the instantaneous convenience yield and the instantaneous interest rate can be represented as follows,

$$dS = (r - \delta)Sdt + \sigma_1 S dZ_1^* \ , \tag{15}$$

$$d\delta = \kappa(\hat{\alpha} - \delta)dt + \sigma_2 dZ_2^* \ , \tag{16}$$

$$dr = a(\hat{m} - r)dt + \sigma_3 dZ_3^* \ , \tag{17}$$

$$dZ_1^* dZ_2^* = \rho_1 dt \ , \qquad dZ_2^* dZ_3^* = \rho_2 dt \ , \qquad dZ_1^* dZ_3^* = \rho_3 dt \ , \tag{18}$$

where

$$\hat{\alpha} = \alpha - \frac{\lambda_1}{\kappa} \ , \qquad \hat{m} = m - \frac{\lambda_2}{a} \ .$$

Here, a and $\hat{m}$ denote the speed of adjustment and the risk-adjusted long term mean of the interest rate, respectively. λ_1 and λ_2 represent the market price of convenience yield risk and interest rate risk, respectively. The three-factor model is a simple extension from the two-factor model, where the interest rate fluctuations can be rigorously modelled with the flexibility to change its value in the short run. This is of particular interests when certain commodities are empirically related to or heavily influenced by monetary policy change (see Frankel (2006)).

Schwartz (1997) shows that the futures prices under the three-factor model, $F(S, \delta, r, \tau)$, must satisfy the following partial differential equation,

$$\frac{1}{2}\sigma_1^2 S^2 F_{SS} + \frac{1}{2}\sigma_2^2 F_{\delta\delta} + \frac{1}{2}\sigma_3^2 F_{rr} + \sigma_1\sigma_2\rho_1 S F_{S\delta} + \sigma_2\sigma_3\rho_2 F_{\delta r} + \sigma_1\sigma_3\rho_3 S F_{Sr} + (r - \delta)SF_S + \kappa(\hat{\alpha} - \delta)F_\delta + a(\hat{m} - r)F_r - F_\tau = 0 \ , \tag{19}$$

subject to the boundary condition,

$$F(S, \delta, r, 0) = S \ .$$

The solution to the above equation can be shown as,

$$F(S, \delta, r, \tau) = Sexp(A(\tau) + B(\tau) + C(\tau)) \ , \tag{20}$$

where,

$$
\begin{aligned}
A(\tau) =& -\frac{\delta(1-e^{-\kappa\tau})}{\kappa} \\
B(\tau) =& \frac{r(1-e^{-a\tau})}{a} \\
C(\tau) =& \frac{(\kappa\hat{\alpha}+\sigma_1\sigma_2\rho_1)((1-e^{-\kappa\tau})-\kappa\tau)}{\kappa^2} \\
&- \frac{\sigma_2^2(4(1-e^{-\kappa\tau})-(1-e^{-2\kappa\tau})-2\kappa\tau)}{4\kappa^3} \\
&- \frac{(a\hat{m}+\sigma_2\sigma_3\rho_3)((1-e^{-a\tau})-a\tau)-a\tau}{a^2} \\
&- \frac{\sigma_3^2(4(1-e^{-a\tau})-(1-e^{-2a\tau})-2a\tau)}{4a^3} \\
&+ \sigma_2\sigma_3\rho_2(\frac{(1-e^{-\kappa\tau})+(1-e^{-a\tau})-(1-e^{-(\kappa+a)\tau})}{\kappa a(\kappa+a)} \\
&+ \frac{\kappa^2(1-e^{-a\tau})+a^2(1-e^{-\kappa\tau})-\kappa a^2\tau-a\kappa^2\tau}{\kappa^2a^2(\kappa+a)}) \ .
\end{aligned}
$$

The price of the geometric average Asian option at time t with maturity T written on a future contract with maturity $\hat{T}$ under $\mathbb{Q}^*$ is represented by the following equation,

$$GA(t,T,\hat{T}) = \mathbb{E}^* \ [e^{-r(T-t)}max(G(t,T,\hat{T})-K,0)] \ , \tag{21}$$

where

$$G(t,T,\hat{T}) = exp(\frac{1}{T}\int_0^T lnF(S(u),\delta(u),r(u),\hat{T}-u) \ du) \ . \tag{22}$$

To find out the solution to equation (21), we perform a change of numeraire, where the risk-neutral measure, $\mathbb{Q}^*$, will be transformed into the T-forward measure, $\mathbb{Q}^T$, using a zero-coupon bond with maturity T as numeraire (the process of the transformation will be provided in the Appendix). The solution to the Asian option price at time $t = 0$ is given by the following equations,

$$
\begin{aligned}
GA(0,T,\hat{T}) &= P(0,T)\mathbb{E}^{\mathbb{T}}max(G(0,T,\hat{T})-K,0) \\
&= P(0,T)[e^{E+\frac{1}{2}V}N(\frac{E-ln(K)+V}{\sqrt{V}})-KN(\frac{E-ln(K)}{\sqrt{V}})]
\end{aligned}
\tag{23}
$$

where $P(t,T)$, the price of the zero-coupon bond with maturity T at time t, are provided by the following equations,

$$P(0,T) = A(0,T)e^{-r_0B(0,T)} \ , \tag{24}$$

where

$$B(0,T) = \frac{1-e^{-aT}}{a}$$
$$A(0,T) = exp[(\hat{m} - \frac{\sigma_3^2}{2a^2})(B(0,T) - T) - \frac{\sigma_3^2}{4a}B^2(0,T)] \quad (25)$$

And the expectation, E, and variance, V, in equation (23), are

$$\begin{aligned}
E_{3-factor} = & \frac{1}{T}\{[ln(S_0) + \frac{r_0}{a} - \frac{\delta_0}{\kappa}]T - \frac{1}{2}(\frac{1}{2}\sigma_1^2 - \hat{m} + \hat{\alpha})T^2 \\
& + \frac{\delta_0 - \hat{\alpha}}{\kappa}Te^{-\kappa\hat{T}} + \frac{\hat{\alpha}}{\kappa^2}e^{-\kappa\hat{T}}(e^{\kappa T} - 1) - \frac{r_0 - \hat{m}}{a}Te^{-a\hat{T}} - \frac{\hat{m}}{a^2}e^{-a\hat{T}}(e^{aT} - 1) \\
& + \frac{\kappa\hat{\alpha} + \sigma_1\sigma_2\rho_1}{\kappa^2}(l_1 - \kappa T\hat{T} + \frac{1}{2}\kappa T^2) - \frac{\sigma_2^2}{4\kappa^3}(4l_1 - l_3 - 2\kappa T\hat{T} + \kappa T^2) \\
& - \frac{a\hat{m} + \sigma_1\sigma_3\rho_3}{a^2}(l_2 - aT\hat{T} + \frac{1}{2}aT^2) - \frac{\sigma_3^2}{4a^3}(4l_2 - l_4 - 2aT\hat{T} + aT^2) \\
& + \frac{\sigma_2\sigma_3\rho_2}{\kappa a(\kappa + a)}(l_1 + l_2 - l_5) + \frac{\sigma_2\sigma_3\rho_2}{\kappa^2 a^2(\kappa + a)}[\kappa^2 l_2 + a^2 l_1 - a\kappa T(\hat{T} - \frac{1}{2}T)(a + \kappa)] \\
& - \frac{\sigma_1\sigma_3\gamma_1}{a}[\frac{T^2}{2} - \frac{1}{a^2} + (\frac{T}{a} + \frac{1}{a^2})e^{-aT}] - \frac{\sigma_2\sigma_3\gamma_2}{\kappa a}[e^{-\kappa\hat{T}}(-\frac{T}{\kappa} + \frac{e^{\kappa T} - 1}{\kappa^2}) \\
& - e^{-\kappa\hat{T} - aT}(-\frac{T}{\kappa + a} + \frac{e^{(\kappa + a)T} - 1}{(\kappa + a)^2}) - \frac{T^2}{2} + \frac{1}{a^2} - \frac{T}{a}e^{-aT} - \frac{e^{-aT}}{a^2}] \\
& + \frac{\sigma_3^2\gamma_3}{a^2}[e^{-a\hat{T}}(-\frac{T}{a} + \frac{e^{aT} - 1}{a^2}) - e^{-a(\hat{T} + T)}(-\frac{T}{2a} + \frac{e^{2aT} - 1}{4a^2}) \\
& - \frac{T^2}{2} + \frac{1}{a^2} - \frac{T}{a}e^{-aT} - \frac{e^{-aT}}{a^2}]\} \ , \quad (26)
\end{aligned}$$

where

$$l_1 = T - \frac{e^{-\kappa\hat{T}}}{\kappa}(e^{\kappa T} - 1)$$
$$l_2 = T - \frac{e^{-a\hat{T}}}{a}(e^{aT} - 1)$$
$$l_3 = T - \frac{e^{-2\kappa\hat{T}}}{2\kappa}(e^{2\kappa T} - 1)$$
$$l_4 = T - \frac{e^{-2a\hat{T}}}{2a}(e^{2aT} - 1)$$
$$l_5 = T - \frac{e^{-(\kappa + a)\hat{T}}}{(\kappa + a)}(e^{(\kappa + a)T} - 1)$$

$$
\begin{aligned}
V_{3-factor} = & \frac{1}{T^2}\{\frac{1}{3}T^3\sigma_1(\alpha_1^2+\beta_1^2+\gamma_1^2) \\
& + \frac{\sigma_2^2(\beta_2^2+\gamma_2^2)}{\kappa^2}(\frac{T^3}{3}+\frac{1}{4\kappa^3}e^{-2\kappa(\hat{T}-T)}-\frac{4}{\kappa^3}e^{-\kappa(\hat{T}-T)}-m_1+m_2) \\
& + \frac{\sigma_3^2\gamma_3^2}{a^2}(\frac{T^3}{3}+\frac{1}{4a^3}e^{-2a(\hat{T}-T)}-\frac{4}{a^3}e^{-a(\hat{T}-T)}-m_3+m_4) \\
& + \frac{2\sigma_1\sigma_2(\beta_1\beta_2+\gamma_1\gamma_2)}{\kappa}(\frac{2}{\kappa^3}e^{-\kappa(\hat{T}-T)}-\frac{1}{2}m_2-\frac{T^3}{3}) \\
& - \frac{2\sigma_1\sigma_3\gamma_1\gamma_3}{a}(\frac{2}{a^3}e^{-a(\hat{T}-T)}-\frac{1}{2}m_4-\frac{T^3}{3}) \\
& - \frac{2\sigma_2\sigma_3\gamma_2\gamma_3}{\kappa a}[(\frac{2}{(\kappa+a)^3}e^{-(\kappa+a)(\hat{T}-T)}-m_5-\frac{T^3}{3}) \\
& - (\frac{2}{\kappa^3}e^{-\kappa(\hat{T}-T)}-\frac{1}{2}m_2-\frac{T^3}{3})-(\frac{2}{a^3}e^{-a(\hat{T}-T)}-\frac{1}{2}m_4-\frac{T^3}{3})]\} \\
,
\end{aligned}
\tag{27}
$$

where

$$
\begin{aligned}
m_1 &= \frac{1}{2\kappa}(T^2+\frac{T}{\kappa}+\frac{1}{2\kappa^2})e^{-2\kappa\hat{T}} \\
m_2 &= \frac{2}{\kappa}(T^2+\frac{2T}{\kappa}+\frac{2}{\kappa^2})e^{-\kappa\hat{T}} \\
m_3 &= \frac{1}{2a}(T^2+\frac{T}{a}+\frac{1}{2a^2})e^{-2a\hat{T}} \\
m_4 &= \frac{2}{a}(T^2+\frac{2T}{a}+\frac{2}{a^2})e^{-a\hat{T}} \\
m_5 &= \frac{1}{\kappa+a}(T^2+\frac{2T}{\kappa+a}+\frac{2}{(\kappa+a)^2})e^{-(\kappa+a)\hat{T}}
\end{aligned}
$$

Similar to the two-factor model, we use the closed-form solution of the geometric average Asian option as a control variate to price the corresponding arithmetic average Asian option. Table 2 lists the comparison of the results from standard Monte Carlo simulation with no variance reduction means, with antithetic method, and with control variate method. The results resemble what we observe in two-factor model, that control variate technique gives the most accurate performance, followed by antithetic method. It is easy to observe that most of the options prices lie in the range of \$2 - \$4. The standard deviations from standard Monte Carlo, antithetic method and control variate method are mostly about 2 cents, 1 cents, and one tenth of a cent, or 1%, 0.5% and 0.05% to the price. On average, control variate method produces the results of option prices with 23 times smaller standard errors than the standard Monte Carlo method, and 9 times smaller than the antithetic method.

3 Options Payoffs in Response to Parameter Changes

3.1 Two-Factor Model

Option prices change correspondingly to the parameters in the model. In this section we will show how and to what extent they change when the parameters in our model are manipulated. The price of the underlying asset and convenience yield follow equations (4) to (6). We manipulate the parameters in pairs to see how they jointly influence the price of an arithmetic average Asian option with 1 year to maturity, written on a futures contract of 2 year to maturity. The valuation is carried out by Monte Carlo simulation, with the analytical solution of the corresponding geometric average Asian option as a control variate for variance reduction purpose. 20,000 paths are generated for each pair of parameters. The interest rate is fixed at 5%. The basic setting of the values for all the parameters are exactly the same as in Table 1 and Table 2, unless being changed and indicated in the graph.

Figure 1 illustrate all the results for each pair of the parameters, and some interesting patterns can be easily identified here. For example, the first graph shows that the Asian option price seems to be rather sensitive to changes of either σ_1 or κ when the other parameter is relatively low. However, when either of them carries higher values, we can hardly see any price movement when we manipulate the value of the other. The first graph of Figure 1 also suggests that given a relatively lower speed of adjustment for the convenience yield (κ), a more volatile market reduce the price of Asian options, and a higher κ given a low level of σ_1 also reduces the price of the Asian option, implying that a more stable inventory level for the commodity lead to a lower option prices, given the relationship between the convenience yield and inventory level (Fama and French, 1988). Similar pattern can be observed in the second graph of Figure 1. When σ_2 is low, the option price fluctuates to a larger extent subject to the change of α. Also, higher volatility of the convenience yield bring the Asian options prices down. The joint influence of the two volatility parameters, σ_1 and σ_2, on the Asian option prices can be observed in the third graph, which reveals convex shape for both parameters. In the final graph of Figure 1, we can also observe a convex shape for λ_1, while option prices monotonically increase with σ_1.

3.2 Three-Factor Model

Under the three-factor model, the Asian option prices react to varying values of parameters in a different way. The price of the underlying asset, convenience yield and interest rate follow equations (15) to (18). Figure 2 shows how the prices change in response to the change of the parameters. In the first graph, α and m are long-term mean of the convenience yield and the interest rate respectively (unadjusted by market price of risk). It is easy to notice that when m is higher than 5%, the price movement becomes very subtle regardless of how α changes. This is plausible since a high discount factor would certainly lower the option price. On the other hand, the Asian option price rises monotonically with the increase of α, given a relatively low level of interest rate over the option period. This is obviously different from the results under two-factor model in Figure 1, where the option price forms a convex shape with the change of α. Hence, the introduction of the third stochastic factor apparently altered the way Asian options prices react to changes of the long-term mean of convenience yield.

The second graph of Figure 2 shows how options prices move subject to the different values of the speed of adjust for both convenience yield, κ, and interest rate, a, respectively. It is obvious to see that higher values of κ and lower values of a will result in higher options prices. This is also inconsistent with our results in two-factor model. While option prices rises dramatically with the decrease of κ given a low value of σ_1, the influence of κ on option prices seems to be the opposite, provided a relatively large σ_1. In Figure 5, the option price appear to increase monotonically with κ, and drops monotonically with a.

The third to fifth graph of Figure 2 are dedicated to show the relationship between the Asian option price and the volatility of the underlying spot price, σ_1, the convenience yield, σ_2, and the interest rate, σ_3. The third graph clearly shows that the option price form convex shape in response to the changes of σ_1 and σ_2. This is similar to the results from two-factor model in Figure 1. It is also apparent to see from the forth and fifth graphs that the option prices form a convex shape when the volatility of interest rate (σ_3) changes its value. However, in neither of the two graphs could we identify any observable changes in option prices when the volatility of the underlying spot (σ_1) and convenience yield (σ_2) are manipulated with different values, given a fixed σ_3. The last graph of Figure 2 demonstrates the response of option prices subject to changes in the market prices of both convenience yield risk, λ_1, and interest rate risk, λ_2. Option prices drop with λ_1, but rise with λ_2.

4 Jump Diffusion

In this section, a jump diffusion is added to the stochastic process of the underlying spot price. The concept of jumps in financial asset pricing starts from Merton (1976), where he questions the critical consumption of continuity and log-normality in the classic Black-Scholes option pricing model. Specifically, although the Black-Scholes model assumes that the movement of the spot price is log-normally distributed in a continuous time framework, there is strong evidence to suggest fat-tails in the real distribution of asset prices. In other words, extraordinarily large movements of the underlying price do happen in the real market. Hence, Merton's model allows these spikes to be captured by a jump diffusion, which, for example, could be caused by a demand or supply shock, or the arrival of some important news, leading to a sudden and profound effect on the underlying spot price. This is extremely important in pricing commodity related products for various reasons. First, although most of the commodities are consumed worldwide, they are particularly rich and thus produced on a large scale in only a handful of areas around the world. Hence, any information of new sources being discovered or existing plants found exhausted can send a shock to the price. Furthermore, due to the importance of the role commodities play in modern country development, they are usually prone to geopolitical conflicts between countries, which can cause dramatic price changes that no one can foresee.

It therefore makes sense to add the jump diffusion into our model, where the underlying price follows the stochastic process with a jump diffusion. We also decide to include jump diffusions only in the three-factor model. The rationale is the following. First, as in Schwartz (1997), the three factor model seems to outperform the two-factor model when the corresponding futures contracts carry a longer maturity, probably because the interest rate becomes more volatile in the long run. Thus, modelling the interest rate with a stochastic process for the futures contracts or Asian options with longer maturity (as in the three-factor model) fits the real market condition better than using a fixed rate (as in two-factor model), which might be more efficient in the short run. Meanwhile, jumps also occur more often in the long run, and hence their accumulated effects are more obvious and profound when the futures or Asian options carry relatively longer maturity. Accordingly, it is logical to include the jump diffusion in the three-factor model rather than in the two-factor model.

With a jump diffusion, the stochastic process for the underlying spot price satisfies the following equation under T-forward measure, $\mathbb{Q}^T$,

$$dS = (r - \delta - \theta\gamma_1 - \lambda_J^T \bar{\kappa}_J^T) S dt + \alpha_1 S dW_1^T + \beta_1 S dW_2^T + \gamma_1 S dW_3^T + \kappa_J^T dq^T \ , \quad (28)$$

where θ, α_1, β_1, γ_1 are identical to equation (A6) and equation (A2) in the Appendix. $\kappa_J^T dq^T$ represents the jump part, governed by two random variables. The first random variable, q^T, is the Poisson process, with λ_J^T as its intensity. Hence, the probability that a jump occurs per unit time is λ_J^T ($Prob(dq^T = 1) = \lambda_J^T dt$ when a jump occurs, and $Prob(dq^T = 0) = 1 - \lambda_J^T dt$ when there is no jump). The second random variable, κ_J^T, denotes the percentage jump size of the underlying spot price, conditional upon the occurrence of a jump. Then, $\kappa_J^T + 1$ follows a log-normal distribution, or $ln(\kappa_J^T + 1) \sim \mathcal{N}(ln(\bar{\kappa}_J^T + 1) - 0.5v_J^2, v_J^2)$. Hence, for most of the time when no jump occurs, $dq^* = 0$ and the stochastic process for the underlying price is analogous to the standard Brownian motion. When it does, the underlying price move abruptly by a random percentage. It is assumed that these jump random variables are pairwise uncorrelated with each other, and with the Brownian motion of the underlying price.

Under the three-factor model with jump diffusion, the stochastic process of the futures price under $\mathbb{Q}^T$ follows

$$dF = -(\lambda_J^T \bar{\kappa}_J^T - h_4 h_3) F dt + h_1 F dW_1^T + h_2 F dW_2^T + h_3 F dW_3^T + \kappa_J^T dq^T \ , \quad (29)$$

where

$$\begin{aligned} h_1 &= \sigma_1\alpha_1 \\ h_2 &= \sigma_1\beta_1 - \frac{\sigma_2\beta_2}{\kappa}(1 - e^{-\kappa(\hat{T}-t)}) \\ h_3 &= \sigma_1\gamma_1 - \frac{\sigma_2\gamma_2}{\kappa}(1 - e^{-\kappa(\hat{T}-t)}) + \frac{\sigma_3\gamma_3}{a}(1 - e^{-a(\hat{T}-t)}) \\ h_4 &= \ -\frac{\sigma_3}{a}(1 - e^{-a(T-t)}) \ . \end{aligned} \quad (30)$$

The introduction of jump diffusion adds certain complication to the pricing of Asian options, because there will be no closed-form solutions for either arithmetic average or geometric average Asian options. The reason is as follows. Although the random jump percentage size is log-normal and so is its product with the log-normal spot price, the jumping time is not. This is not a problem for pricing plain vanilla European options, because it is path independent. As a result, only the underlying price at maturity is relevant, and it is log-normal regardless of the jumping times. Nevertheless, Asian options are path dependent, and hence when the price jumps during the option period matters to its average over that period. A jump at the beginning of the option period clearly affects the options price differently than a jump near the maturity date. Since the jumping time is Poisson distributed, the average of the underlying spot price is no longer log-normal.

Nonetheless, we are able to find a unique analytical solution of geometric average Asian option, conditional on knowing when the jumps occur during the option period. This is because if we know when the jumps occur during the option period, we are left with only one random variable, the percentage jump size. It is log-normally distributed, and so is its product with the futures price. Hence, the geometric average of the futures prices also follows a log-normal distribution. Then, the solution of the geometric average Asian option can be used as a control variate to price arithmetic average Asian option by Monte Carlo simulation. We follow the same approach as in the three-factor model discussed earlier to find the solution, which is to find the mean and variance of the geometric average of the futures prices over the option period, namely E_{Jump} and V_{Jump}, respectively, under the T-forward measure. We use N_J and $T_{Ji}(i = 1, 2, ..., N_J)$ to denote the total number of jumps and the exact time of each jump, respectively, which are both assumed to be known. The closed-form solutions to the geometric average Asian option at time 0, given N_J and T_{Ji}, are presented in the following equation,

$$\begin{aligned} GA(0, T, \hat{T}) &= \mathbb{E}[e^{-rT} max(G(0, T, \hat{T}) - K, 0) \mid N_J, T_{Ji}] \ , \\ &= P(0, T)\mathbb{E}^{\mathbb{T}} max(G(0, T, \hat{T}) - K, 0 \mid N_J, T_{Ji}) \\ &= P(0, T)[e^{E_J + \frac{1}{2}V_J} N(\frac{E_J - ln(K) + V_J}{\sqrt{V_J}}) - KN(\frac{E_J - ln(K)}{\sqrt{V_J}})] \end{aligned} \tag{31}$$

where $G(0, T, \hat{T})$ follows equation (22).

The solution to the expectation, E_J, and variance, V_J, are given as follows,

$$E_J = \frac{1}{T}(E_1 + E_2), \qquad V_J = \frac{1}{T^2}(V_1 + V_2) \ , \tag{32}$$

where both E_2 and V_2 are related to and conditional on the jumps.

$$
\begin{aligned}
E_1 = & T ln(F_0) - \frac{1}{2}\bar{\kappa}_J^T \lambda^T T^2 \\
& - \frac{1}{2}[\frac{1}{2}\sigma_1^2 T^2(\alpha_1^2 + \beta_1^2 + \gamma_1^2) + \frac{\sigma_2^2}{\kappa^2}(\beta_2^2 + \gamma_2^2)(\frac{1}{2}T^2 - 2n_1 + n_3) \\
& + \frac{\sigma_3^2\gamma_3^2}{a^2}(\frac{1}{2}T^2 - 2n_2 + n_4) - \frac{2\sigma_1\sigma_2}{\kappa}(\beta_1\beta_2 + \gamma_1\gamma_2)(\frac{1}{2}T^2 - n_1) \\
& + \frac{2\sigma_1\sigma_3\gamma_1\gamma_3}{a}(\frac{1}{2}T^2 - n_2) - \frac{2\sigma_2\sigma_3\gamma_2\gamma_3}{\kappa a}(\frac{1}{2}T^2 - n_1 - n_2 + n_5)] \\
& - \frac{\sigma_1\sigma_3\gamma_1}{a}(\frac{1}{2}T^2 - n_6) + \frac{\sigma_2\sigma_3\gamma_2}{\kappa a}(\frac{1}{2}T^2 - n_1 - n_6 + n_7) \\
& - \frac{\sigma_3^2\gamma_3}{a^2}(\frac{1}{2}T^2 - n_6 - n_2 + n_8)
\end{aligned} \tag{33}
$$

$$
E_2 = (N_J T - \sum_{i=1}^{i=N_J} T_{Ji})(ln(\bar{\kappa}_J^T + 1) - \frac{1}{2}\upsilon_J) \tag{34}
$$

where

$$
\begin{aligned}
n_1 &= \frac{e^{-\kappa\hat{T}}}{\kappa}[\frac{1}{\kappa}(e^{\kappa T} - 1) - T] \\
n_2 &= \frac{e^{-a\hat{T}}}{a}[\frac{1}{a}(e^{aT} - 1) - T] \\
n_3 &= \frac{e^{-2\kappa\hat{T}}}{2\kappa}[\frac{1}{2\kappa}(e^{2\kappa T} - 1) - T] \\
n_4 &= \frac{e^{-2a\hat{T}}}{2a}[\frac{1}{2a}(e^{2aT} - 1) - T] \\
n_5 &= \frac{e^{-(\kappa+a)\hat{T}}}{\kappa + a}[\frac{1}{\kappa + a}(e^{(\kappa+a)T} - 1) - T] \\
n_6 &= \frac{e^{-aT}}{a}[\frac{1}{a}(e^{aT} - 1) - T] \\
n_7 &= \frac{e^{-aT-\kappa\hat{T}}}{\kappa + a}[\frac{1}{\kappa + a}(e^{(\kappa+a)T} - 1) - T] \\
n_8 &= \frac{e^{-aT-a\hat{T}}}{2a}[\frac{1}{2a}(e^{2aT} - 1) - T]
\end{aligned}
$$

and

$$
\begin{aligned}
V_1 = & \frac{1}{3}T^3\sigma_1(\alpha_1^2+\beta_1^2+\gamma_1^2) \\
& + \frac{\sigma_2^2(\beta_2^2+\gamma_2^2)}{\kappa^2}[\frac{1}{4\kappa^3}e^{-2\kappa(\hat{T}-T)} - \frac{4}{\kappa^3}e^{-\kappa(\hat{T}-T)} - m_1 + m_2 + \frac{T^3}{3}] \\
& + \frac{\sigma_3^2\gamma_3^2}{a^2}[\frac{1}{4a^3}e^{-2a(\hat{T}-T)} - \frac{4}{a^3}e^{-a(\hat{T}-T)} - m_3 + m_4 + \frac{T^3}{3}] \\
& + \frac{2\sigma_1\sigma_2(\beta_1\beta_2+\gamma_1\gamma_2)}{\kappa}[\frac{2}{\kappa^3}e^{-\kappa(\hat{T}-T)} - \frac{1}{2}m_2 - \frac{T^3}{3}] \\
& - \frac{2\sigma_1\sigma_3\gamma_1\gamma_3}{a}[\frac{2}{a^3}e^{-a(\hat{T}-T)} - \frac{1}{2}m_4 - \frac{T^3}{3}] \\
& - \frac{2\sigma_2\sigma_3\gamma_2\gamma_3}{\kappa a}[(\frac{2}{(\kappa+a)^3}e^{-(\kappa+a)(\hat{T}-T)} - m_5 - \frac{T^3}{3}) \\
& - (\frac{2}{\kappa^3}e^{-\kappa(\hat{T}-T)} - \frac{1}{2}m_2 - \frac{T^3}{3}) - (\frac{2}{a^3}e^{-a(\hat{T}-T)} - \frac{1}{2}m_4 - \frac{T^3}{3})]
\end{aligned} \tag{35}
$$

$$
V_2 = \sum_{i=1}^{i=N_J}(T-T_{Ji})^2\ v_J \tag{36}
$$

where

$$
\begin{aligned}
m_1 &= \frac{1}{2\kappa}(T^2+\frac{T}{\kappa}+\frac{1}{2\kappa^2})e^{-2\kappa\hat{T}} \\
m_2 &= \frac{2}{\kappa}(T^2+\frac{2T}{\kappa}+\frac{2}{\kappa^2})e^{-\kappa\hat{T}} \\
m_3 &= \frac{1}{2a}(T^2+\frac{T}{a}+\frac{1}{2a^2})e^{-2a\hat{T}} \\
m_4 &= \frac{2}{a}(T^2+\frac{2T}{a}+\frac{2}{a^2})e^{-a\hat{T}} \\
m_5 &= \frac{1}{\kappa+a}(T^2+\frac{2T}{\kappa+a}+\frac{2}{(\kappa+a)^2})e^{-(\kappa+a)\hat{T}}
\end{aligned}
$$

Note that V_1 is exactly the same as $V_{3-factor}$, the solution of variance to the three-factor model. It also implies that the solution is only intermediate to reach the final result of pricing arithmetic average Asian option, because it is based on one unique series of jumping times over the option period. The next step is to simulate a large number of different series of jumping times, each with a unique analytical solution to the geometric average Asian option as a control variate, leading to a unique numerical solution of arithmetic average Asian option. The final unbiased solution of the arithmetic average Asian options can be derived by averaging all the intermediate results.

Table 3 lists the results from the Monte Carlo simulation. By comparing the results and standard deviations from different methods, it is easy to see that control variate method outperforms the standard Monte Carlo simulation and antithetic method. This is consistent with our results without jump diffusions. It is also observable that the antithetic method hardly improves the simulation accuracy, since the standard errors are very close to those with standard Monte Carlo. Nevertheless, a comparison between Table 3 and Table 2 show that the control variate method can improve the simulation accuracy in both cases, but to a much more limited extent with jumps. This is mainly due to the fact that in the three-factor model without jumps, only one closed-form solution to the geometric average Asian option is generated and used as a control variate. However, in our jump model, for every specific series of jumping times, there is a unique analytical solution as a control variate. Since we need to simulate different series of jumping times, the number of control variates grow accordingly. As a result, the control variate method is effective and outperforms the standard Monte Carlo and the antithetic method, but to a limited extent with jump diffusion in the spot price.

5 Conclusion

We extend the Schwartz (1997) model of stochastic spot prices, convenience yield and interest rates to price Asian options. Given that analytical solutions are unattainable for arithmetic average Asian options, we use numerical methods such as the Monte Carlo simulation to price them. Furthermore, we obtain closed-form solutions of geometric average Asian options and use them as control variates to reduce simulation errors and thus improve the accuracy. The comparison of the results derived from the standard Monte Carlo simulation without any variance reduction methods, with the antithetic method and with the control variate methods show a significant improvements when control variates method is implemented. Next, we manipulate the parameters of the model to see how the options prices behave accordingly. The results show that option prices react to changes of parameters very differently from the two-factor model to the three-factor model. Then, we add a jump diffusion to the spot price process. Since Asian options are path-dependent, there appears to be no analytical solutions for either the arithmetic average or the geometric average Asian options. However, we find that conditional upon knowing when jumps occur over the option period, there is a unique closed-form solution to the geometric average Asian option. This solution is then used as a control variate to price arithmetic average Asian option numerically with Monte Carlo simulation. The result is intermediate, because it is conditional on a specific series of jumping times over the option period. We then simulate a large number of different series of jumping times and repeat the previous step, which finally leads to the unbiased final result by taking the average of all intermediate results. Our methods of control variate shows observable but relatively limited improvement in terms of accuracy over the standard Monte Carlo without any variance reduction methods and with antithetic method.

However, our paper is also not without limits, one of which is that the approach to price Asian option with Monte Carlo simulation is arguably outdated. A fairly large number of new methods have been introduced by different academics, such as Novikov and Kordzakhia (2014), Cai *et al* (2015) and Susai and Kyriakou (2016). We believe that further investigations should be conducted to compare different methods and see how ours differ from the others in terms of accuracy and efficiency. Furthermore, since Asian options, like most of the exotic options, are traded over-the-counters, lacking market data to calibrate our model remains to be a challenge. However, we hope that pricing Asian options using Monte Carlo simulation with control variate for variance reduction purpose can shed some light on further extension in the realm of commodity options pricing. A vast number of different financial instruments other than the usual futures, forwards and plain vanilla options have been invented for trading and hedging purposes. Exotic options, such as Asian options and lookback options, draw significant attentions for their unique feature that carry significant values for market participants to construct their investment portfolios and hedge against risks. Hence, we also appeal for further efforts to extend our research on wider range of exotic options under similar pricing models.

Appendix

Recall that the price of the geometric average Asian option at time t with maturity T written on a future contract with maturity $\hat{T}$ under the risk-neutral measure, $\mathbb{Q}^*$, is represented by equation (21),

$$GA(t,T,\hat{T}) = \mathbb{E}^* \; max[e^{-r(T-t)}(G(t,T,\hat{T}) - K, 0)] \; . \tag{21}$$

Since the interest rate is now a stochastic process that appears in both the discount factor and the underlying price, solving equation (21) requires a change of numeraire, where we transform the risk-neutral measure to the T-forward measure, $\mathbb{Q}^T$, where the zero-coupon bond is used as the new numeraire. Accordingly, equation (21) is equivalent to equation (23).

We first attempt to decompose the three correlated Brownian motions in equation (15) to (18) into three independent Brownian motions, namely dW_1^*, dW_2^*, dW_3^* under $\mathbb{Q}^*$. The result is shown as follows,

$$\begin{aligned} dZ_1^* &= \alpha_1 dW_1^* + \beta_1 dW_2^* + \gamma_1 dW_3^* \\ dZ_2^* &= \beta_2 dW_2^* + \gamma_2 dW_3^* \\ dZ_3^* &= \gamma_3 dW_3^* \end{aligned} \tag{A1}$$

where

$$\begin{aligned} \alpha_1 &= \sqrt{1 - \rho_3^2 - (\frac{\rho_1 - \rho_2\rho_3}{\sqrt{1-\rho_2^2}})^2} \\ \beta_1 &= \frac{\rho_1 - \rho_2\rho_3}{\sqrt{1-\rho_2^2}} \\ \beta_2 &= \sqrt{1-\rho_2^2} \\ \gamma_1 &= \rho_3 \\ \gamma_2 &= \rho_2 \\ \gamma_3 &= 1 \end{aligned} \tag{A2}$$

We then attempt to derive the corresponding Brownian motion under the T-forward measure $\mathbb{Q}^T$. In our model where the stochastic interest rate process is governed by equation (17), the price of the zero-doupon bond at time t with maturity T, $P(t,T)$, satisfies the following stochastic process under $\mathbb{Q}^*$,

$$dP(t,T) = r(t)P(t,T)dt - \sigma_3 B(t,T)P(t,T)dZ_3^* \ , \tag{A3}$$

where

$$B(t,T) = \frac{1 - e^{-a(T-t)}}{a} \ , \tag{A4}$$

Define the discount factor $D(t)$ at time t, then the discounted price of the zero-coupon bond can be represented by the following stochastic equation,

$$d(D(t)P(t,T)) = \theta D(t)P(t,T)dZ_3^* \ , \tag{A5}$$

where

$$\theta = -\sigma_3 B(t,T) \ , \tag{A6}$$

Hence, according to the rules of changing numeraire, the following process,

$$dW_3^T = \theta dt + dW_3^* \ , \tag{A7}$$

is a Brownian motion under $\mathbb{Q}^T$, and so are dW_1^*, dW_2^*, because they are independent of dW_3^* or the interest rate process, and will be denoted as dW_2^T, dW_3^T in the following context. Hence, the three stochastic processes in our model can now be shown as follows,

$$dS = (r - \delta - \theta\gamma_1)Sdt + \alpha_1 SdW_1^T + \beta_1 SdW_2^T + \gamma_1 SdW_3^T \ , \tag{A8}$$

$$d\delta = [\kappa(\hat{\alpha} - \delta) - \theta\gamma_2]dt + \beta_2 dW_2^T + \gamma_2 dW_3^T \ , \tag{A9}$$

$$dr = [a(\hat{m} - r) - \theta\gamma_3]dt + \gamma_3 dW_3^T \ , \tag{A10}$$

where

$$\hat{\alpha} = \alpha - \frac{\lambda_1}{\kappa} \ , \qquad \hat{m} = m - \frac{\lambda_2}{a} \ .$$

Under the T-forward measure, Equation (A8) to (A10) will then be used to calculate the price of the geometric average Asian option, represented by equation (23).

Bibliography

Agca, S. (2005) The Performance of Alternative Interest Rate Risk Measures and Immunization Strategies under a Heath-Jarrow-Morton Framework. *The Journal of Financial and Quantitative Analysis*, 40 (3), 645 - 669.

Bjerksund, P. (1991) Contigent Claims Evaluation When the Convenience Yield is Stochastic: Analytical Results. *Working Paper, Norwegian School of Economics and Business Administration.*

Boyle, P., Broadie, M. and Glasserman, P. (1997) Monte Carlo Methods for Security Pricing. *The Journal of Economic Dynamics and Control*, 21 (1997), 1267-1321.

Brennan, M.J. (1958) The Supply of Storage. *American Economic review*, 48 (March), 50-72.

Cai, N., Song, Y. and Kou, S. (2015) A General Framework for Pricing Asian Options Under Markov Processes. *Operations Research*, 63 (3), 540 - 554.

Casassus, J. and Collin-Dufresne, P. (2005) Stochastic Convenience Yield Implied from Commodity Futures and Interest Rates. *The Journal of Finance*, 60 (5), 2283-2331.

Fama, E.F. and French, K.R. (1987) Commodity Futures Prices: Some Evidence on Forecast Power, Premiums and the Theory of Storage. *The Journal of Business*, 60 (1), 55-73.

Fama, E.F. and French, K.R. (1988) Business Cycles and the Behavior of Metals Prices. *The Journal of Finance*, 43 (5), 1075-1093.

Frankel, J.A. (2006) The Effect of Monetary Policy on Real Commodity Prices. *Working Paper, National Bureau of Economic Research.*

Fusai, G. and Kyriakou, I. (2016) General Optimized Lower and Upper Bounds for Discrete and Continuous Arithmetic Asian Options. *Mathematics of Operations Research*, 41 (2), 531 - 559.

Gibson, R. and Schwartz, E. (1990) Stochastic Convenience Yield and the Pricing of Oil Contingent Claims. *The Journal of Finance*, 45 (3), 959-976.

Heath, D., Jarrow, R. and Morton, A. (1992) Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation. *Econometrica*, 60 (1), 77 - 105.

Heinkel, R., Howe, M.E. and Hughes, J.S. (1990) Commodity convenience yield as an Option Profit. *The Journal of Futures Markets*, 10 (5), 519-533.

Hilliard, J.E. and Reis, J. (1998) Valuation of Commodity Futures and Options under Stochastic convenience yield, Interest Rates, and Jump Diffusions in the Spot. *The Journal of Financial and Quantitative Analysis*, 33 (1), 61-86.

Inui, K. and Kijima, M. (1998) A Markovian Framework in Multi-Factor Heath-Jarrow-Morton Models. *The Journal of Financial and Quantitative Analysis*, 33 (3), 423 - 440.

Jamshidian, F. and Fein, M. (1990) Closed-Form Solutions for Oil Futures and European Options in the Gibson-Schwartz model: A Note. *Working Paper, Merrill Lynch Capital Markets*.

Jong, F. and Clara, P. (1999) The Dynamics of the Forward Interest Rate Curve: A Formulation with State Variables. *The Journal of Financial and Quantitative Analysis*, 34 (1), 131 - 157.

Kaldor, N. (1939) *Treatise on Money*. London: Macmillan.

Kemna, A.G.Z. and Vorst, A.C.F. (1990) A Pricing Method for Options Based on Average Asset Values. *The Journal of Banking and Finance*, 14 (1990), 113-129.

Merton, R.C. (1976) Option Pricing When Underlying Stock Returns Are Discontinuous. *The Journal of Financial Economics*, 3, 125-144.

Milonas, N.T. and Thomadakis, S.B. (1997) convenience yield as Call Options: an Empirical Analysis. *The Journal of Futures Markets*, 17 (1), 1-15.

Miltersen, K.R. and Schwartz, E.S. (1998) Pricing of Options on Commodity Futures with Stochastic Term Structures of convenience yield and Interest Rates. *The Journal of Financial and Quantitative Analysis*, 33 (1), 33-59.

Novikov, A. A. and Kordzakhia, N. E. (2014) Lower and Upper Bounds for Prices of Asian-Type Options. *Proceedings of the Steklov Institute of Mathematics*, 2014 (287), 234 - 241.

Routledge, B.R., Seppi, D.J. and Spatt, C.S. (2000) Equilibrium Forward Curves for Commodities. *The Journal of Finance*, 55 (3), 1297-1338.

Schwartz, E.S. (1997) The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging. *The Journal of Finance*, 52 (3), 923-973.

Telser, L.G. (1958) Futures Trading and the Storage of Cotton and Wheat. *The Journal of Political Economy*, 66, 233-255.

Vasicek, O. (1977) An Equilibrium Characterization of the Term Structure. *The Journal of Financial Economics*, 5 (1977), 177-188.

Working, H. (1948) Theory of te hInverse Carrying Charge in Futures Markets. *The Journal of Farm Economics*, 30, 1-28.

Table 1

Comparison of Monte Carlo Simulation Results*

(Two-Factor Model)

		Standard	Antithetic	Control Variate
α	0.1	2.306617 (0.025629)	2.289985 (0.013974)	2.271700 (0.001324)
	0.2	0.457325 (0.010928)	0.452497 (0.007369)	0.448185 (0.001082)
	0.3	0.046903 (0.003081)	0.046883 (0.002180)	0.046227 (0.000656)
	0.4	0.002167 (0.000558)	0.002028 (0.000431)	0.002377 (0.000257)
κ	1	3.845916 (0.033409)	3.886217 (0.014462)	3.887534 (0.001377)
	1.4	2.713750 (0.027067)	2.761657 (0.013609)	2.768098 (0.001212)
	1.8	2.205209 (0.024835)	2.254735 (0.013818)	2.267960 (0.001264)
	2.2	1.968732 (0.024265)	2.021045 (0.014214)	2.038581 (0.001384)
σ_1	0.3	2.478853 (0.021796)	2.486202 (0.009646)	2.461492 (0.000671)
	0.4	2.294381 (0.025333)	2.299143 (0.013970)	2.269943 (0.001298)
	0.5	2.393610 (0.030962)	2.398914 (0.018597)	2.372258 (0.002401)
	0.6	2.607746 (0.037780)	2.614441 (0.023671)	2.588267 (0.004024)

*The basic setting is as follows, unless indicated otherwise: $S_0 = 40$, $K = 40$, $\delta_0 = 0.2$, $\kappa = 1.8$, $\alpha = 0.1$, $\lambda_1 = 0.3$, $\rho = 0.8$, $\sigma_1 = 0.4$, $\sigma_2 = 0.5$. The interest rate is fixed at 5%. 20,000 paths are simulated. Standard deviations are put in brackets.

Table 2

Comparison of Monte Carlo Simulation Results*

(Three-Factor Model)

		Standard	Antithetic	Control Variate
a	0.8	2.510291 (0.023235)	2.493358 (0.010638)	2.489474 (0.001039)
	1.2	2.595693 (0.023585)	2.578850 (0.010587)	2.574607 (0.001041)
	2	2.707930 (0.024027)	2.691050 (0.010510)	2.686557 (0.001044)
	3	2.788461 (0.024334)	2.771541 (0.010448)	2.767069 (0.001046)
λ_2	0.15	4.814316 (0.031320)	4.843531 (0.009310)	4.836389 (0.001140)
	0.2	3.481873 (0.027644)	3.506576 (0.010758)	3.498777 (0.001114)
	0.25	2.374521 (0.023415)	2.396399 (0.011613)	2.386823 (0.001088)
	0.3	1.514725 (0.018927)	1.533245 (0.011079)	1.523934 (0.001057)
σ_3	0.01	3.469342 (0.027471)	3.478973 (0.010552)	3.498801 (0.001098)
	0.1	3.546264 (0.028072)	3.564702 (0.010790)	3.581590 (0.001149)
	0.3	4.263503 (0.033618)	4.304707 (0.012688)	4.313600 (0.001693)
	0.5	5.585277 (0.042904)	5.651460 (0.015322)	5.650241 (0.002837)

*The basic setting is as follows, unless indicated otherwise: $S_0 = 40$, $K = 30$, $\delta_0 = 0.2$, $r_0 = 0.05$, $\kappa = 1.8$, $\alpha = 0.1$, $a = 1.2$, $m = 0.05$, $\lambda_1 = 0.3$, $\lambda_2 = 0.2$, $\rho_1 = 0.8$, $\rho_2 = -0.01$, $\rho_3 = -0.02$, $\sigma_1 = 0.4$, $\sigma_2 = 0.5$, $\sigma_3 = 0.01$. 20,000 paths were simulated. Standard deviations are put in brackets.

Table 3

Comparison of Monte Carlo Simulation Results*

(Three-Factor Model with Jump Diffusion)

		Standard	Antithetic	Control Variate
λ_J	20	10.544340 (0.059832)	10.509634 (0.058269)	10.496171 (0.026282)
	40	10.397485 (0.075554)	10.378429 (0.074805)	10.269188 (0.031691)
	60	12.631421 (0.104631)	12.606723 (0.104180)	12.505597 (0.042692)
	80	13.902382 (0.126056)	13.884201 (0.125444)	13.585535 (0.053697)
	125	15.790436 (0.185479)	15.780606 (0.185122)	15.637354 (0.087090)
	175	16.611694 (0.259135)	16.599103 (0.258992)	16.408216 (0.144561)
	250	17.593515 (0.304078)	17.592879 (0.303905)	17.515748 (0.159361)

*The basic setting is as follows, unless indicated otherwise: $F_0 = 40$, $K = 40$, $r_0 = 0.01$, $\kappa = 1.3$, $a = 0.2$, $\rho_1 = 0.9$, $\rho_2 = -0.01$, $\rho_3 = -0.02$, $\sigma_1 = 0.3$, $\sigma_2 = 0.4$, $\sigma_3 = 0.01$, $\bar{\kappa}_J^T = 0.05$, $v_J = 0.01$. 40,000 paths were simulated. Standard deviations are put in brackets.

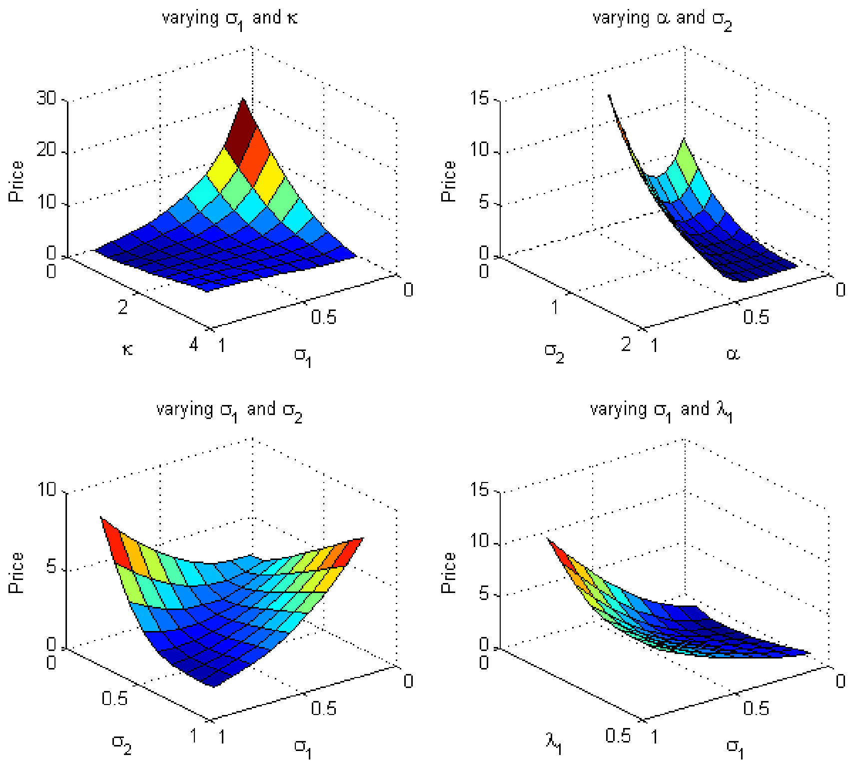

Figure 1: **Payoff of Asian Option, under varying parameters (two-factor model)**

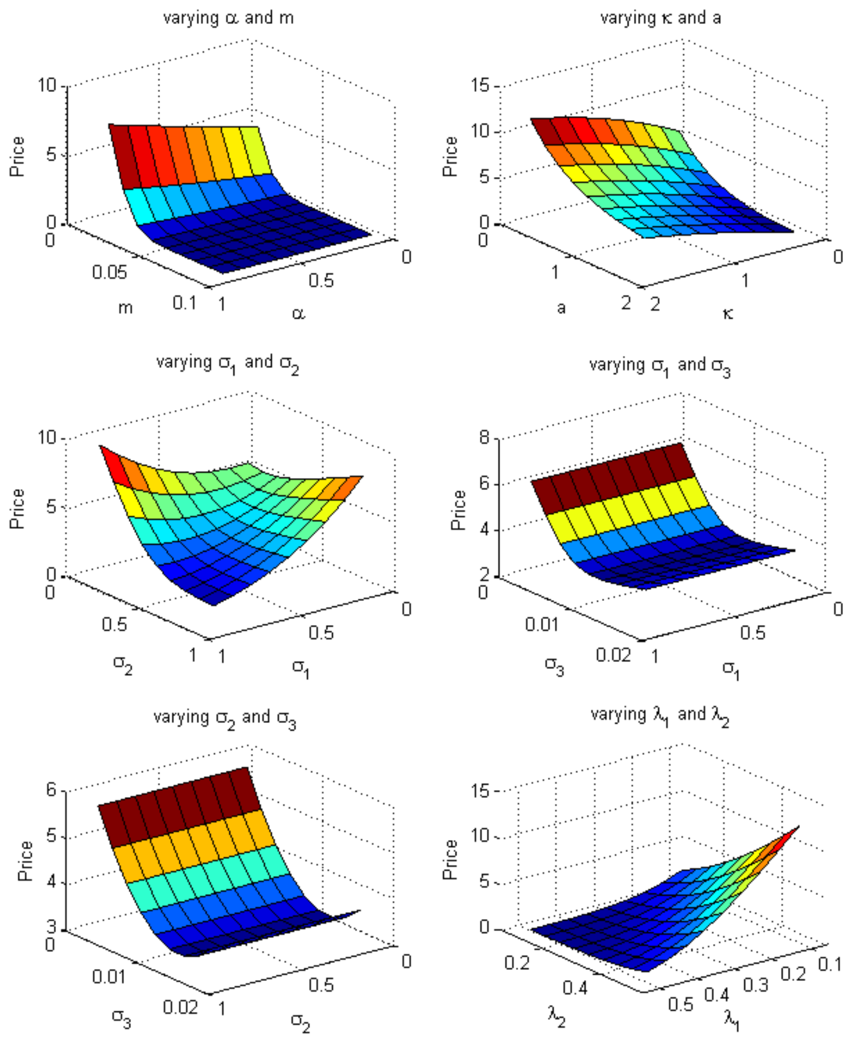

Figure 2: **Payoff of Asian Option, under varying parameters (three-factor model)**

Study of A New Seasonality Pattern in the Futures Prices of the Commodity Market

Abstract

The traditional idea of seasonality in the future market relates to the maturity date of a future contract. However, we find a new seasonal pattern in the futures prices that relates to the trading dates. We decide to explore such phenomenon in three energy commodity markets, namely, natural gas, gasoline, and crude oil. To conduct our initial empirical research, we design the so-called backward curve, as opposite to the forward curve, to visually illustrate the pattern of the trading-date seasonality. We find that when the prices of a collection of future contracts with the same maturity month can be averaged over the different years, the seasonality of trading dates is obvious to observe. We also find an interesting change of behaviour in the natural gas futures prices. Then, we conduct multiple statistical tests to further confirm our findings, which include the Kruskal-Wallis test, the autocorrelation test, and the power spectrum test. The results show strong evidence to support the existence of the trading-date seasonality.

1 introduction

Commodity trading has always been one of the major aspects in the financial market, which could, to a very large extent, help major commodity producers and consumers control their risk exposure to the market fluctuation. In light of this, understanding the behaviour of the commodity prices becomes increasingly important. Of all the factors that can influence the market prices of the commodities, seasonality could be a major one. However, its causes and effects differ for different types of commodities. For example, some agriculture products may show very strong seasonal pattern due to its supply side, because its price level in a year could largely been determined by whether it is the harvest season or not. On the other hand, the seasonality of some energy commodities usually occur thanks to the demand of its consumers, such as the increase of the natural gas price during the winter time when more is needed for heating purpose, or the fact that the gasoline price usually rises during the summer time when more people may choose to travel and thus consume more fuel for transportation.

A vast amount of researches has been dedicated to this specific realm of financial study. Early work include Samuelson (1965), which identifies what later known as the Samuelson effect, where volatility of a future contract with longer time till maturity tend to be lower than one closer to mature. Although this is not directly related to seasonality, it arguably starts the discussion in the academia regarding the price movement of a future contract over time during its active trading period. Sorensen (2002) confirms the Samuelson effect by the empirical study on agriculture products which carries significant seasonality impact, and further extends the Schwartz (1997) and Schwartz and Smith (2000) models by adding a deterministic seasonal factor, governed by a linear combination of trigonometric functions, and tests the new model with agriculture commodity futures prices. Lucia and Schwartz (2002) and Cartea and Figueroa (2005), among others, attempt to price the spot and forward prices in the electricity market, where seasonality plays a crucial role in determining the price. The paper of Lucia and Schwartz (2002) proposes a one-factor and a two-factor models with seasonal components, while a mean-reverting model with jump diffusion is introduced in Cartea and Figueroa (2005). The energy commodity market has also been largely studied. For example, Mirantes *et al* (2012, 2013) proposes several pricing models with seasonality as a stochastic factor. Borovkova and Geman (2006) specifically studies the seasonal pattern in the forward curves of the commodity prices. Furthermore, there are also papers such as Suenaga and Smith (2011) and Back *et al* (2013) that attempt to model seasonality in the volatility of the commodity prices. Shao *et al* (2015) develop a model to include time-varying and seasonal risk premiums in the US natural gas market.

In this paper, we focus on energy commodities and their prices. To begin with, the seasonal factor affects both their spot prices and futures prices. Given the brief description above, it is fairly straightforward to see how the spot price is affected. As for the futures prices, previous studies suggest that seasonality must be closely related to the maturity dates of the future contracts, since this is very close to the time when the product is delivered to the customers. As a result, although one can only speculate the price at a future delivery time, it is fairly reasonable to assume a higher or lower price if the maturity time of the contract coincides with the peaks or troughs of the prices in a year according to the seasonality pattern.

Nevertheless, there has been no theory to suggest a relation between seasonality and the trading dates. In other words, when a contract is traded should, in theory, have no influence on the price, with all other factors fixed. However, during our exploration of seasonality in some energy commodities, we find that the reality appears to contradict such conventional belief. We study the futures prices of Henry Hub natural gas, gasoline and generic crude oil, and all of them reveal evidence to suggest that the trading months actually influence the prices to a very observable extent, independent of the maturity dates. This even includes crude oil, which, in the past literature, is believed to carry no seasonality. In the case of crude oil, on average, the futures prices are always the highest when traded in July, and lowest in December, regardless of when the contract matures. On the other hand, the highest and lowest prices of the gasoline futures usually appear in July and December, respectively. In light of this, we create the backward curves, as opposed to the forward curves in the past literature, to further study the new seasonal pattern. When we aggregate all the prices of the future contracts that represent the delivery of one of the twelve months across our observation period and take average, the results show very clear and strong seasonal pattern. All of the findings are further tested by various statistical tests, which lead to some very promising results. We believe that this could be a very interesting and important contribution to the existing literature, and help understand the behaviour of commodity prices.

The rest of this paper is organized as follows. Section 2 briefly describes the raw data that we are going to use in this paper. Next, we discuss all the empirical results in section 3, including the preliminary findings of the data, and the introduction of the backward curves, which could significantly help visualize the seasonal pattern that relates to the trading dates in the futures prices. We also discuss a very interesting finding of the change of behaviour in the natural gas price during our observation period. Then, in section 4, we use multiple statistical tests, including the Kruskal Wallis test, the autocorrelation function, and the power spectrum to further confirm our finding of a new seasonality feature in the future prices. We draw our conclusions in Section 5, with limitations of this paper explained and appeal for further investigations.

2 The Data

The data we use consists of the daily futures prices of three energy commodities obtained from Bloomberg Terminal[1], including the Henry Hub natural gas, gasoline, and generic crude oil. Since we focus on seasonality, we decide to take the average of all daily observed prices over a month to derive the monthly price that we use in this paper. Our notation for the future contracts in most of the paper follows F1, F2, F3, ..., where F1 indicates the future contracts that mature in the next month, F2 in two months time, and so on. In the case of natural gas, we have the data from 1997 to 2017. However, from 1997 till 2002, only futures contracts with maturity dates expanding to the next 36 months from the trading dates (F1 $\sim$ F36) were traded. From 2002 till 2008, F37 $\sim$ F72 were added. Since 2008, the futures contracts with maturity dates up to 144 months from the trading dates (F73 $\sim$ F144) have become available. As a result, we divide the natural gas data into three groups accordingly. The first one includes all the futures prices of F1 $\sim$ F36 from April, 1997 till March, 2017. The second group is constituted by the futures contracts from January, 2002 to April, 2017, with maturity dates expanding to the next 60 months[2]. The third group includes the futures contracts from March, 2008 to October, 2017, with maturity dates up to 144 months. As for the contracts of gasoline and crude oil, the data we use are from March, 2007 and March, 2006 respectively to December, 2017, with maturity dates up to 36 and 60 months ahead. Table 1 shows the brief statistical summary of the data set, after taking monthly average of the raw daily observations.

[1] What we use for the prices is called "closing price one day ahead".

[2] we do not include F61 $\sim$ F72 due to a large number of missing data of these longer term contracts in the early years

3 The Empirical Results

In this section, we show the empirical evidence we gather from the analysis of our data to prove our findings of a new pattern of seasonality that is related to the trading dates.

3.1 Preliminary Findings

To demonstrate the preliminary findings of seasonality in our three observed energy commodities, Figure 1 to 3 show the average futures prices of natural gas, gasoline and crude oil with respect to the different maturity months and trading months. It is easy to see from the left part of Figure 1 that the futures prices of natural gas tend to be higher when maturing during the winter time such as December and January, and remains relatively low during the rest of the year. The opposite is true for gasoline, as shown in the left hand side of Figure 2, when the prices is higher in the summer time, and lower in the winter time. This is consistent with the previous research of both commodities. When it comes to the effects of the trading months, as demonstrated from the right hand side of Figure 1 and 2, both of natural gas futures and gasoline futures appear to reach their peak price of a year when being traded around the summer time in around June or July, and drop to their lowest when traded in the cold season, regardless of their maturity dates. The most interesting case here seems to be that of crude oil, which, according to the conventional theory, possesses no seasonality. This can easily be confirmed by the left part of Figure 3, which shows an almost completely flat line. It means that generally the different maturity months have nearly no impact on the futures prices of crude oil. However, the right hand side of Figure 3 clearly show that the crude oil futures share very similar seasonal pattern with natural gas and gasoline. The highest price of the contracts is usually reached when being traded in July, and the lowest price in December.

Also, in each year, there are twelve months when a contract can be traded or mature in. Hence, in the twelve months of each year, there is a trading month and a maturity month, respectively, with the highest price of the year, and one with the lowest price. We decide to count how many times a maturity or trading month has the highest or lowest price in each year over the observation period. The results, as illustrated in from Figure 4 to 8, seem to be less conclusive. We first look at the three groups of natural gas. The upper left graphs in Figure 4, 5, and 6 show that almost all of the maximum prices appear when the contracts mature in the winter time (December or January). However, the upper right graphs in these three figures suggest that the minimum prices in a year frequently occur in April as well as November, which is slightly inconsistent from the traditional theory that the lowest prices usually happen in the summer time. When it comes to the trading months, the evidence is less clear as no one particular month seem to have a dominant number to produce the highest prices of that year, as illustrated in the lower left graphs of Figure 4, 5, and 6. Nevertheless, it is quite obvious that contracts traded in January and December often have the lowest price of that year for all three groups. This is also slightly inconsistent with our observations in Figure 1, where the lowest average prices are usually found when being traded in February.

As for gasoline, the evidence is also mixed, as shown in Figure 7. From the upper left graph, it can be seen that for more than 60 times, a contract that mature in April has the highest price of that year. However, other than April, February, March, October and November, the rest of the maturity months seem to share similar proportion to have the maximum price of the year. The upper right graph, on the other hand, shows that the majority of the contracts that mature in December and January have the lowest price. As for the trading months, the two lower graphs in Figure 7 shows no discernible pattern across a year, as the maximum or minimum prices can occur in any month of a year.

The most interesting case is again the crude oil. The traditional idea believes that there should be no seasonality pattern for crude oil. This is further confirmed by our research, shown in the left graph of Figure 3, which turns out to be a complete flat line. However, it is very obvious to see in the two upper graphs of Figure 8 that, for some reason, the contracts that mature in winter months, especially December and January, consistently produce both the highest and lowest price of that year. As for the trading months, there seems to be no discernible pattern, as shown by the two lower graphs in Figure 8.

3.2 The Forward and Backward Curves

In this section, we adopt the idea of forward and backward curves to further investigate seasonality in our three commodities. First, the forward curve has been used for a long time in the past literature to describe the expected future price movements on the trading date, which can easily be drawn by using the different contracts that are traded on the same day (see Borovkova and Geman (2006)). The price shown on a forward curve indicates the expected price of a particular maturity month. Denoting $F(t_i, T_j)$ as the price of a future contract that is traded on t_i and matures on T_j, a forward curve illustrates the series of contracts $F(t^*, T_1), F(t^*, T_2), \ldots, F(t^*, T_N)$, where $j = 1, 2, \ldots, N$. Given the existence of maturity-date seasonality, the forward curve should present a wave and periodic pattern. Since most energy commodities carry annual seasonality, the pattern should repeat itself after every 12 months. Each of Figure 9 to 11 demonstrates four examples of forward curves of some randomly selected trading months for the three commodities. It is very easy to observe the seasonal pattern for natural gas and gasoline, and the lack of any seasonality for crude oil.

However, the forward curves fail to capture any seasonality that relates to the trading dates, since its x-axis represents the maturity dates. Hence, we introduce the so-called backward curve. The only technical difference between the forward and the backward curve is that the latter includes all the contracts that are traded on different dates but mature on the same date in one graph. As a result, each contract marked on the x-axis now indicates the different trading dates, prior to the maturity date. To be more specific, the series of contracts that appear on a backward curve can be denoted as $F(t_1, T^*), F(t_2, T^*), \ldots, F(t_N, T^*)$, where $i = 1, 2, \ldots, N$. As can be seen clearly, all the contracts in one backward curve mature in the same month, but they are traded i months before the maturity date. The larger the i is, the earlier the contract was traded. Therefore, following the x-axis, the curve illustrates a backward-looking view of the futures prices of contracts that are traded on different months and have the same maturity date. As a result, if there is a seasonality that relates to the trading date, it should be captured and shown in the backward curve.

Nevertheless, in the case of the backward curves, the trading-date seasonal effect seems very hard to detect, as shown in Figure 12 to 14. Each of the figures includes four backward curves of some randomly chosen maturity months, and none of the twelve backward curves show any obvious seasonal pattern. Therefore, although our preliminary findings of Figure 1 to 3 suggest the existence of seasonality that relates to the trading dates independent of the maturity dates, the backward curves fail to provide any discernible evidence.

Nonetheless, despite the failure, we decide to take a further look at the forward and backward curve. According to our knowledge, in the past literature, the forward curves are usually based on a single specific day, or a specific month, as shown in Figure 9 to 11. By the same definition, our backward curves are also based on a single maturity months. We wonder, however, if the effect of trading-date seasonality would be more conspicuous, when we plot the backward curves based on the average price of a collection of the contracts that are going to mature in one of the twelve months of a year over our entire observation periods. The same procedure is also applied to obtain the new forward curves.

To explain how to obtain the new forward curves as an example, let us change our notation of the future price to $F_{t_i}(K)$, where t_i here indicates the month t of year i when the contract is traded. To be more specific, $t = 1$ for January, $t = 2$ for February and so on, and $i = 1$ for the first year in our observation period, $i = 2$ for the second year and so on. K here indicates the lengths of time till maturity. We take the average of the prices of all the contracts that are traded in month t over the observation period to plot the new forward curve. The x-axis is the series of K, or time till maturity. Each point on the new forward curve of month t now represent the averaged futures prices with the maturity dates $t + K$ over the years from $i = 1$ to N, which is calculated by $\frac{\sum_{i=1}^{N} F_{t_i}(K)}{N}$, where N here represents the latest year in our observation period. A similar procedure is implemented on the new backward curves, with the only difference being to gather all the contracts that mature in the same month. If the future price can be marked as $F_{T_i}(K)$, where T_i here represents the maturity month T of year i. Then, each point on the new backward curve of month T are calculated by $\frac{\sum_{i=1}^{N} F_{T_i}(K)}{N}$. The x-axis represents the trading date that is K months before the maturity date.

Figure 15 to 19 show the new forward curves of all three commodities, including the three groups of natural gas. Each figure has 12 graphs, indicating the 12 months in a year. It is very easy to see the seasonal pattern for all three groups of natural gas as well as gasoline, of which the wave pattern repeats for every 12 months. A closer look at each graph shall prove that the peak and bottom of the price during a calender year is consistent with Figure 1 and 2. As for crude oil, its new forward curve shows almost no fluctuation, which further proves the conventional idea that crude oil possesses no seasonal feature. An interesting finding can also be observed in Figure 17 of natural gas group 3, where there is obviously a trend of prices rising with respect to time till maturity. This is consistent with a phenomenon of contango, where futures prices exceeds the expected spot price in the future. On average, since the futures prices must converge to the spot price at maturity, they would drop over time when the contracts approach maturity. However, such trend does not seem to exist in group 1 or 2 of natural gas. We will discuss about this in a separate subsection below. In summary, the new forward curve of collective months provides consistent evidence with the existing literature about seasonality for natural gas and gasoline, and lack of seasonality for crude oil.

We then move on to the new backward curves, which is exhibited from Figure 20 to 24. It is very clear to see strong graphical evidence of seasonality that is related to the trading dates in most of the figures, which includes crude oil that is previously deemed as a non-seasonal commodity. Hence, crude oil, just like the other two commodities, also features seasonality in its prices, but is subject to the trading dates, instead of the maturity dates. In all the cases, the seasonal pattern repeats itself on an annual basis. This is obviously in contrast to what we observe in the old backward curves in Figure 12 to 14. As a result, although the trading-date seasonality effect may be very hard to find in the old backward curves of a specific maturity month, it is very obvious to be observed in the new backward curves. We also speculate if this is why the trading-date seasonality seems to be largely ignored by the previous researches. Moreover, only Figure 22 of natural gas group 3 shows evidence that is consistent with contango, which is similar to the forward curves in Figure 17 above. We will discuss this phenomenon in detail in the next subsection.

We also attempt to find potential explanations behind the trading-seasonality, which turns out to be less fruitful than expected. First of all, as far as our knowledge extends, no existing literature appear to focus on the relation between seasonality in prices and the trading dates in any commodity market. Samuelson (1965) is one of the first papers to discuss the effect of time till maturity on the futures volatility and price. Although the Samuelson effect does no indicate strong seasonal pattern during the futures life time, it does identify the volatility changes over the course of the future contract. Recent papers such as Suenaga and Smith (2011) and Back *et al* (2013) model seasonal volatility in their study, and Shao *et al* (2015) develop models with seasonal risk premium. However, neither of their papers indicate that the trading date of a contract should have any influence on the volatility or the futures prices. We believe that the relation between volatility, risk premium, or other factors and the future prices could still hold valuable secrets to be discovered to explain what we observe in this paper.

We then decide to speculate the possible reasons in reality. The first one that we can think of is related to the major commodity producers or consumers and their timing of establishing their portfolios of financial products to hedge their existing positions. Given their enormous scales, if a group of major market participants such as major oil producers or airlines decide to build their financial portfolios at the same time of a year, the volume of the financial contracts to build such portfolios could increase of decrease the price of the underlying commodity to a very large extent. It is very hard to simply claim a mere coincidence that for all three of our energy commodities, their prices tend to be higher during the summer time. We speculate that maybe it is because the large companies mostly decide to build their portfolios during such time when trading is more active and thus liquidity is abundant. Nevertheless, it is extremely difficult to prove such claim, due to the difficulty to obtain relevant information from any of these companies at this stage. Although the data of the market trading volume might, in theory, reveal some traces of their trading activities during a year, we are unable to obtain consistent and long-term data on trading volume.

3.3 From Backwardation to Contango for Natural Gas Price during 1997 and 2017

This subsection is dedicated to discuss what we observe as a change of behaviour of the future price verses the time till maturity of the three natural gas groups in the above subsection. It is easy to see that both the forward and backward curves of natural gas group 1 and group 2 in Figure 15, 16, 20, 21 show no discernible trend of price against time to maturity. Nonetheless, Figure 17 and 22 of natural gas group 3 reveals an obvious trend of increase of future price with longer maturity term.

Since one of the two major differences among the three groups of natural gas is the observation period[3], we decide to look deeper into group 1 and 2, by breaking both of them into several sub-groups, and then compare them to group 3. Group 1 is divided into three sub-groups: group 1.1 involves the observation period of 1997 $\sim$ 2002, group 1.2 of 2002 $\sim$ 2008 and group 1.3 of 2008 $\sim$ 2017. Group 2 is broken into two sub-groups: group 2.1 covers the observation period of 2002 $\sim$ 2008, and group 2.2 of 2008 till 2017. The reason for dividing the groups in such a way is because of the time when new contracts are introduced into the market to trade. As we have mentioned in the Data section, F37 $\sim$ F72 started to trade only after 2002, and F73 $\sim$ F144 only after 2008. Hence, we speculate if the introduction of new contracts may have any effects on the behaviour of prices, albeit no confirmed theory to prove such relation.

Figure 25 to 34 demonstrate all the new forward and backward curves of the 5 sub-groups. We would like to mention 2 points from our observations. First, Figure

[3]The other difference is the maximum length of time to maturity, which should play no role here.

28 of the new backward curves of sub-group 1.1 seems to show almost no identifiable seasonality, as oppose to Figure 25, where the seasonal effects in the new forward curve is very clear during the same observation period. Also, in contrast to almost no trading-day seasonality in Figure 28 for natural gas contracts of F1 $\sim$ F36 during 1997 and 2002, Figure 29, Figure 30, Figure 33 and Figure 34 that covers more recent observation periods reveal clear seasonal pattern that relates to the trading dates. Hence, the seasonal effect of trading dates may be more significant only in recent years after 2002.

The second, and arguably more interesting, observation is that before 2008, there seems to be either no apparent trend of natural gas prices verses time till maturity, or a slight decreasing one, as can be seen from both the forward and backward curves of Figure 25, 26, 28, 29, 31 and 33. In other words, there are either moderate signs that are consistent with normal backwardation, or neither backwardation nor contango in the natural gas market before 2008. However, after that, natural gas prices seem to rise significantly with the maturity term, as shown in Figure 27, 30, 32, and 34. In other words, after 2008, market participants seem to give higher premium to futures contracts traded with longer maturity term. Therefore, when we combine all the sub-groups together into three groups to cover the entire observation period as we have done in the previous section, the joined result show almost no signs of any trend in the forward or backward curves (Figure 15, 16, 20 and 21 of group 1 and 2), but very obvious rising trend after 2008 (Figure 17 and 22 of group 3). Unfortunately, the reason behind such change of the natural gas price behaviour appears hard to ascertain and beyond the scope of this paper. Although our research shows that the change may happen in around 2008 when the longer term future contracts were introduced, we fail to find any solid evidence or theory to suggest a strong relationship between them. We appeal for further investigations into this rather interesting phenomenon.

4 The Statistical Tests

In this section, we use multiple statistical tests to further confirm our findings of seasonality that is related to the trading dates of the futures contracts in all three commodities. Although the graphical evidence in the previous sections seems strong to suggest the new seasonality exists in the energy future market, we would like to conduct further statistical tests to consolidate previous findings. The tests include the Kruskal-Wallis test, the autocorrelation tests, and the power spectrum tests.

We adopt multiple tests to have a comprehensive analysis, because each of them involves different techniques and tests the existence of seasonality from a different perspective. Also, in this section, we will mostly omit the forward curves, for the reason that the maturity-related seasonality, shown by the forward curves, has been well confirmed and studied in the past literature. Hence, we will focus on the maturity-date seasonality, and the backward curves in this section. Furthermore, we adopt the autocorrelation and power spectrum tests on both the old backward curves of some randomly selected dates where the seasonality pattern is hard to observe, and the new backward curves derived from the averaged prices of a collection of contracts over our entire observation period.

First, we perform the Kruskal-Wallis test, which is a non-parametric one-way analysis of variance (ANOVA) for testing two or more groups of samples of data if they follow the same distribution. Since we have 12 months in a year, our samples will be in 12 groups, hence a degree of freedom of 11. The null hypothesis is that there is no seasonality related to trading date in our sample. The results of the Kruskal-Wallis test are listed in Table 2, which clearly show the rejection of null for natural gas group 2 and 3, gasoline and crude oil. However, we fail to reject the null for natural gas group 1. Since the Kruskal-Wallis test is adopted directly on the monthly data averaged from the raw daily observations of prices without any further manipulation of the data, the conclusion here largely confirms our findings that there exists a trading-date seasonal pattern in natural gas (group 2 and 3), gasoline and crude oil.

Next, one of the most popular tests for seasonality is the autocorrelation function, which is to find if there is a pattern in the time series of data that repeat itself by various time lags. To perform the autocorrelation test, we detrend all the data first. Figure 35 to 37 show the results of the autocorrelation functions on the old backward curve of some randomly selected maturity dates. In all cases of natural gas, there is almost no sign in the autocorrelation curves to suggest any annual seasonality. When it comes to gasoline and crude oil, although some autocorrelation curves show some periodic evidence, none of them corresponds to a 12-month time lag. Nevertheless, when the new backward curves are tested (Figure 38 to 42), the results are completely different, where figures of natural gas group 2 and 3 (Figure 39 and 40), gasoline (Figure 41) and crude oil (Figure 42) show very strong signs of annual seasonality that is consistent with our previous findings. In figure 38 of natural gas group 1, although not every graph produces statistically significant evidence to support seasonality, almost all of them show signs of periodic repetition to a certain degree.

The third technique we use is the power spectrum test, which is based on the Fourier transformation (see Wei (2006), Chapter 12, for detailed explanations.), of which the results are shown in Figure 43 to 53[4]. Each point on the spectrum distribution graph indicates the power with respect to its frequency. First, we calculate and plot the power spectrum of the old backward curves of four randomly selected maturity months in Figure 43 to 45. Although almost all the figures show a peak in the spectral curve, very few of them coincide with the frequency of 1. In other words, though there seems to be evidence of some periodic pattern hidden in the data, it almost never repeats itself after one calender year. As a matter of fact, many of the peaks appear with a frequency of lower than 1, meaning that the periodic pattern has a longer period than 1 year. In contrast, we also plot the power spectrum of the old forward curves of the same four months to compare in Figure 46 to 48. In both cases of natural gas and gasoline, there are obvious peaks in the spectral curve that coincide with the frequency of 1, indicating a strong evidence of annual maturity-date seasonality in the data. Also, there is no peak in Figure 48 of crude oil, since it carries no seasonality that would show in the forward curve. Nevertheless, the power spectrum tests of the new backward curves appear to tell a completely different story, as illustrated in Figure 49 to 53. All of the graphs in each of the figures provide very strong evidence to suggest the existence of seasonality in the new backward curve, which proves the existence of seasonality that relates to the trading dates. The conclusion here is again the same with the autocorrelation function, that when it comes to the old backward curves, we can hardly find any evidence to suggest the existence of any seasonal pattern in the data. However, when we further investigate the new backward curves, there are very strong evidence to prove that seasonality related to the trading dates exists.

5 Conclusion

In this paper, we present our findings of a new seasonal pattern that relates to the trading dates in three commodities, namely, natural gas, gasoline, and crude oil. We discover that the trading date of a future contract appears to have rather significant influence on its price level. This is especially interesting in the case of crude oil, which is believed in the previous literature to have no seasonality in its prices. The preliminary findings suggest that the peaks and troughs of prices during a calender

[4]The data is not detrended for the power spectrum test.

year differ slightly among different commodities. For example, the highest prices being traded usually appear in June for natural gas, and July for both gasoline and crude oil. Then, we introduce the backward curves. After some data manipulation, it provides strong evidence of seasonality when we collect and take average of the futures prices over our observation period. Next, we adopt multiple statistical tests to confirm the existence of the trading-day related seasonality in the three commodities. The Kruskal-Wallis test provide solid evidence to suggest the existence of the trading-date seasonality. Both the autocorrelation tests and power spectrum tests fail to confirm the trading-date seasonal effect on the old backward curves, but show strong support of the seasonal pattern in the new backward curves. We conclude that the trading-date seasonality may be hard to observe on an individual old backward curve, but does exist and can easily be identified on a collection of contracts that mature in the same month over a long period of time.

However, in light of our new discovery in the commodity market, we appeal for further investigations. For example, the cause of the new seasonality that relates to the trading dates remains unidentified. It is rather straightforward to explain the conventional seasonal pattern that relates to the maturity dates, since the time that the futures contracts mature is also exactly, or very close to, the time when the commodity is harvested or mined, produced, delivered and consumed. Hence, some simple knowledge of supply or demand is enough to explain such seasonality. Nevertheless, the same knowledge fails to explain why the trading dates have any influence on the futures prices. We speculate that it may be related to the time when some large market participants such as the major commodity producers or traders build their portfolio for the purpose of hedging, investing or speculating. Although the evidence to support such claim remains to be found, and the data needed to conduct the research could be very hard to obtain, the journey to unravel such mystery of the commodity futures prices will not end here.

6 References

Back, J., Prokopczuk, M. and Rudolf, M. (2013) Seasonality and the Valuation of Commodity Options. *Journal of Banking and Finance*, 37 (2), 273 - 290.

Borovkova, S. and Geman, H. (2006) Seasonal and Stochastic Effects in Commodity Forward Curves. *Review of Derivatives Research*, 9 (2), 167 - 186.

Cartea, A. and Figueroa, M.G. (2005) Pricing in Electricity Markets: A Mean Reverting Jump Diffusion Model with Seasonality. *Applied Mathematical Finance*, 12 (4), 313 - 335.

Lucia, J.J. and Schwartz, E.S. (2002) Electricity Prices and Power Derivatives: Evidence from the Nordic Power Exchange. *Review of Derivatives Research*, 5 (1), 5 - 50.

Mirantes, A.G., Poblacion, J., and Serna, G. (2012) The Stochastic Seasonal Behaviour of Natural Gas Prices. *European Financial Management*, 18 (3), 410 - 443.

Mirantes, A.G., Poblacion, J., and Serna, G. (2013) The Stochastic Seasonal Behavior of Energy Commodity Convenience Yields. *Energy Economics*, 40, 155 - 166.

Samuelson, P. A. (1965) Proof That Properly Anticipated Prices Fluctuate Randomly. *Industrial Management Review*, 6 (2), 41 - 49.

Schwartz, E.S. (1997) The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging. *The Journal of Finance*, 52 (3), 923 - 973.

Schwartz, E. and Smith, J. (2000) Short-Term Variations and Long-Term Dynamics in Commodity Prices. *Management Science*, 46 (7), 893 - 911.

Shao, C., Bhar, R. and Colwell, D.B. (2015) A Multi-Factor Model with Time-Varying and Seasonality Risk Premium for the Natural Gas Market. Energy Economics, 50 (2015), 207 - 214.

Sorensen, C. (2002) Modeling Seasonality in Agricultural Commodity Futures. *Journal of Futures Markets*, 22 (5), 393 - 426.

Wei, W.S. (2006) *Time Series Analysis: Univariate and Multivariate Methods*, 2nd edn. Pearson Education, Boston.

Table 1

Average Prices and Standard Deviations of Selected Contracts

Commodity	F1	F4	F10	F18	F30	F42	F60	F90	F120
Natural Gas, Group 1	3.1393	3.1702	3.1201	3.0475	2.9899				
(4/1997 ∼ 3/2017)	(1.3457)	(1.0512)	(0.9712)	(0.8127)	(0.7316)				
Natural Gas, Group 2	6.7186	7.1273	7.2300	7.0764	6.7427	6.4944	6.1947		
(1/2002 ∼ 4/2017)	(2.3399)	(2.4185)	(2.3261)	(2.3360)	(2.1826)	(2.0572)	(1.1770)		
Natural Gas, Group 3	3.9999	4.2504	4.5225	4.6899	4.8147	4.9239	5.1037	5.5636	6.0069
(3/2008 ∼ 10/2017)	(1.8666)	(1.9104)	(1.8871)	(1.9102)	(1.8625)	(1.8330)	(1.7689)	(1.8555)	(1.8692)
Gasoline	221.3155	219.4478	217.5467	215.4578	213.4606				
(3/2007 ∼ 12/2017)	(62.9581)	(58.8160)	(55.2958)	(49.2234)	(45.0019)				
Crude Oil	81.8602	83.1176	83.7683	83.5904	83.0481	82.7449	82.8280		
(3/2006 ∼ 12/2017)	(20.4189)	(19.0312)	(17.2767)	(15.6588)	(14.3093)	(13.7424)	(13.5869)		

* Numbers in the brackets refer to observation period in the first column, and standard deviation in the rest columns.

Table 2

Results of Kruskal-Wallis test

Commodity	Chi-Square (P-value)	Sample Size (N)
Natural Gas, Group 1	6.8572 (0.8105)	8640
Natural Gas, Group 2	26.938 (0.004695)**	10800
Natural Gas, Group 3	282.98 ($<$ 2.2E-16)***	15552
Gasoline	107.91 ($<$ 2.2E-16)***	4320
Crude Oil	96.014 (1.10E-15)***	7920

a: Degree of Freedom: 11

b: 10% significance *, 5% significance **, 1% significance *** .

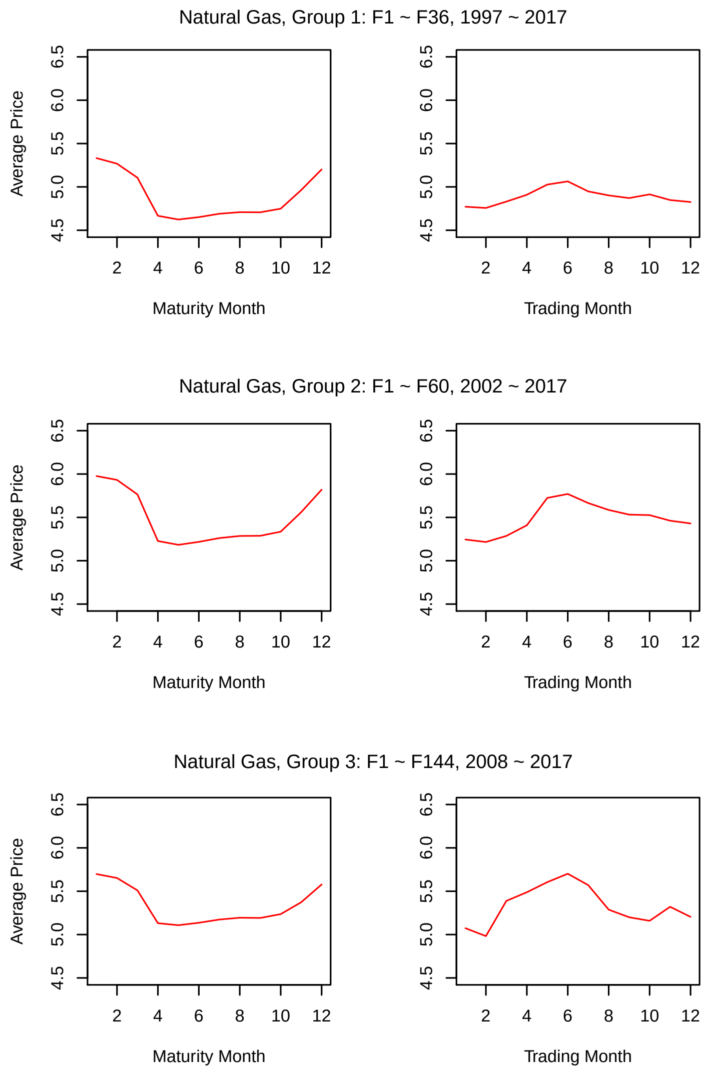

Figure 1: **Average Natural Gas Monthly Price Vs. Maturity Months and Trading Months**

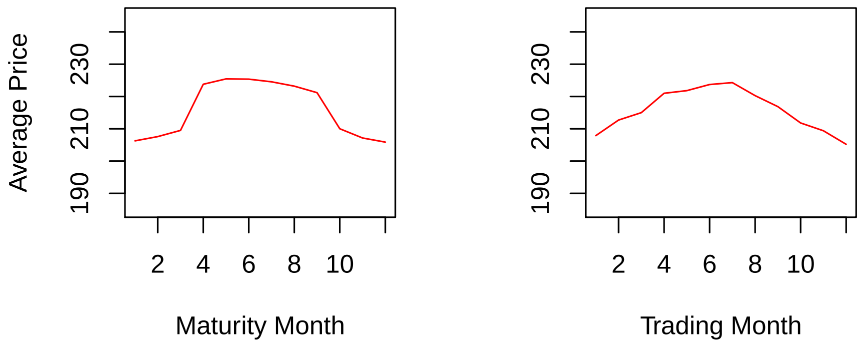

Figure 2: **Average Gasoline Monthly Price Vs. Maturity Months and Trading Months (F1 $\sim$ F36, 2007 $\sim$ 2017)**

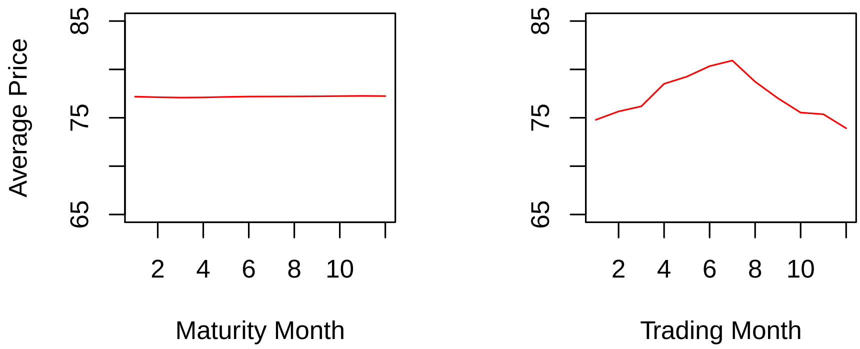

Figure 3: **Average Crude Oil Monthly Price Vs. Maturity Months and Trading Months (F1 $\sim$ F60, 2006 $\sim$ 2017)**

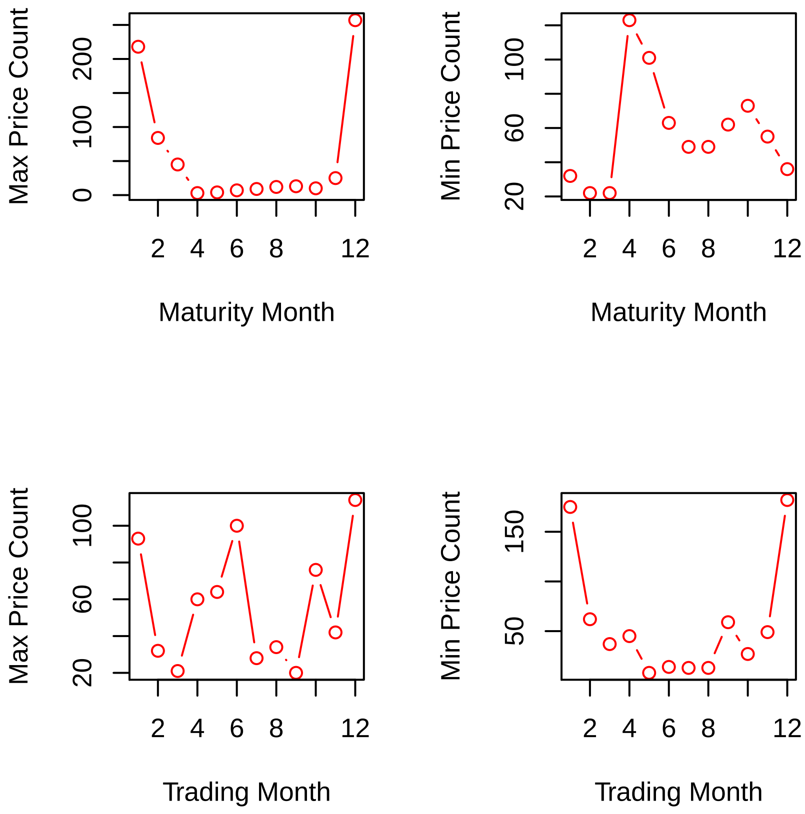

Figure 4: **No. of Counts w.r.t. maturity and trading months, Natural Gas, Group 1 (F1 $\sim$ F36, 1997 $\sim$ 2017)**

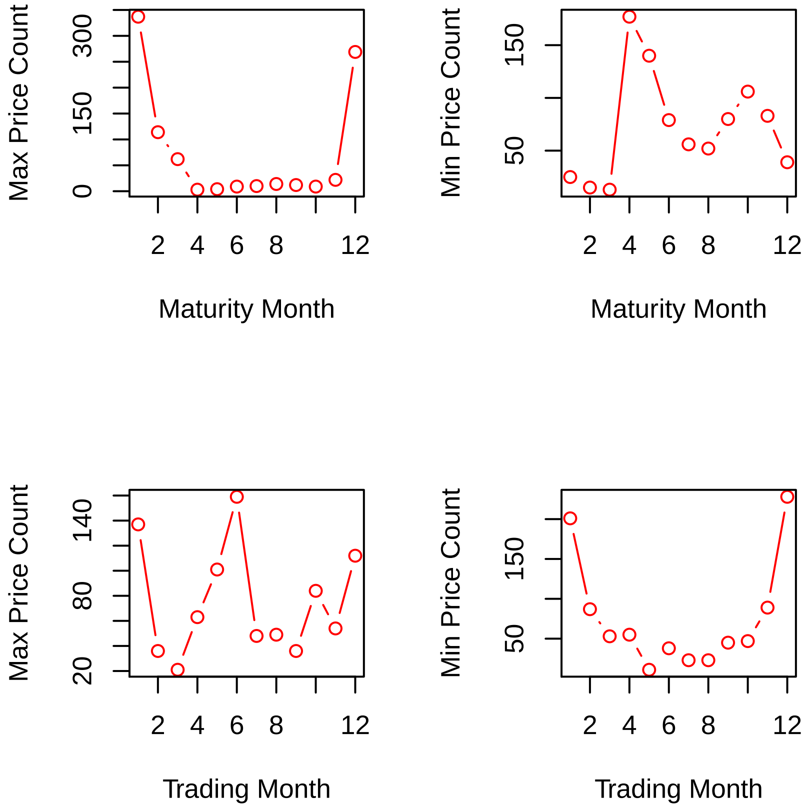

Figure 5: **No. of Counts w.r.t. maturity and trading months, Natural Gas, Group 2 (F1 $\sim$ F60, 2002 $\sim$ 2017)**

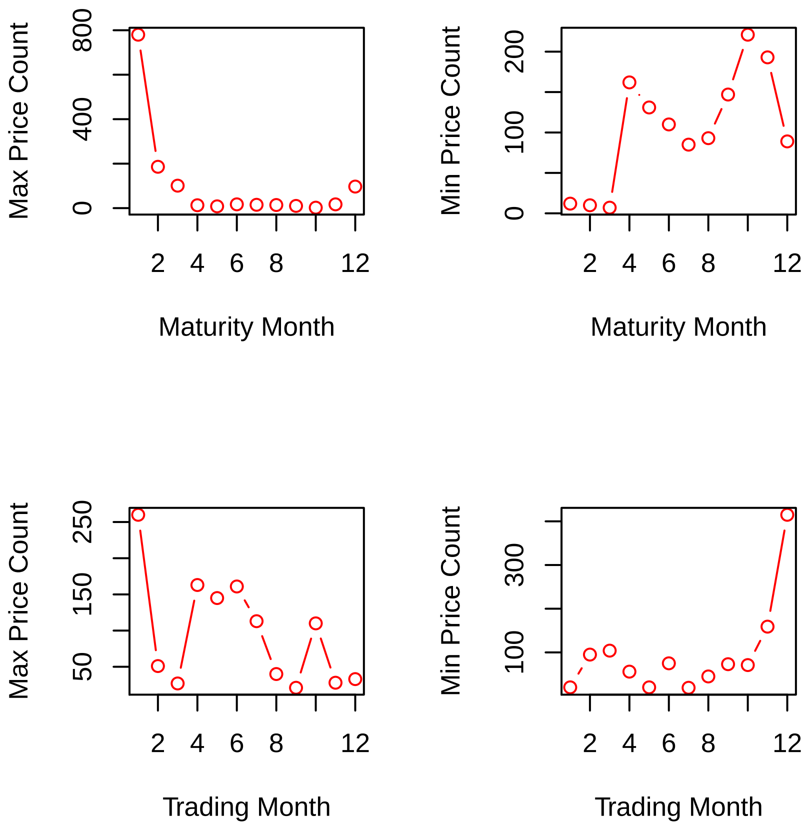

Figure 6: **No. of Counts w.r.t. maturity and trading months, Natural Gas, Group 3 (F1 $\sim$ F144, 2008 $\sim$ 2017)**

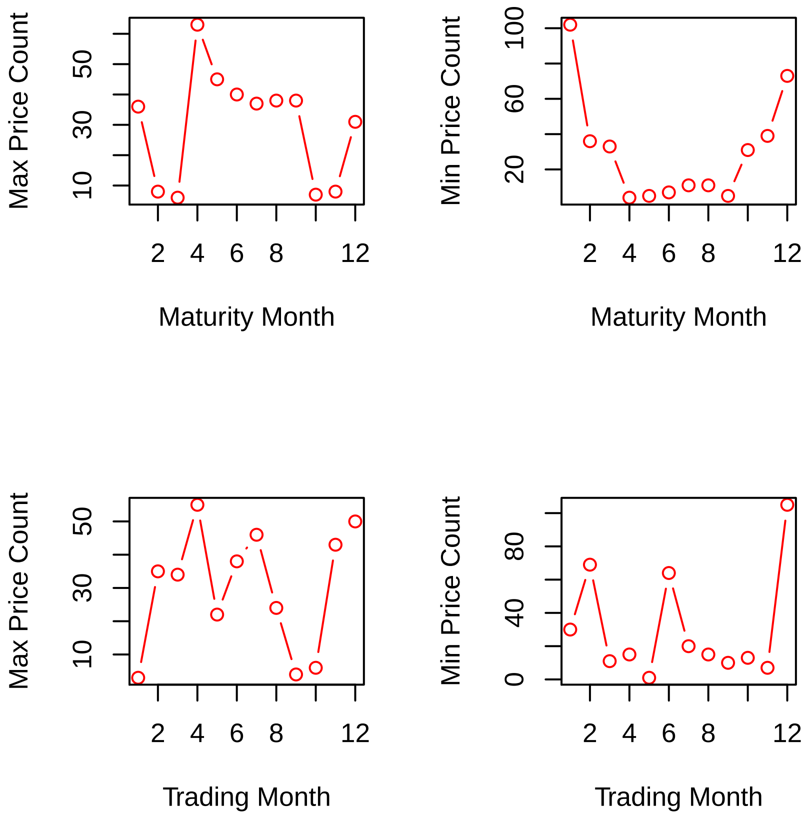

Figure 7: **No. of Counts w.r.t. maturity and trading months, Gasoline (F1 $\sim$ F36, 2007 $\sim$ 2017)**

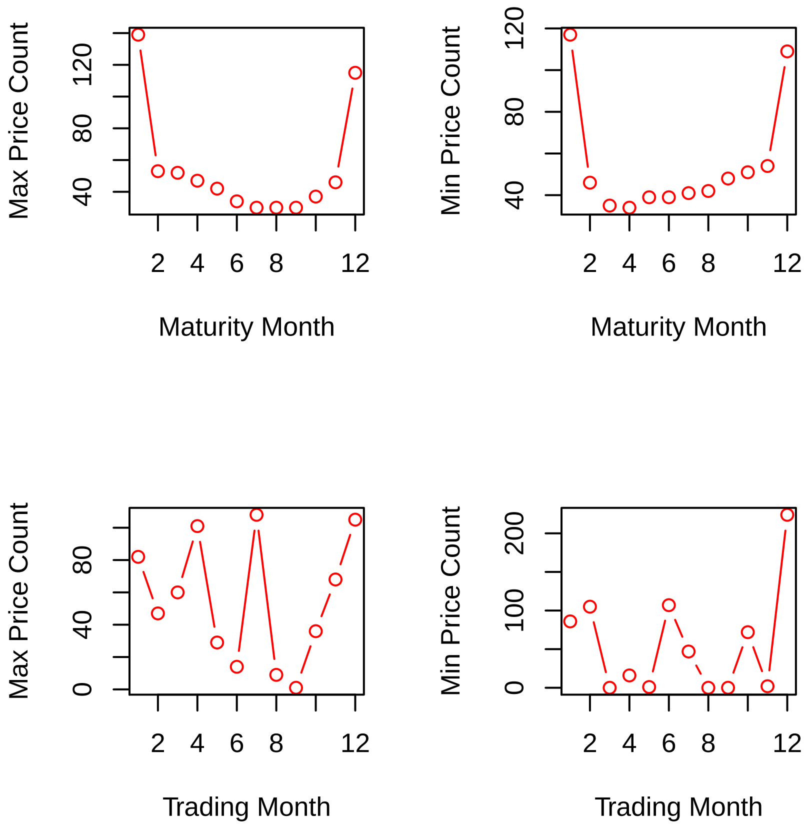

Figure 8: **No. of Counts w.r.t. maturity and trading months, Crude Oil (F1 $\sim$ F60, 2006 $\sim$ 2017)**

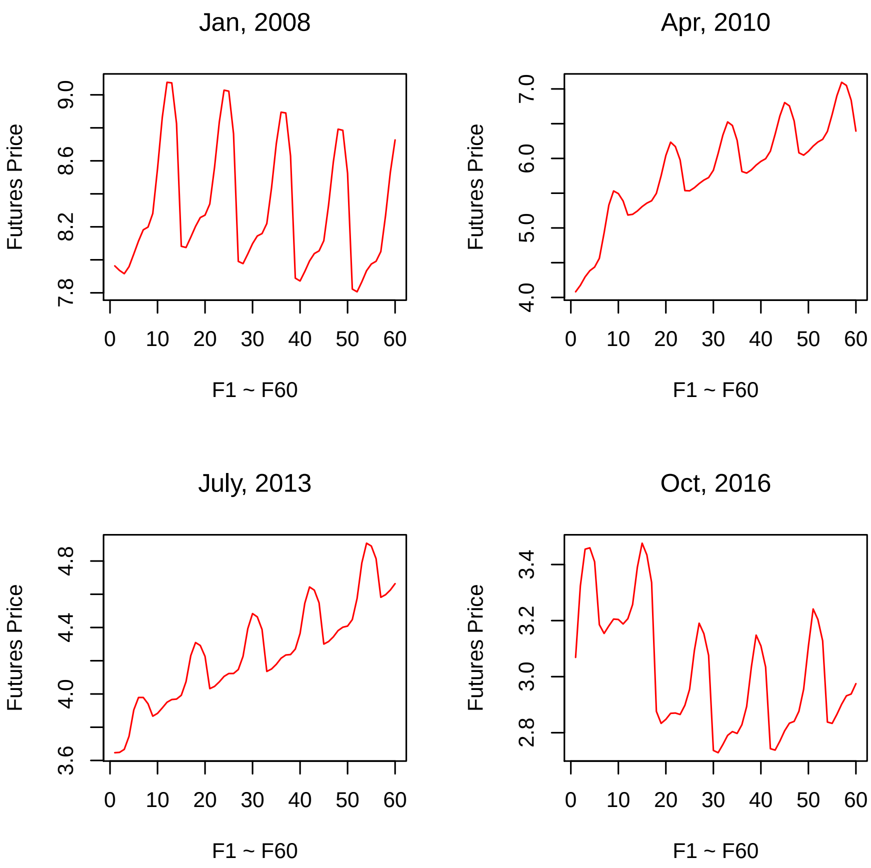

Figure 9: **Forward Curve of a Specific Month, Natural Gas**

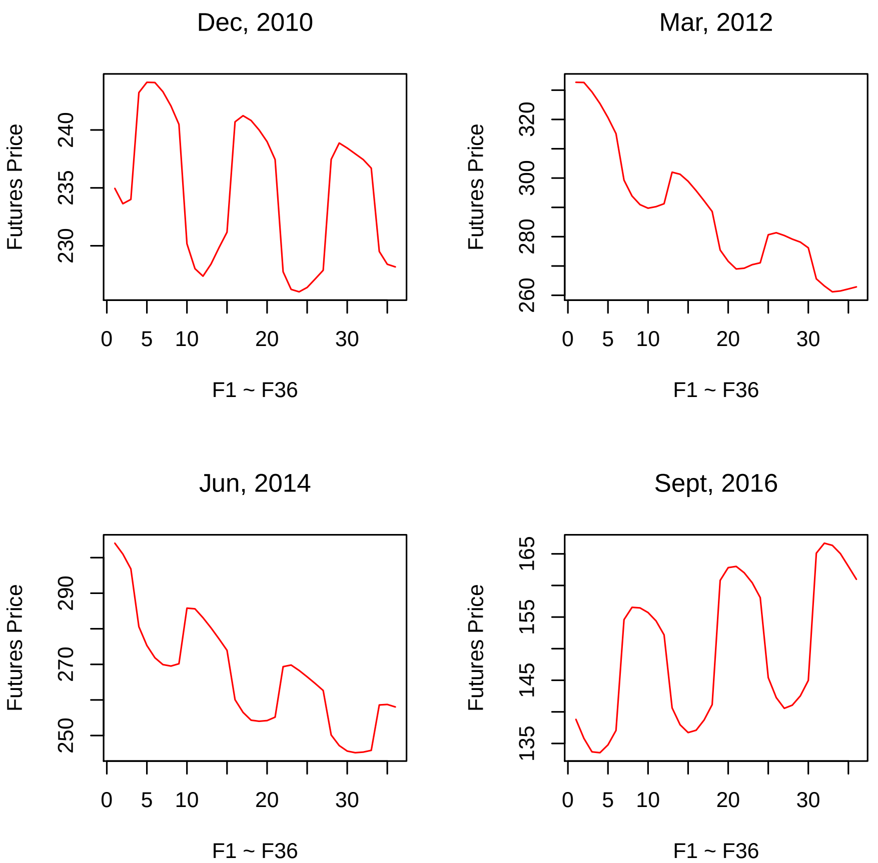

Figure 10: **Forward Curve of a Specific Month, Gasoline**

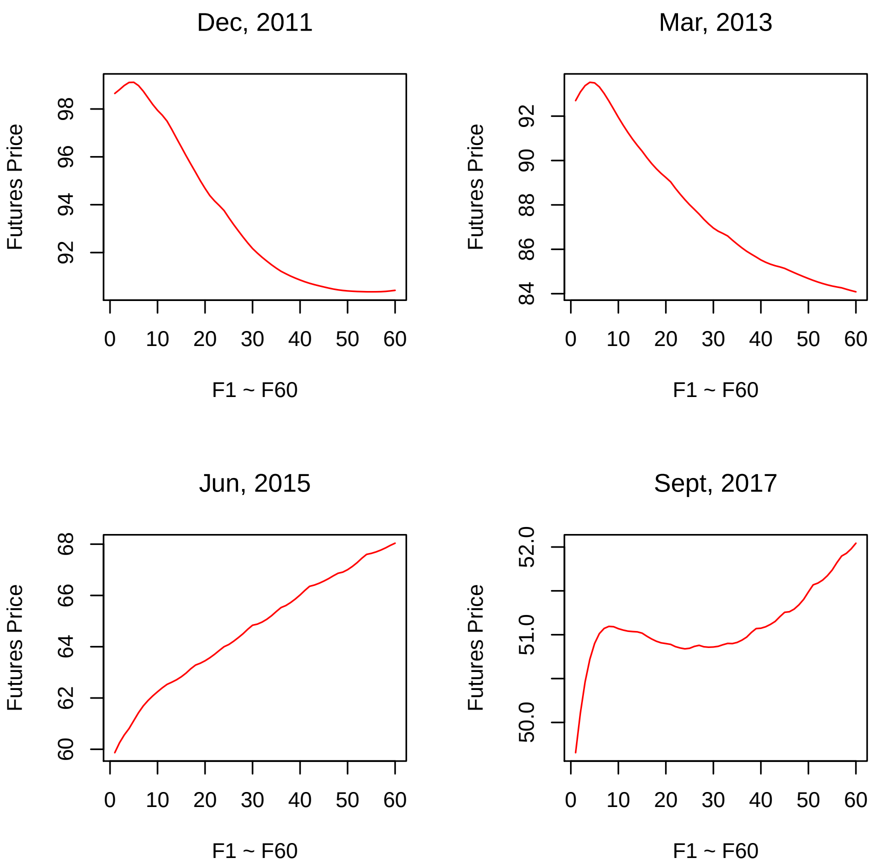

Figure 11: **Forward Curve of a Specific Month, Crude Oil**

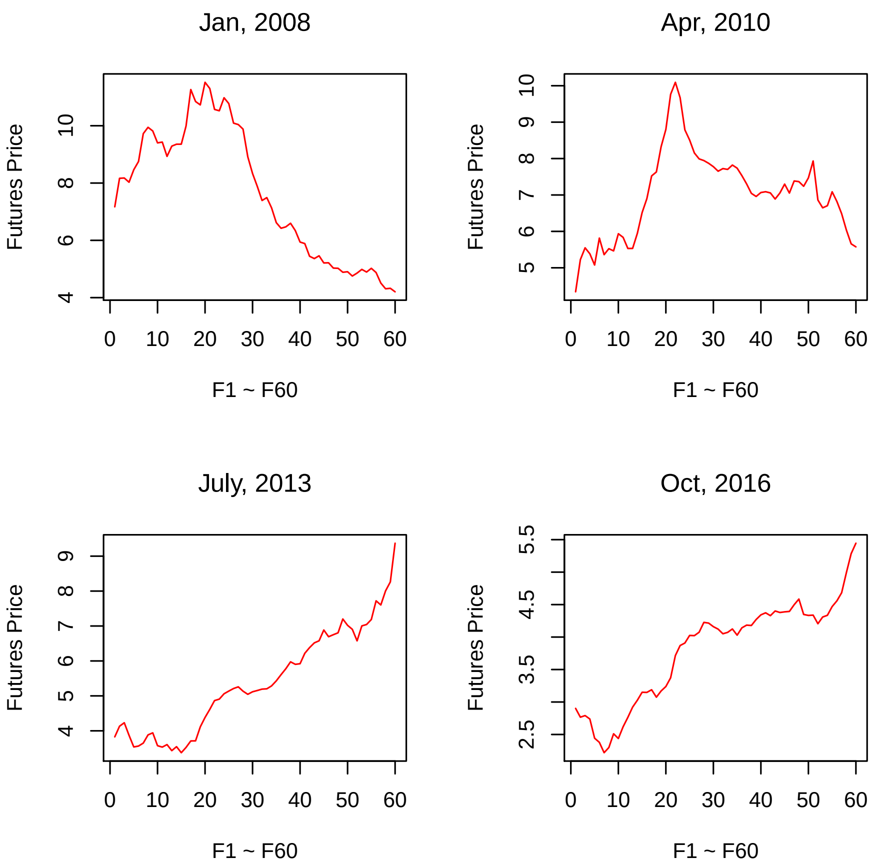

Figure 12: **Backward Curve of a Specific Month, Natural Gas**

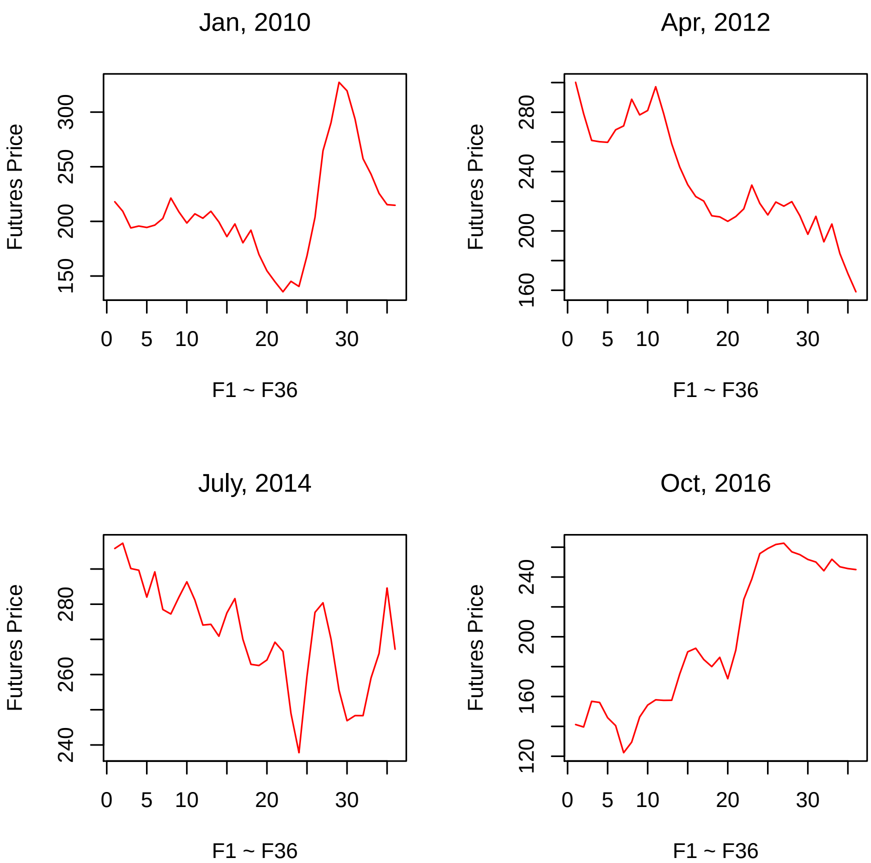

Figure 13: **Backward Curve of a Specific Month, Gasoline**

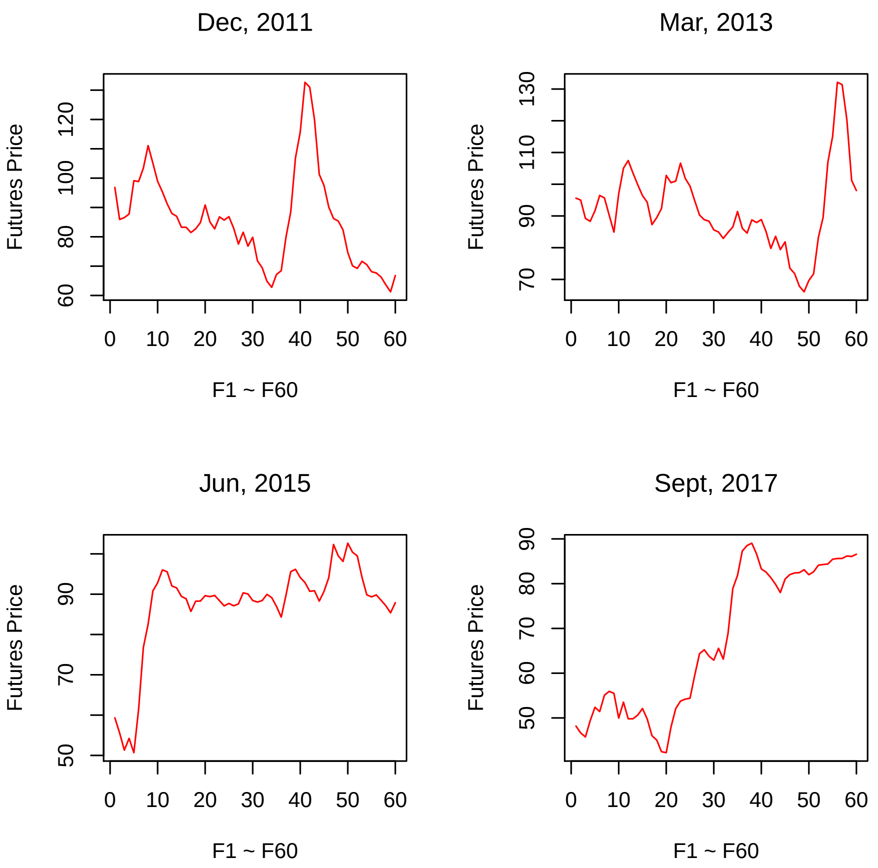

Figure 14: **Backward Curve of a Specific Month, Crude Oil**

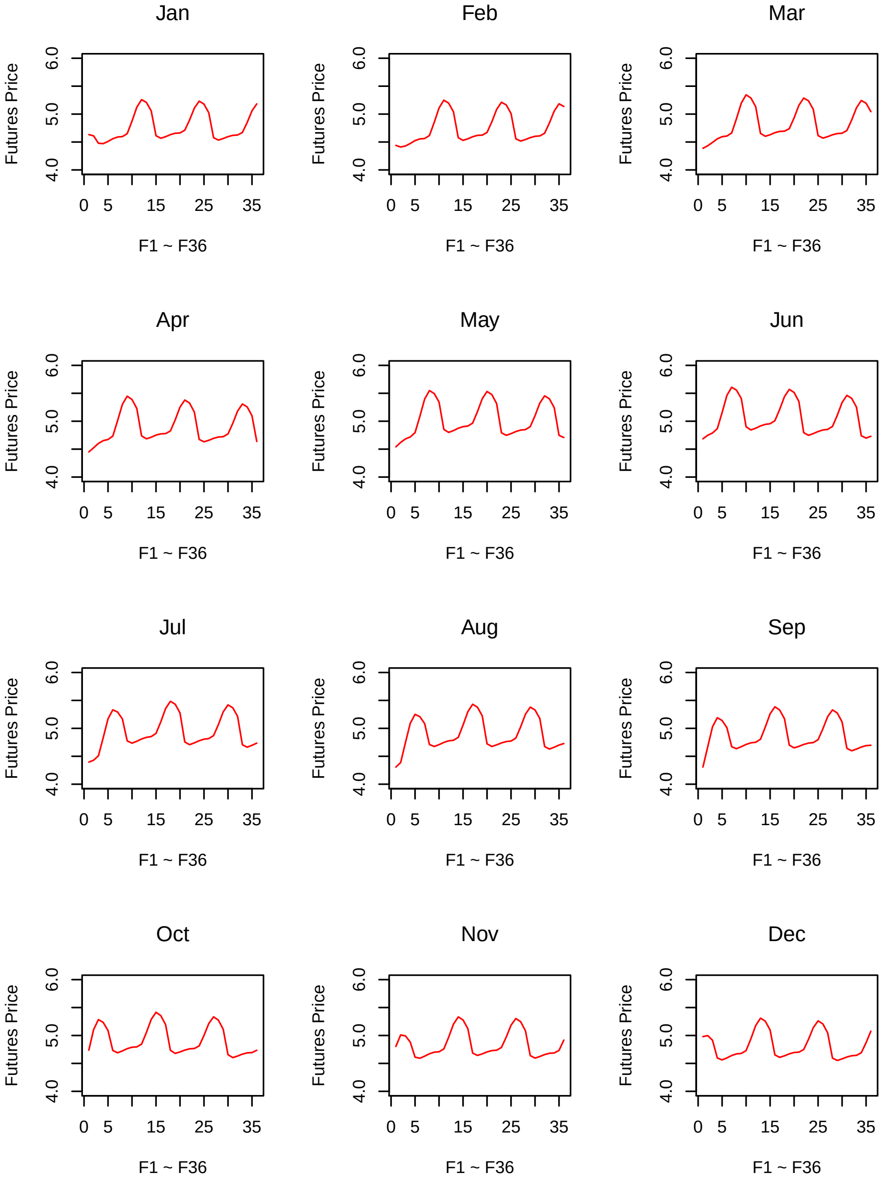

Figure 15: **New Forward Curve, Natural Gas, Group 1 (F1 ∼ F36, 1997 ∼ 2017)**

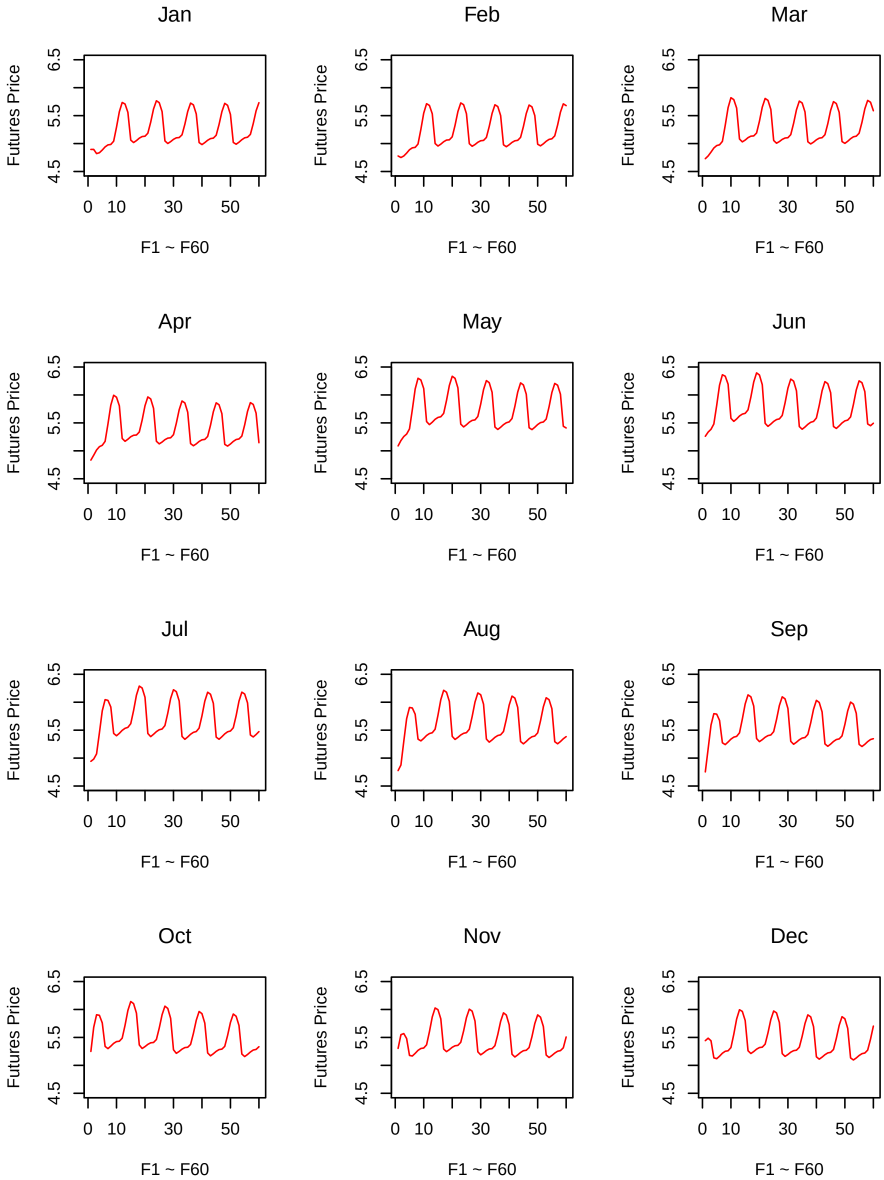

Figure 16: **New Forward Curve, Natural Gas, Group 2 (F1 $\sim$ F60, 2002 $\sim$ 2017)**

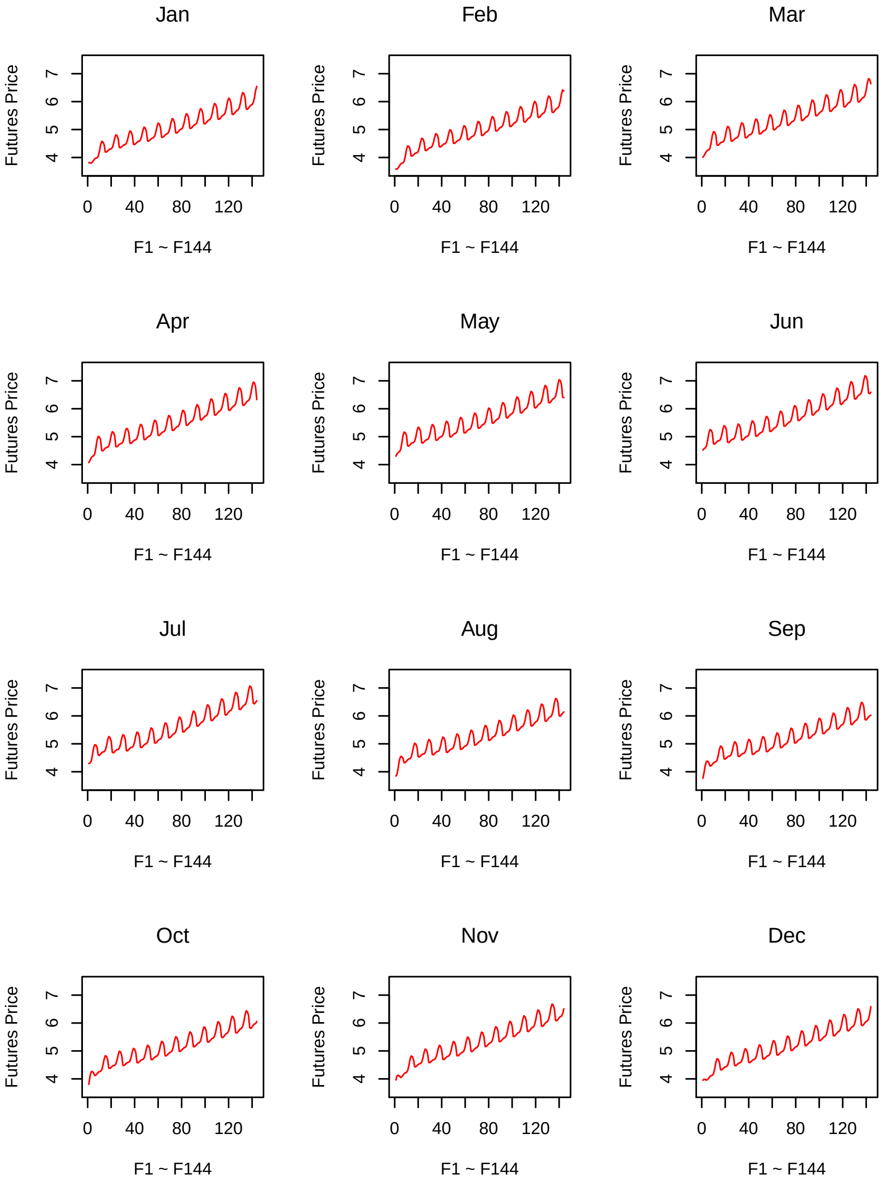

Figure 17: **New Forward Curve, Natural Gas, Group 3 (F1 $\sim$ F144, 2008 $\sim$ 2017)**

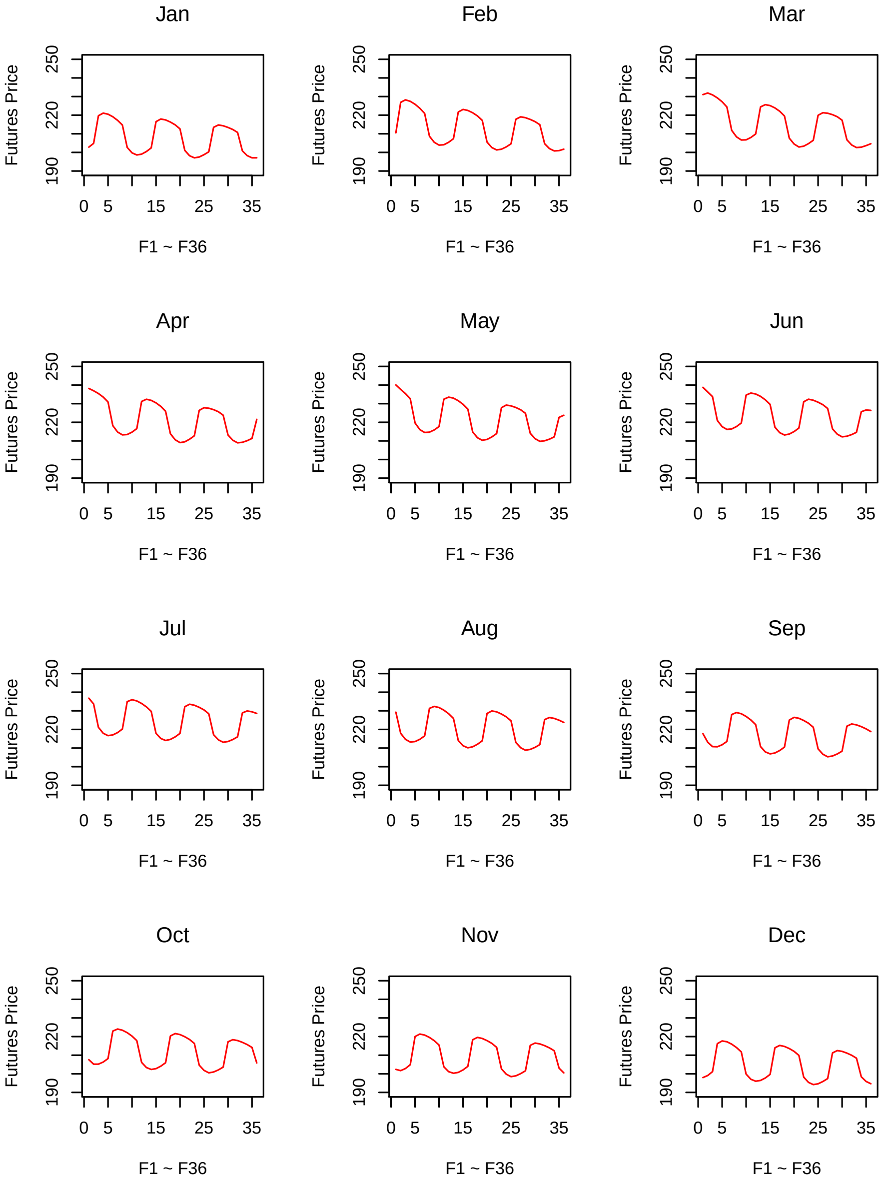

Figure 18: **New Forward Curve, Gasoline (F1 $\sim$ F36, 2007 $\sim$ 2017)**

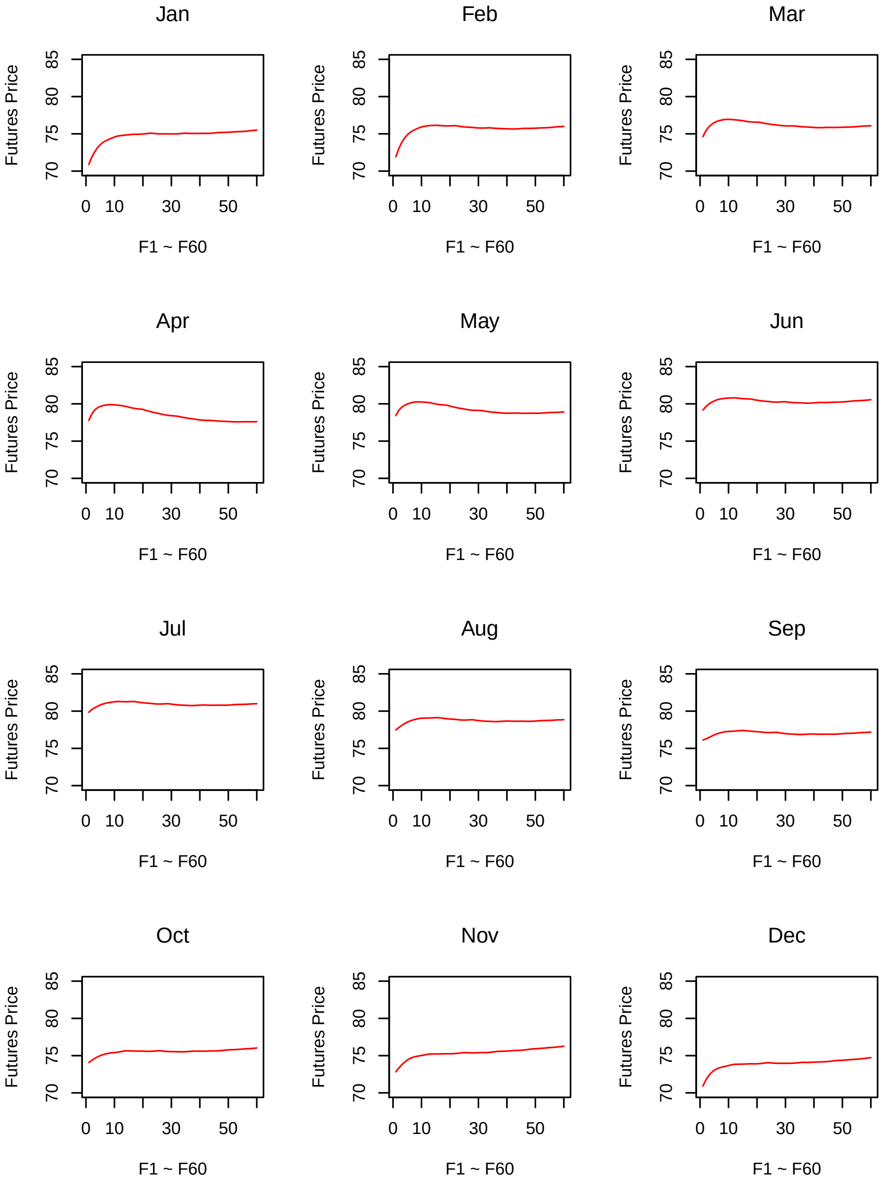

Figure 19: **New Forward Curve, Crude Oil (F1 $\sim$ F60, 2006 $\sim$ 2017)**

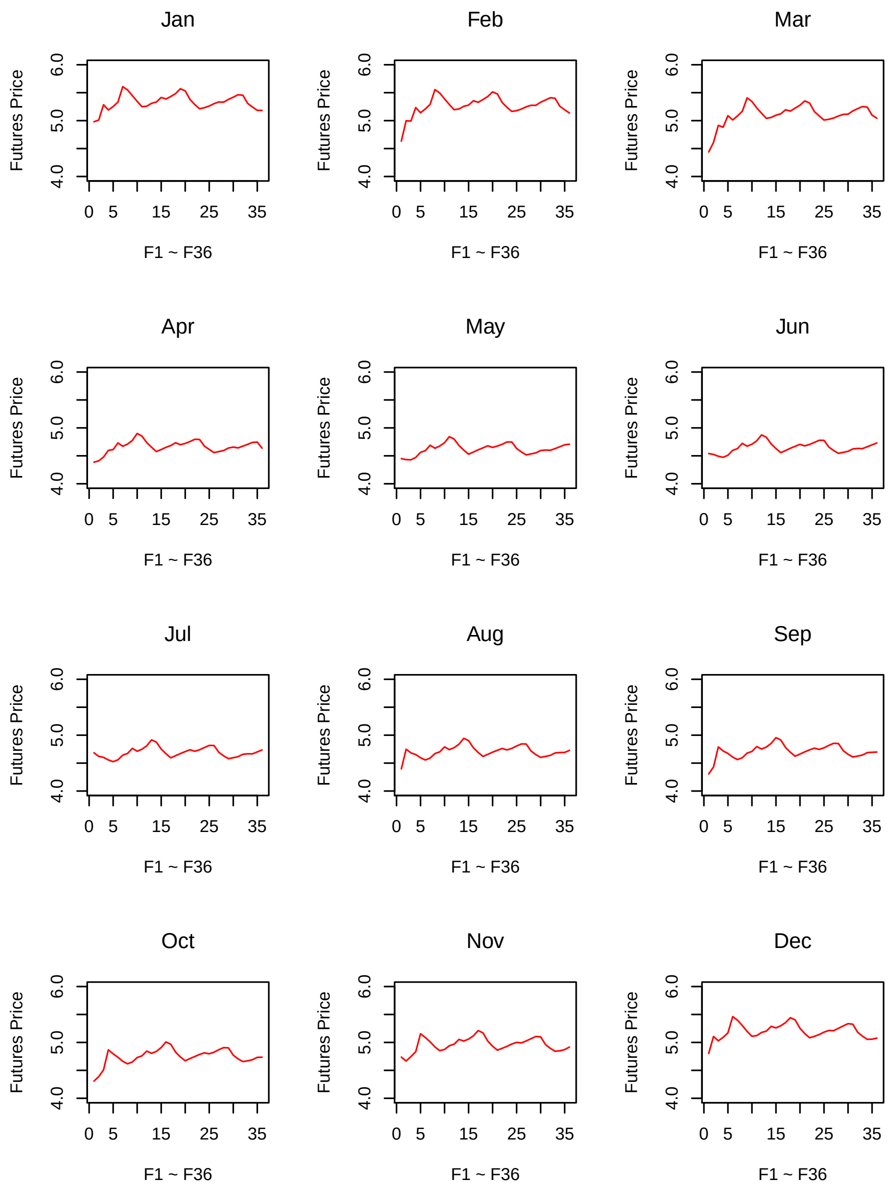

Figure 20: **New Backward Curve, Natural Gas, Group 1 (F1 $\sim$ F36, 1997 $\sim$ 2017)**

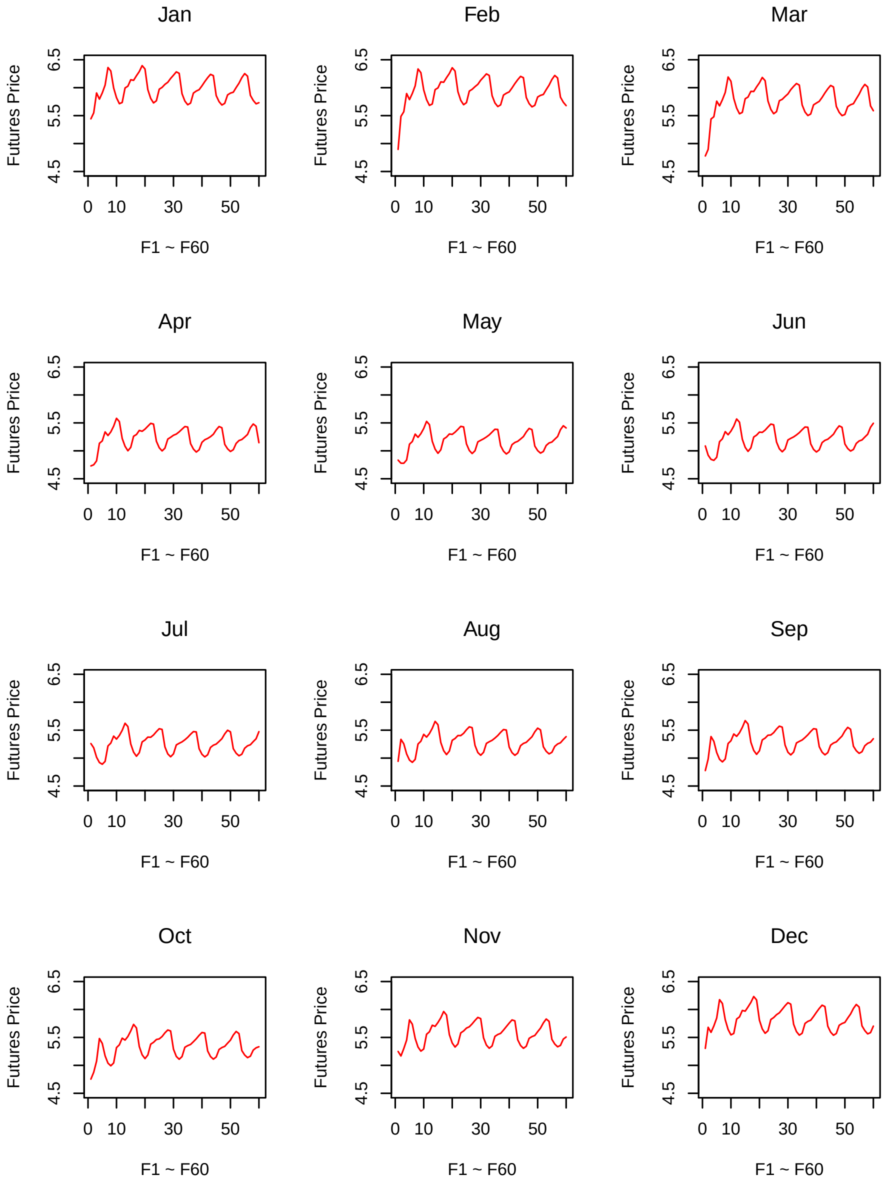

Figure 21: **New Backward Curve, Natural Gas, Group 2 (F1 $\sim$ F60, 2002 $\sim$ 2017)**

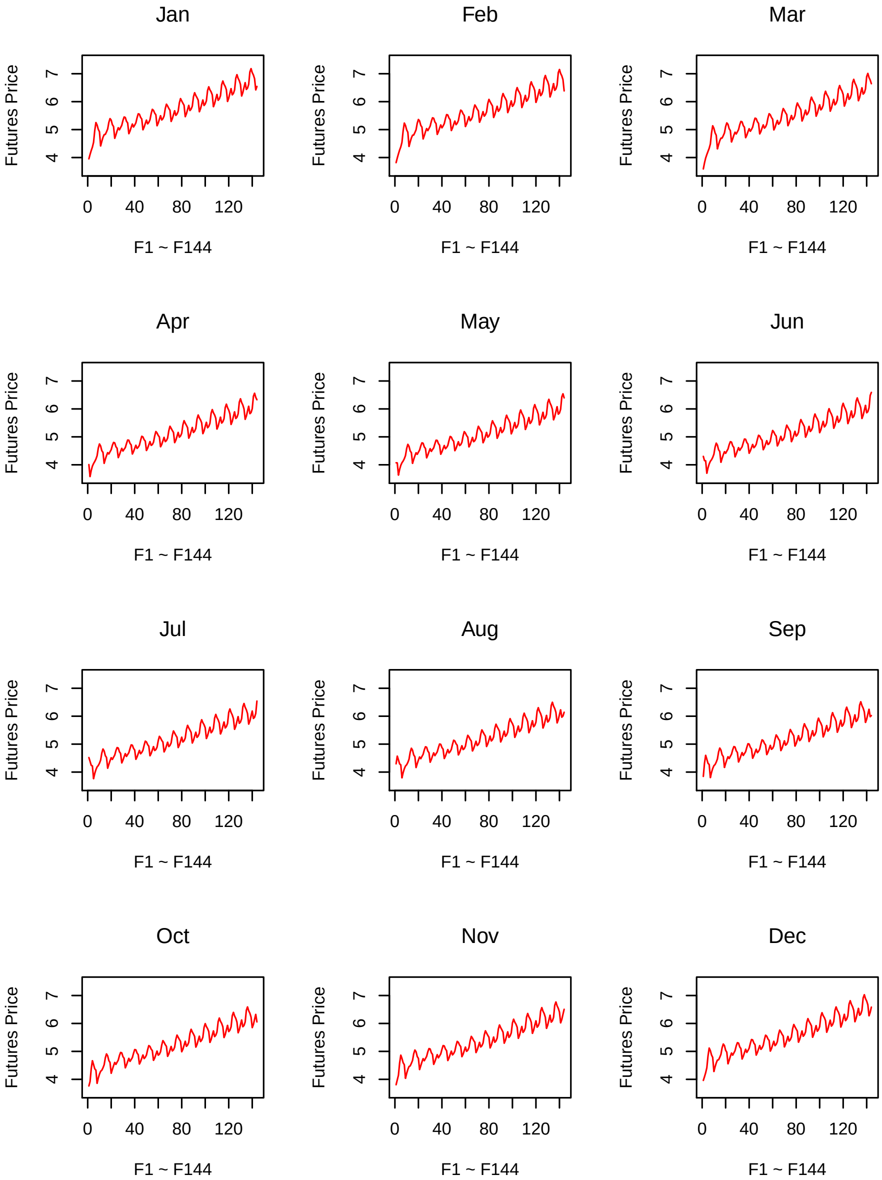

Figure 22: **New Backward Curve, Natural Gas, Group 3 (F1 $\sim$ F144, 2008 $\sim$ 2017)**

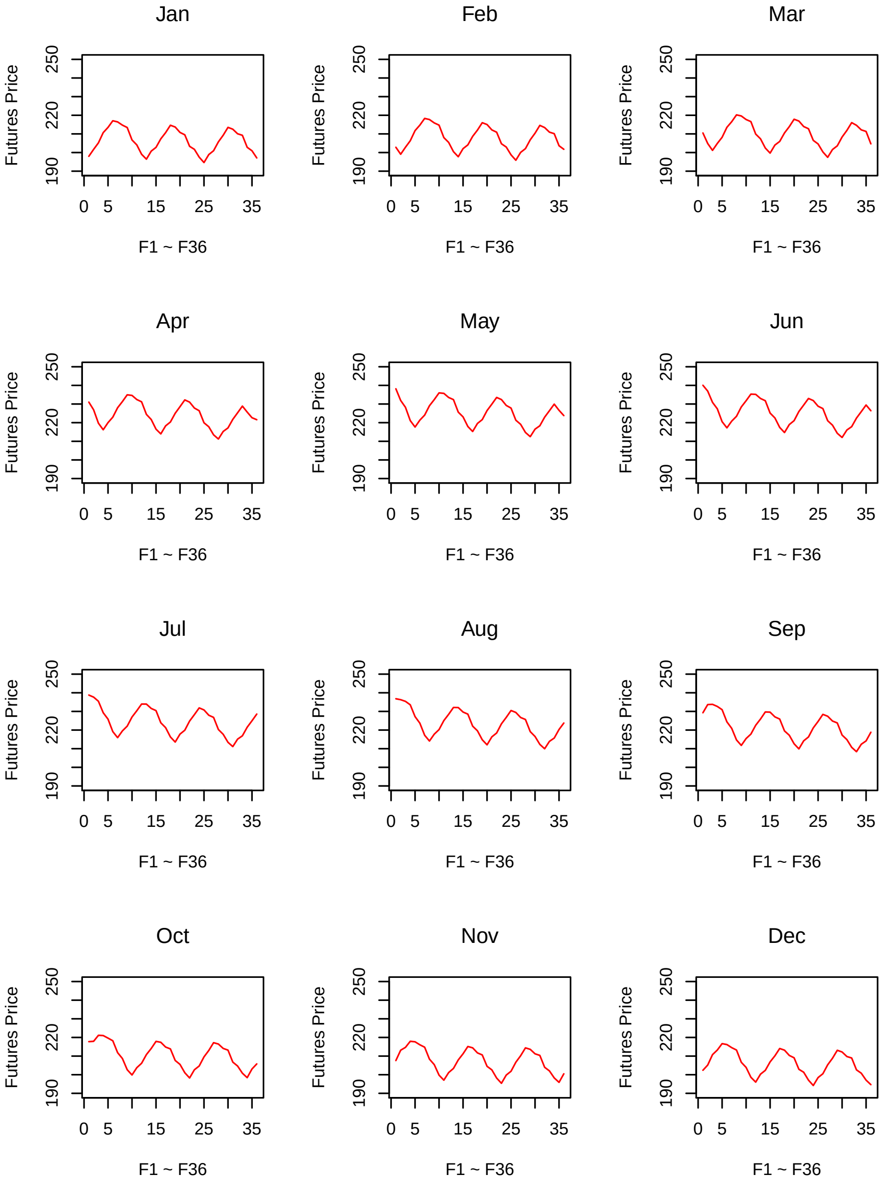

Figure 23: **New Backward Curve, Gasoline (F1 ~ F36, 2007 ~ 2017)**

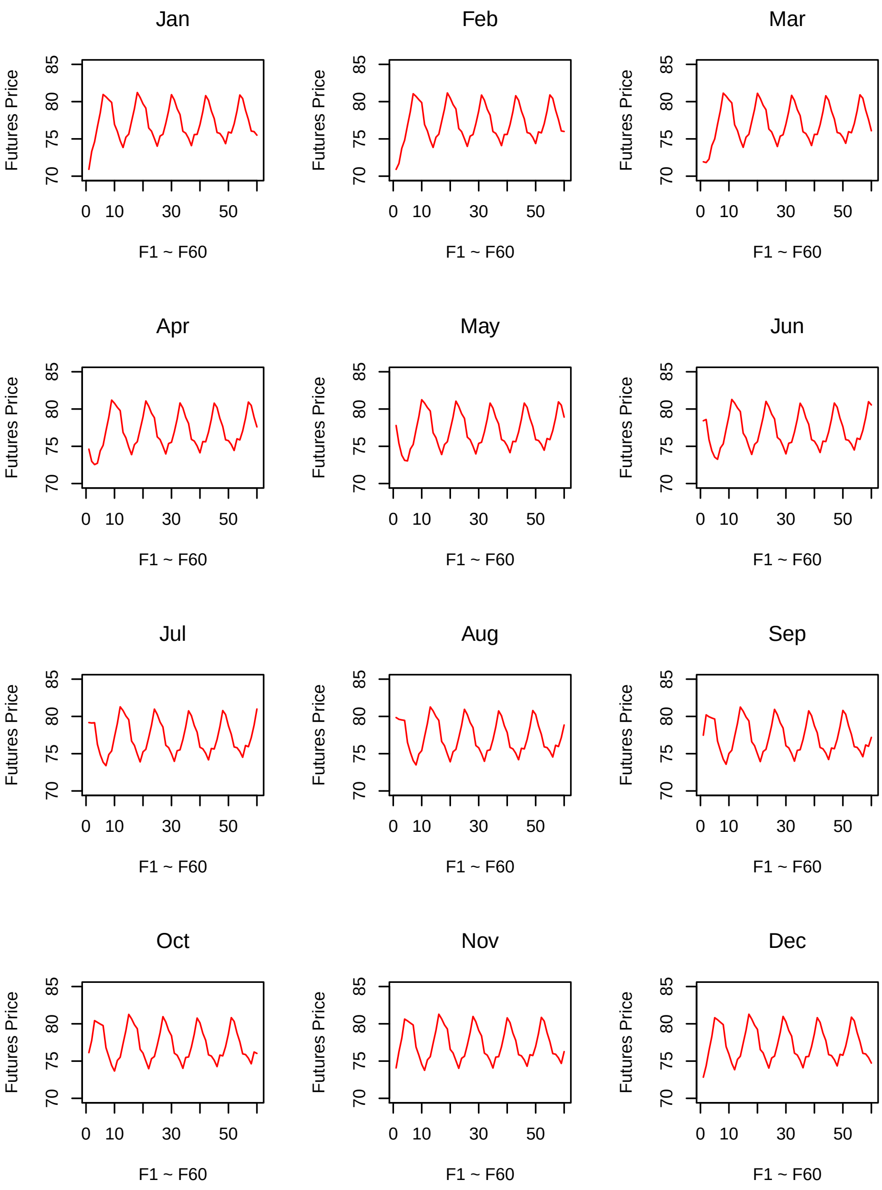

Figure 24: **New Backward Curve, Crude Oil (F1 $\sim$ F60, 2006 $\sim$ 2017)**

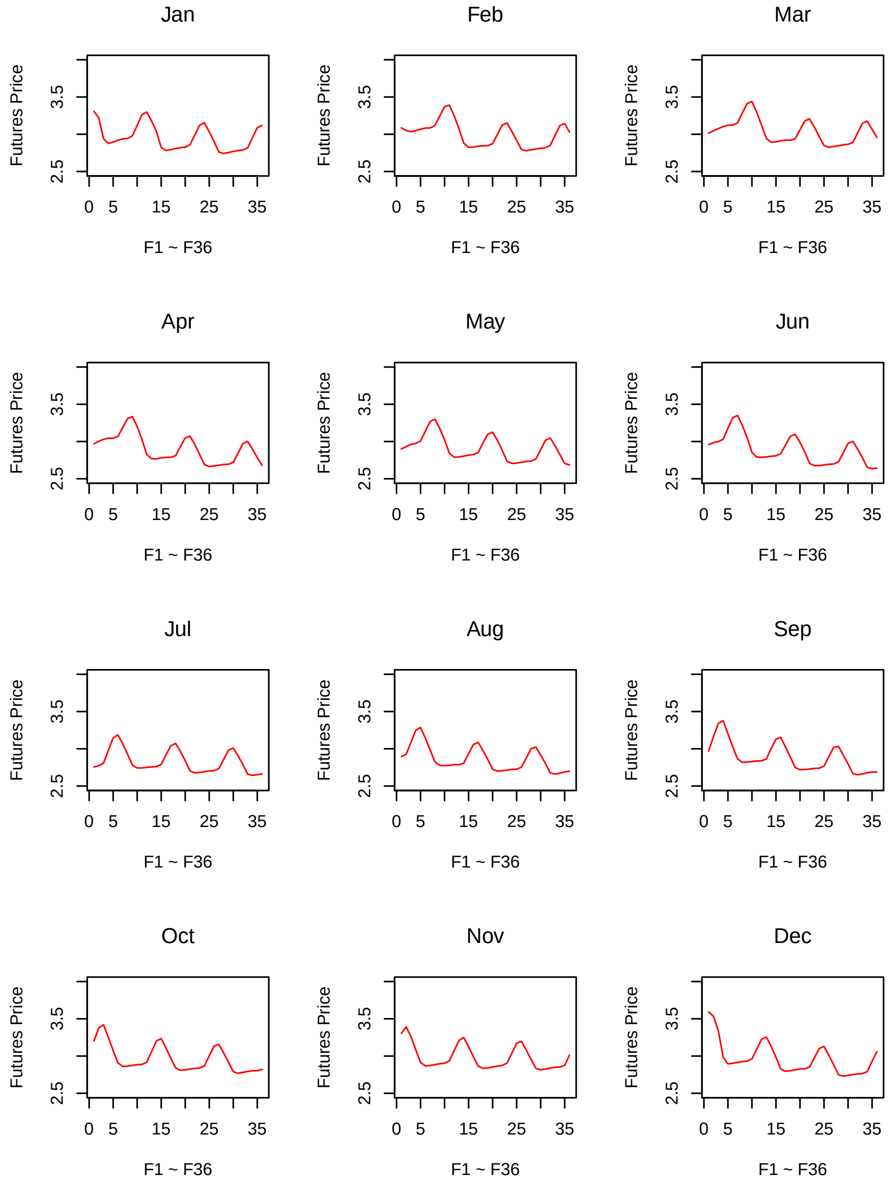

Figure 25: **New Forward Curve, Natural Gas, Sub-Group 1.1 (F1 ∼ F36, 1997 ∼ 2002)**

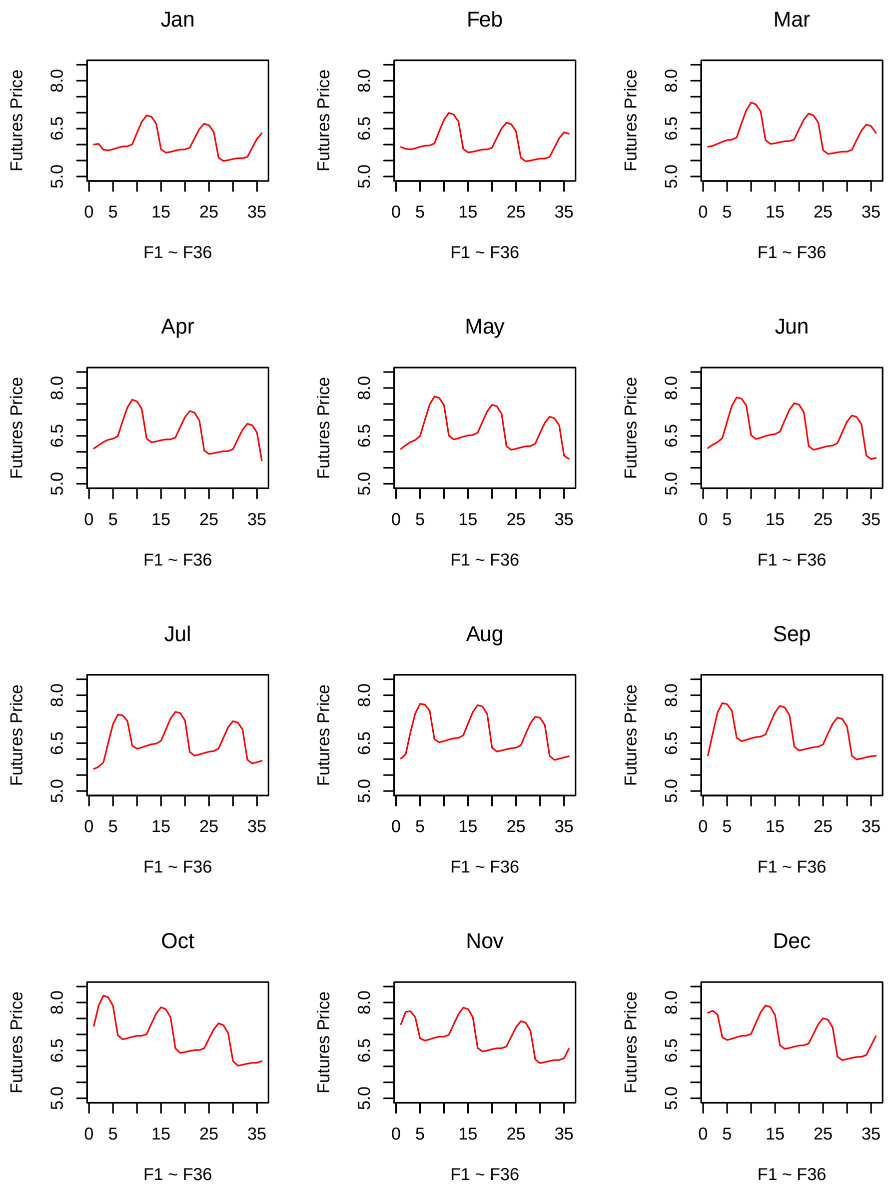

Figure 26: **New Forward Curve, Natural Gas, Sub-Group 1.2 (F1 ∼ F36, 2002 ∼ 2008)**

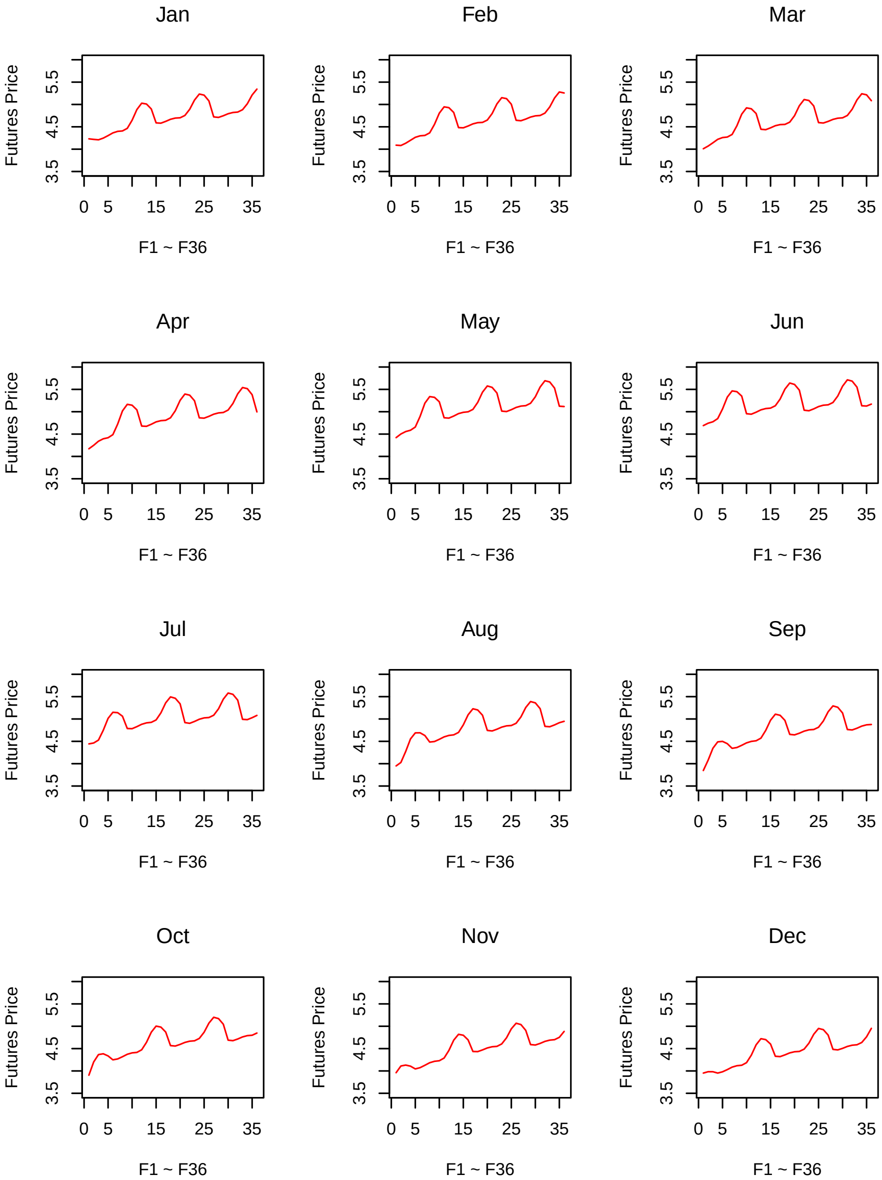

Figure 27: **New Forward Curve, Natural Gas, Sub-Group 1.3 (F1 ∼ F36, 2008 ∼ 2017)**

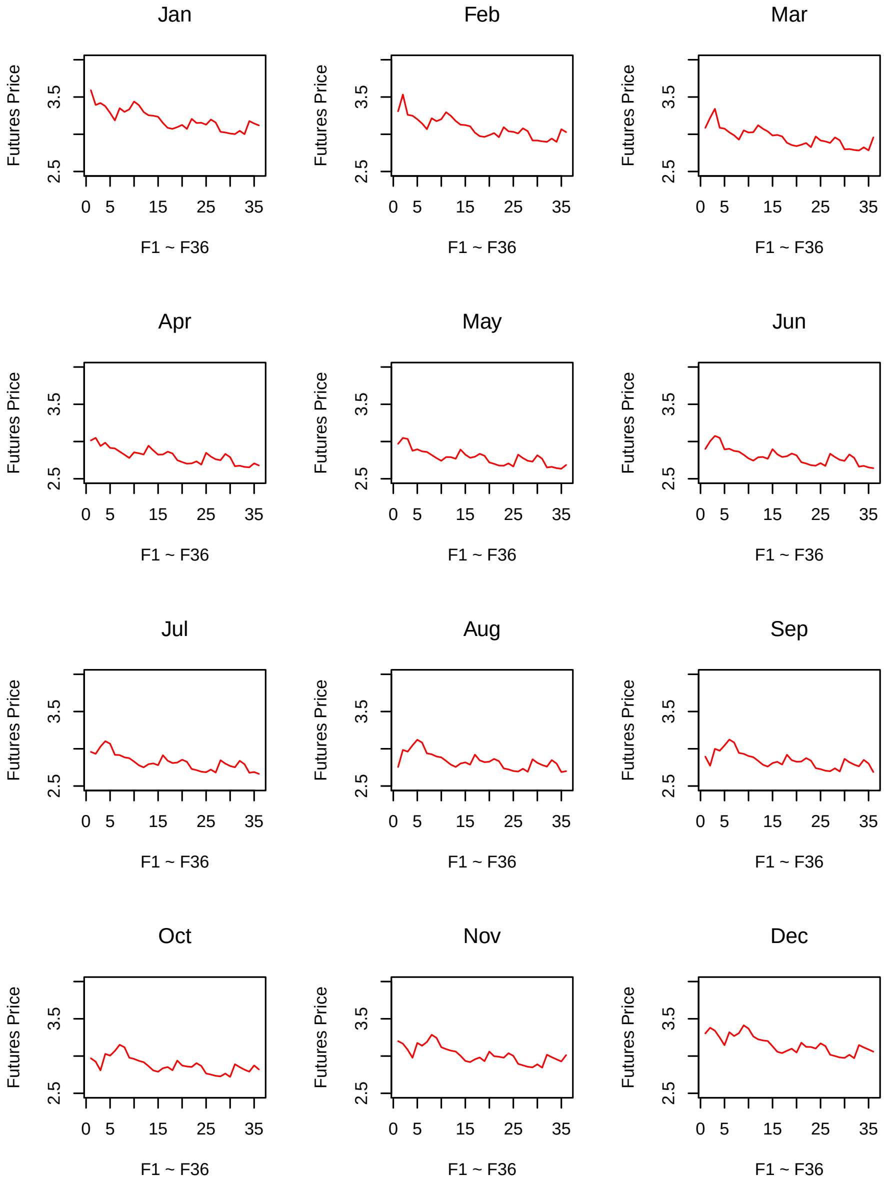

Figure 28: **New Backward Curve, Natural Gas, Sub-Group 1.1 (F1 $\sim$ F36, 1997 $\sim$ 2002)**

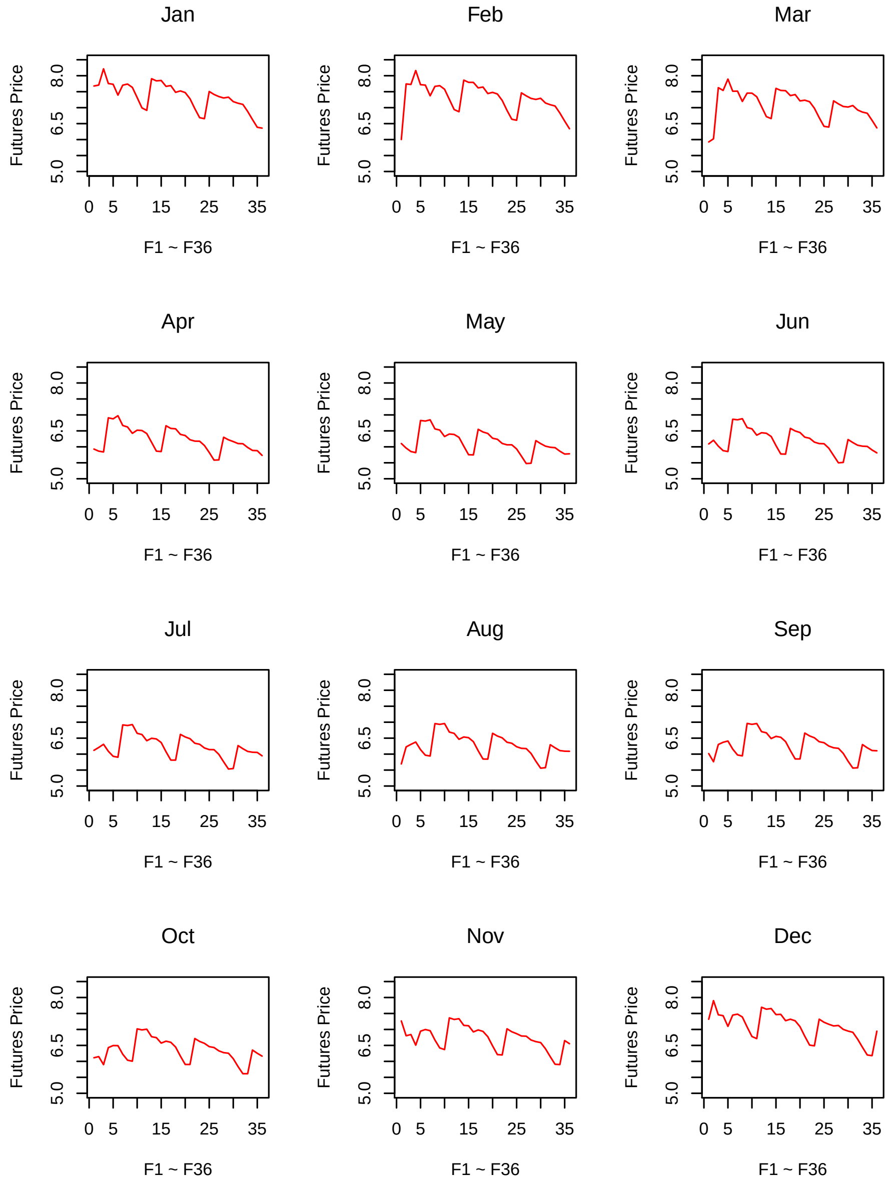

Figure 29: **New Backward Curve, Natural Gas, Sub-Group 1.2 (F1 $\sim$ F36, 2002 $\sim$ 2008)**

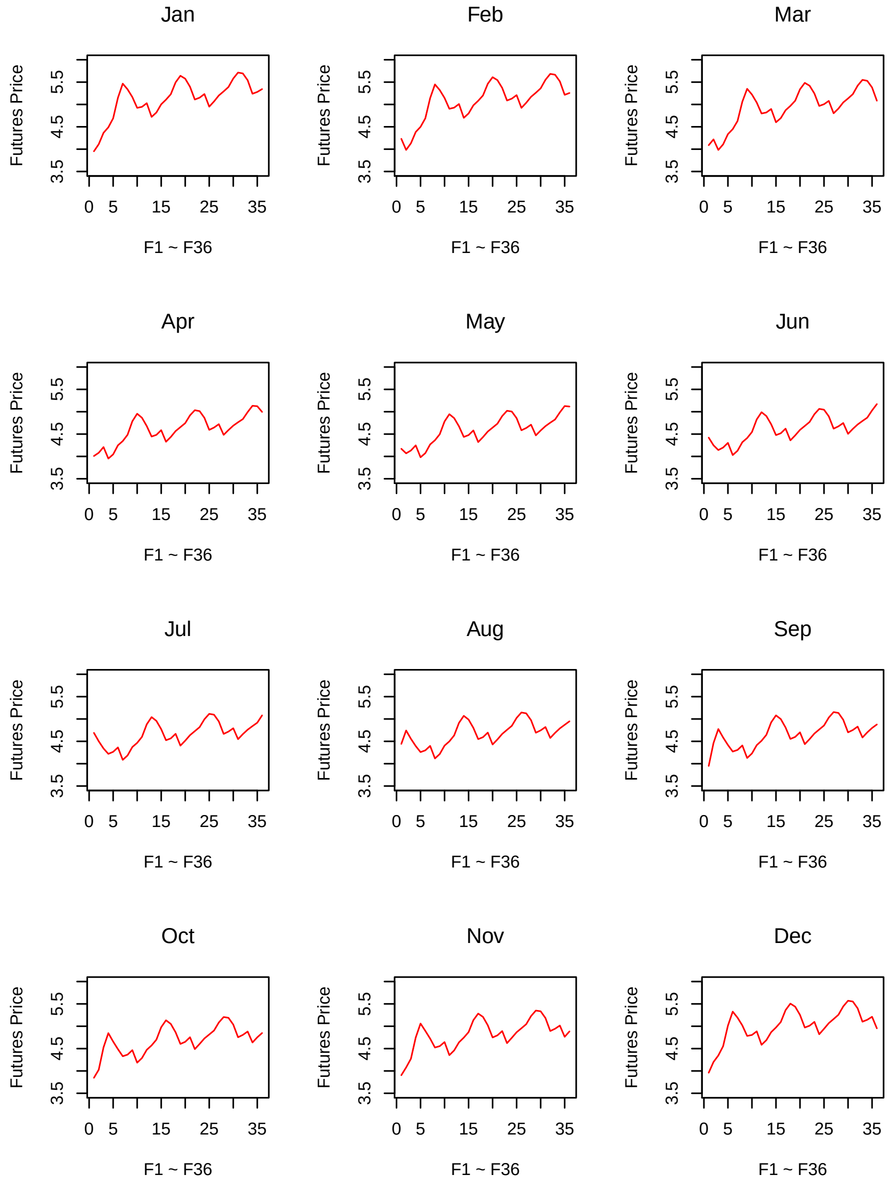

Figure 30: **New Backward Curve, Natural Gas, Sub-Group 1.3 (F1 $\sim$ F36, 2008 $\sim$ 2017)**

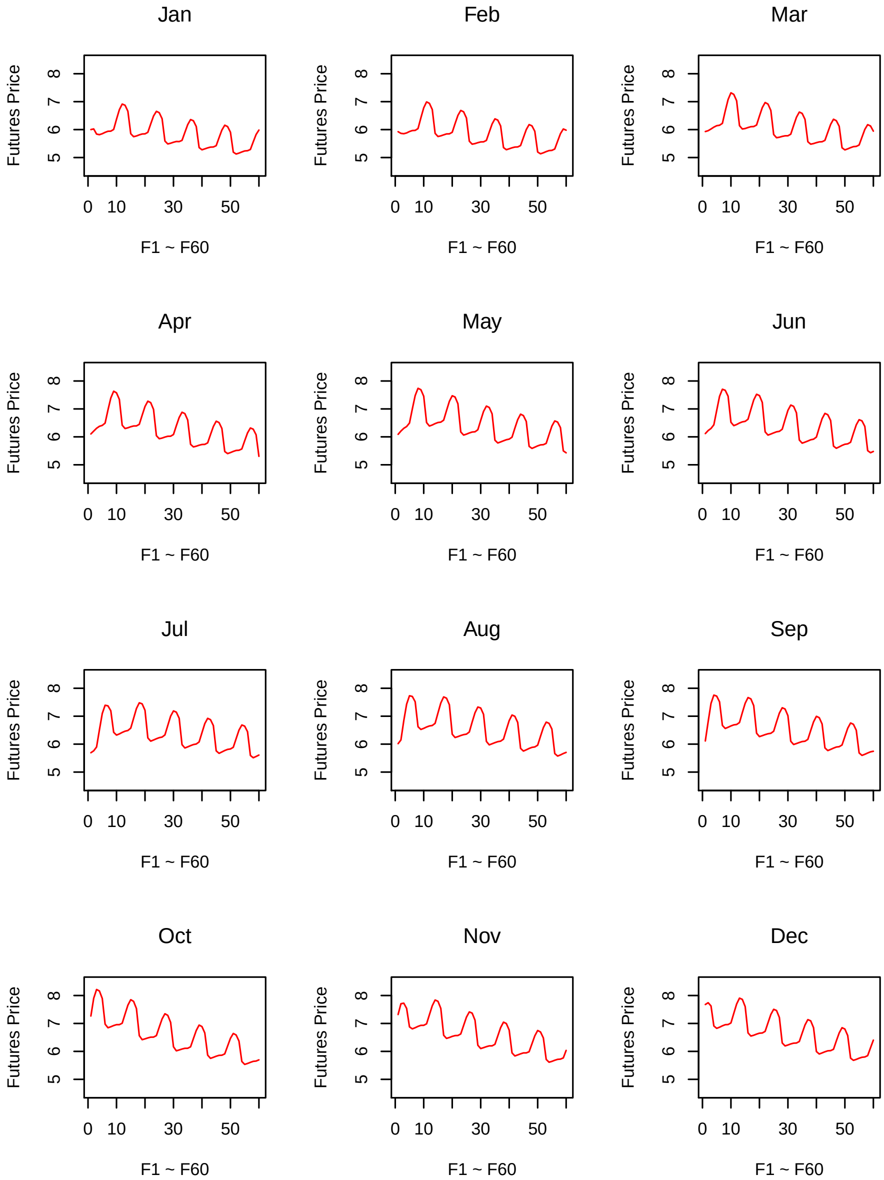

Figure 31: **New Forward Curve, Natural Gas, Sub-Group 2.1 (F1 $\sim$ F60, 2002 $\sim$ 2008)**

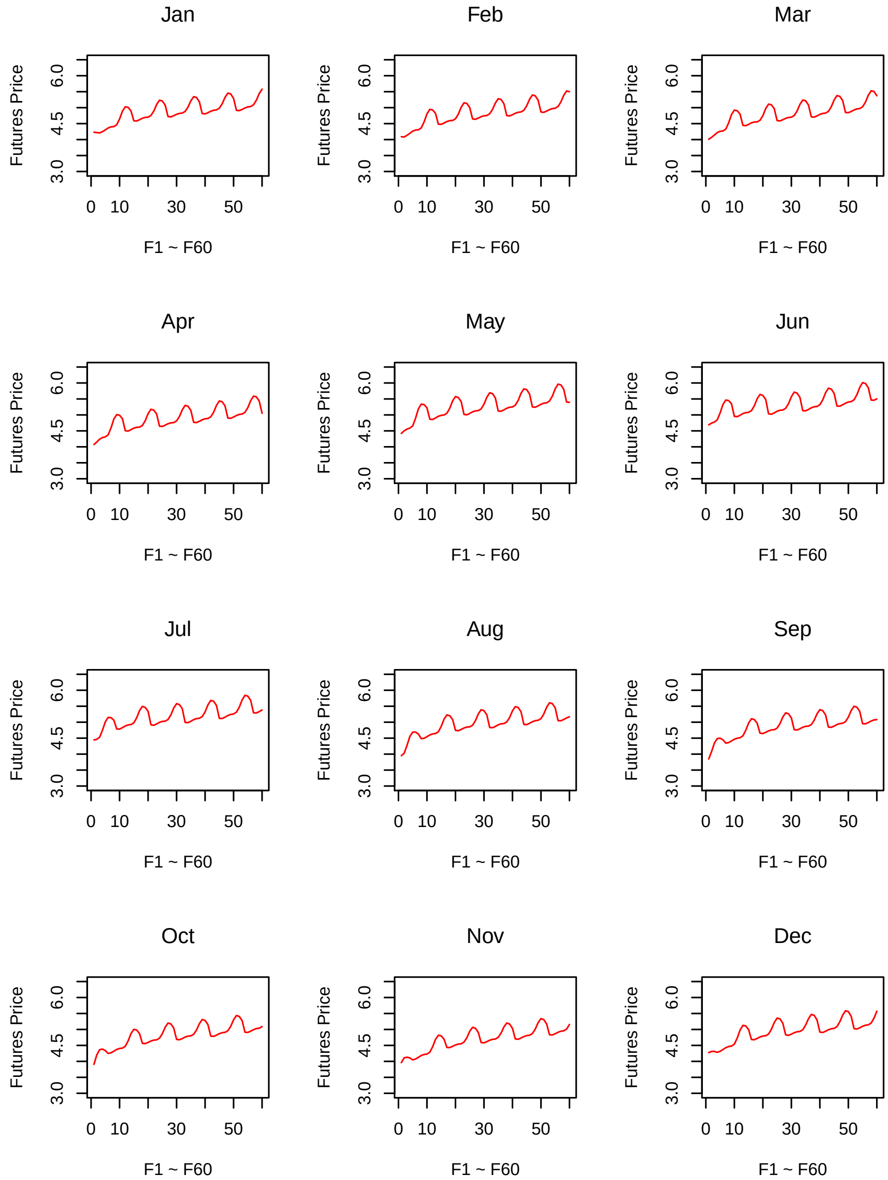

Figure 32: **New Forward Curve, Natural Gas, Sub-Group 2.2 (F1 ∼ F60, 2008 ∼ 2017)**

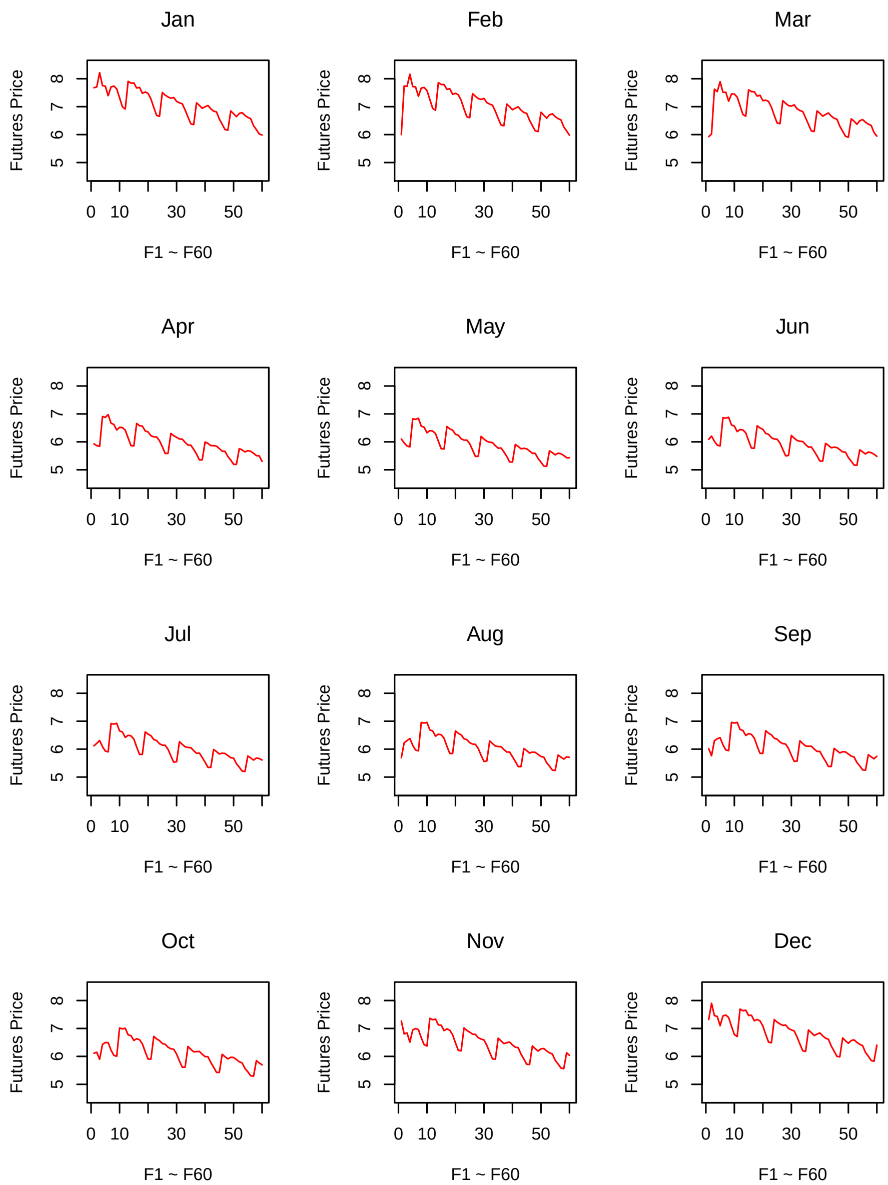

Figure 33: **New Backward Curve, Natural Gas, Sub-Group 2.1 (F1 $\sim$ F60, 2002 $\sim$ 2008)**

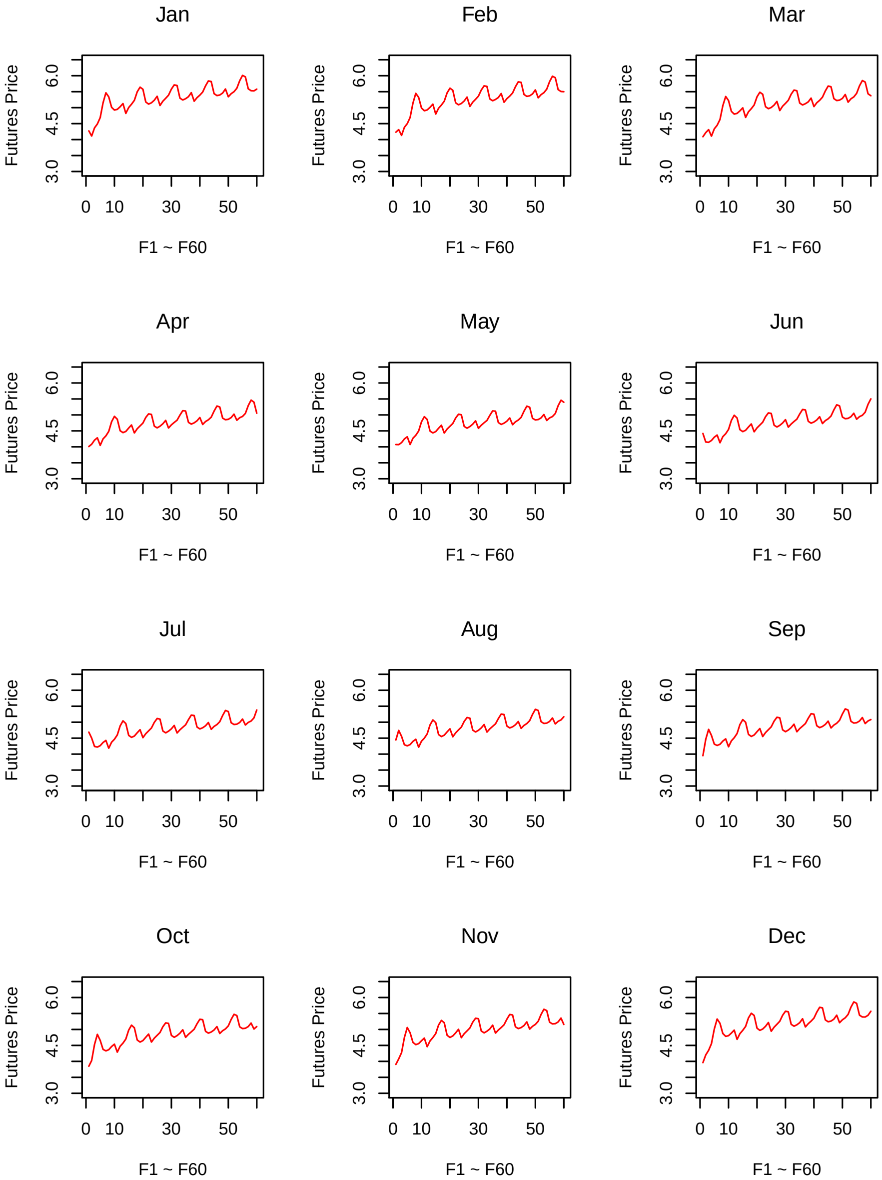

Figure 34: **New Backward Curve, Natural Gas, Sub-Group 2.2 (F1 $\sim$ F60, 2008 $\sim$ 2017)**

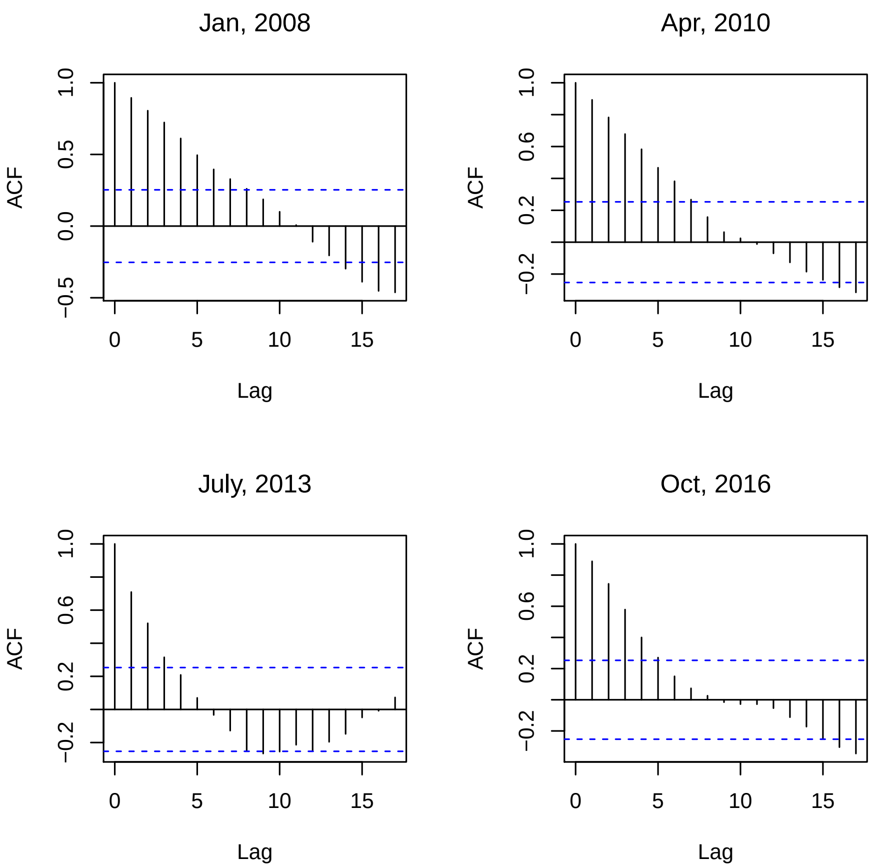

Figure 35: **Autocorrelation Function on (Detrended) Old Backward Curves of 4 Specific Months, Natural Gas (F1 $\sim$ F60, 2002 $\sim$ 2008)**

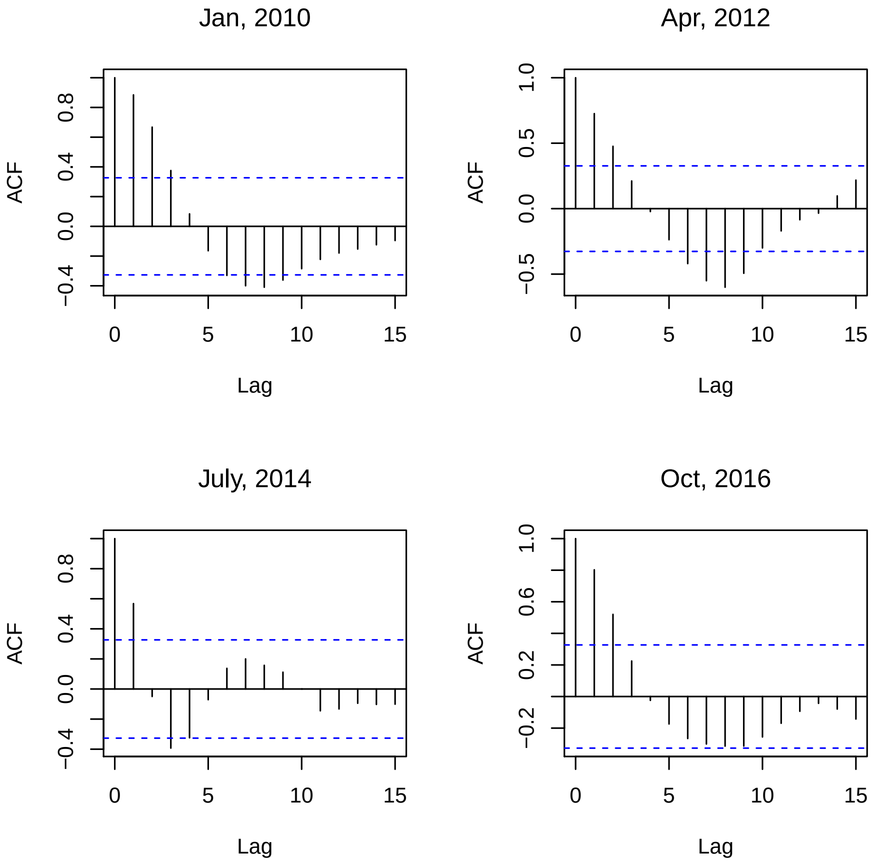

Figure 36: **Autocorrelation Function on (Detrended) Old Backward Curves of 4 Specific Months, Gasoline (F1 $\sim$ F36, 2007 $\sim$ 2017)**

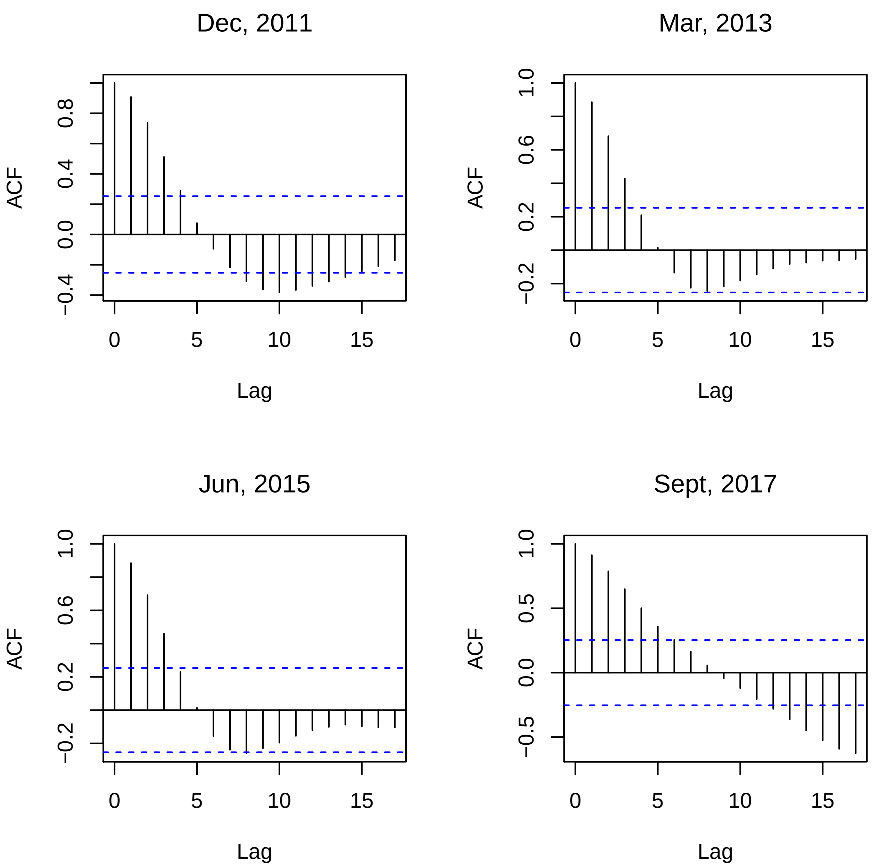

Figure 37: **Autocorrelation Function on (Detrended) Old Backward Curves of 4 Specific Months, Crude Oil (F1 ∼ F60, 2006 ∼ 2017)**

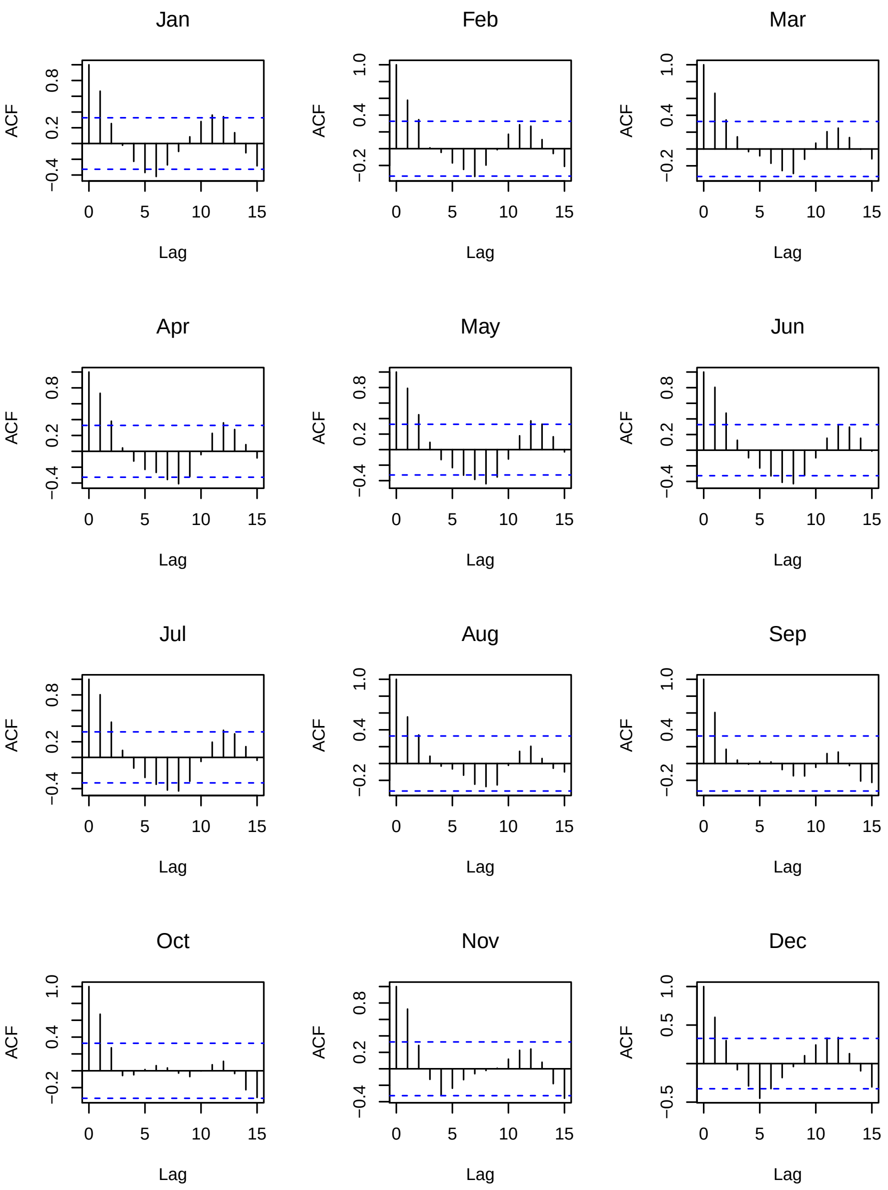

Figure 38: **Autocorrelation Function on (Detrended) New Backward Curve, Natural Gas, Group 1 (F1 ∼ F36, 1997 ∼ 2017)**

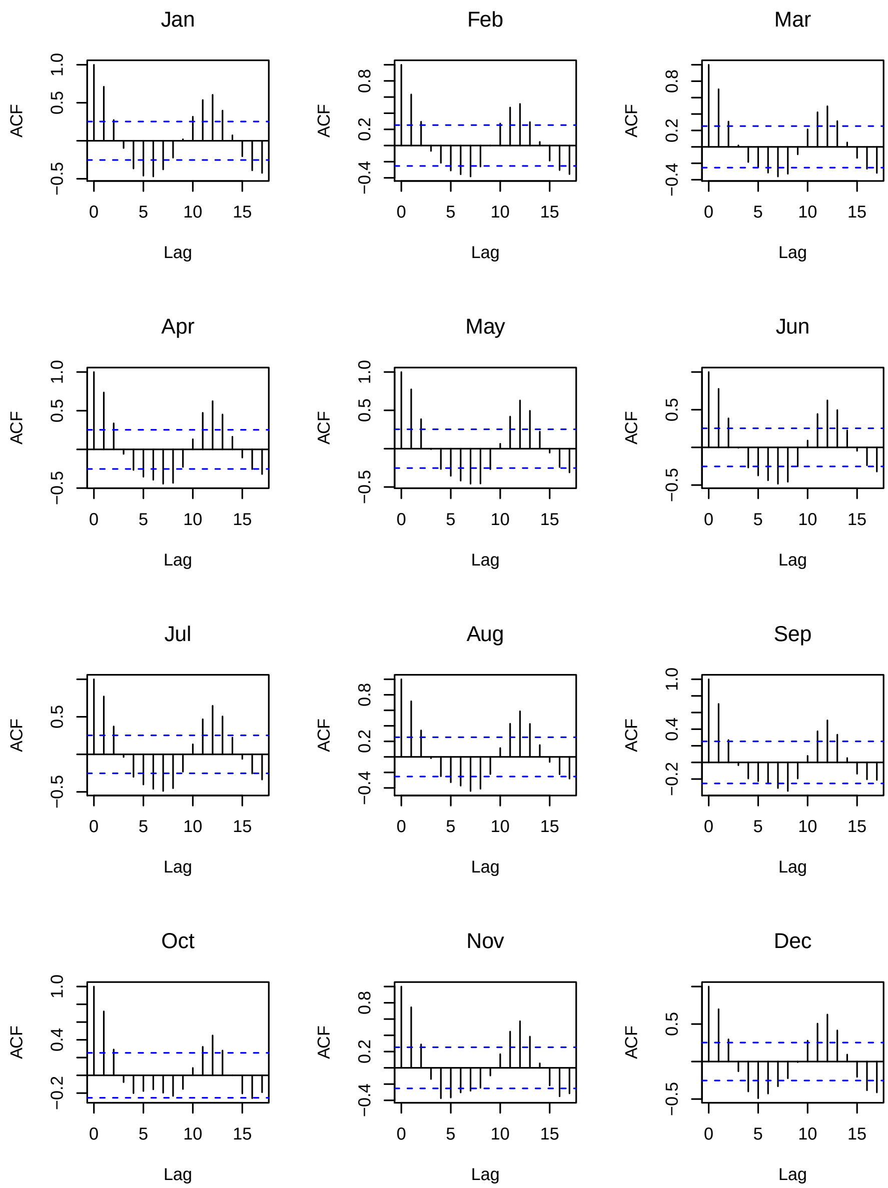

Figure 39: **Autocorrelation Function on (Detrended) New Backward Curve, Natural Gas, Group 2 (F1 $\sim$ F60, 2002 $\sim$ 2017)**

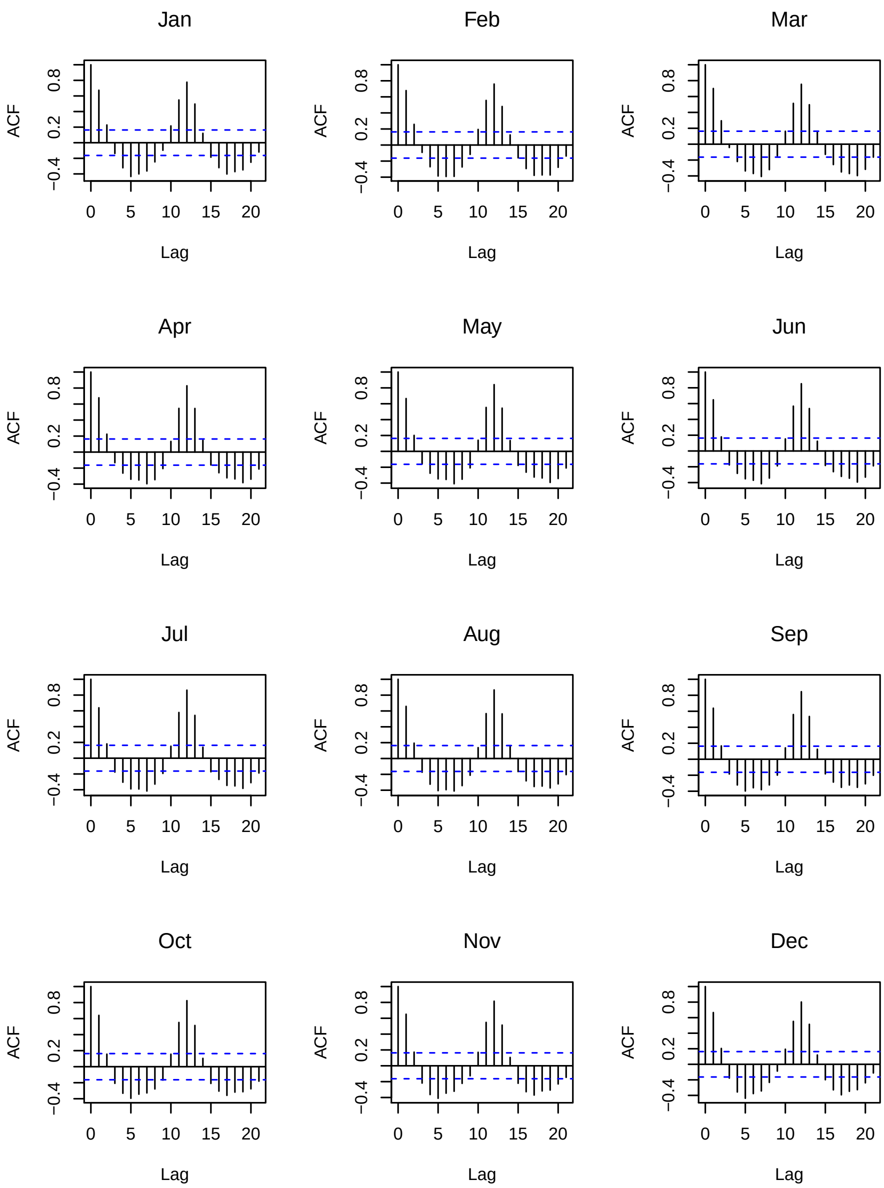

Figure 40: **Autocorrelation Function on (Detrended) New Backward Curve, Natural Gas, Group 3 (F1 ∼ F144, 2008 ∼ 2017)**

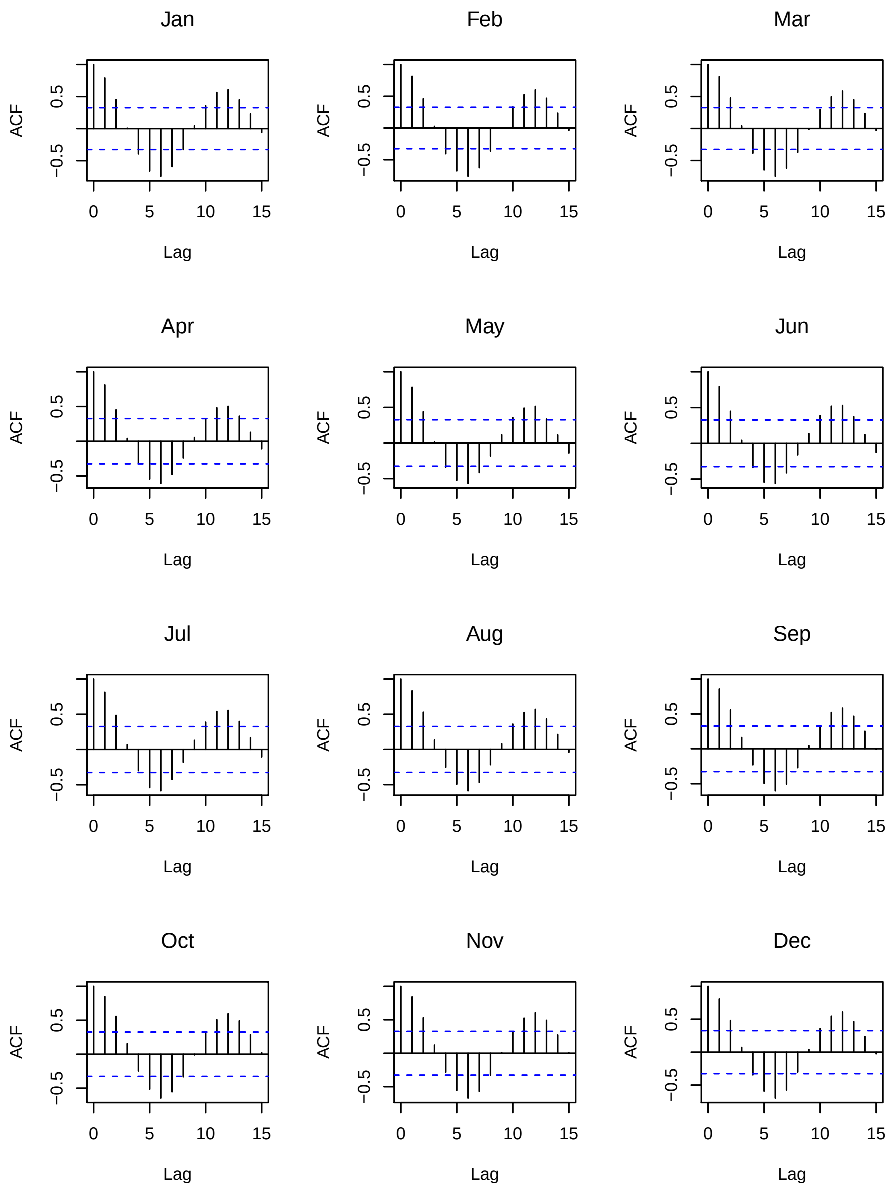

Figure 41: **Autocorrelation Function on (Detrended) New Backward Curve, Gasoline (F1 ∼ F36, 2007 ∼ 2017)**

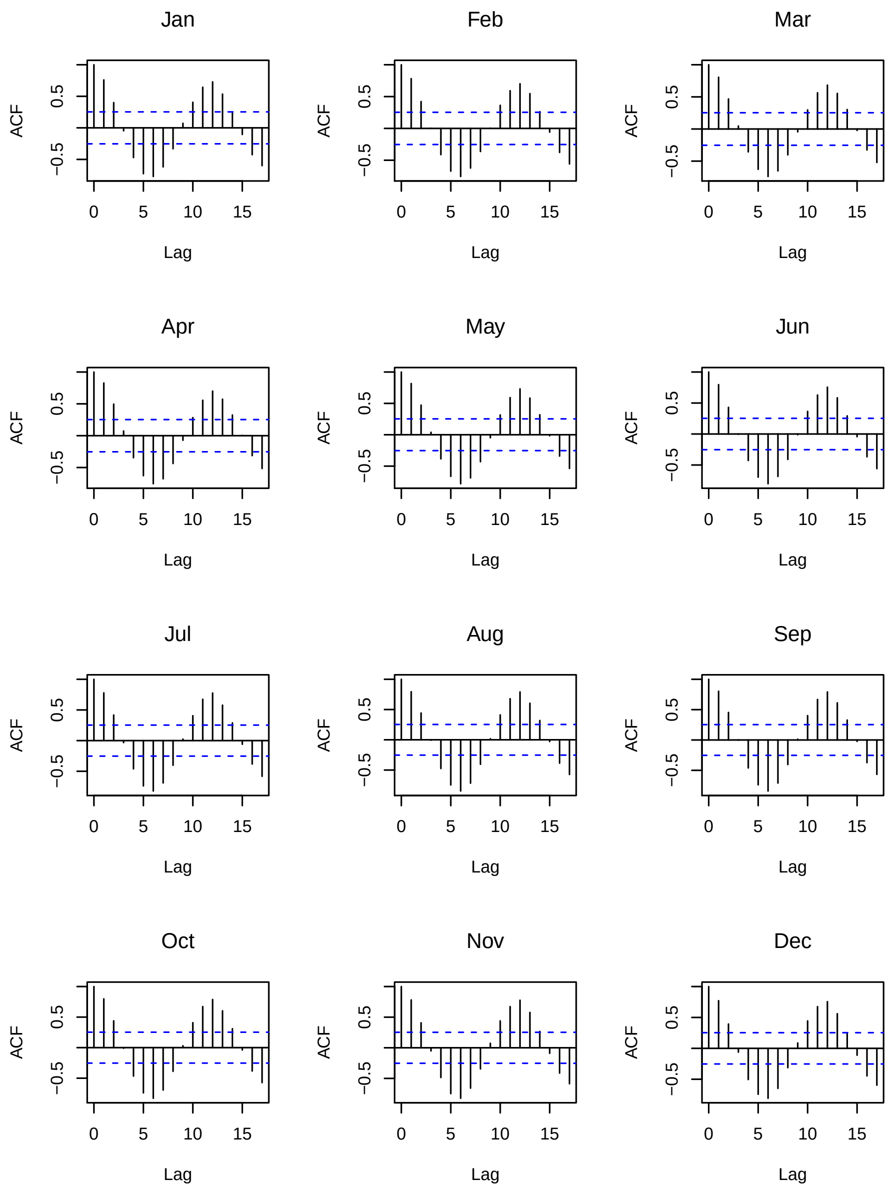

Figure 42: **Autocorrelation Function on (Detrended) New Backward Curve, Crude Oil (F1 $\sim$ F60, 2006 $\sim$ 2017)**

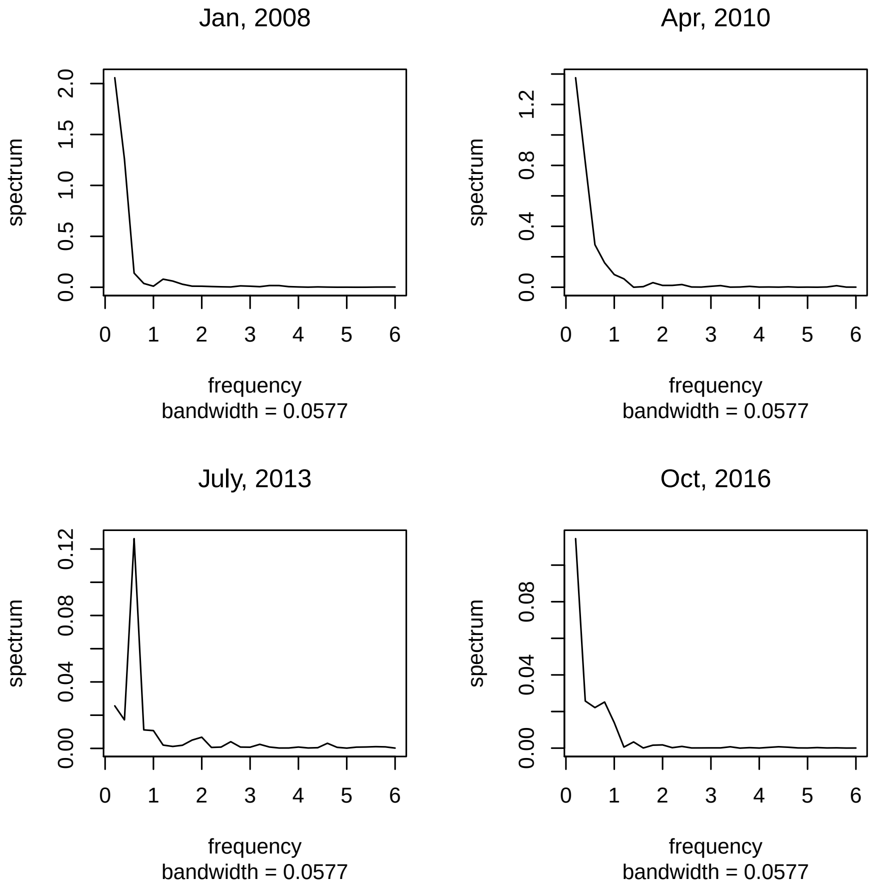

Figure 43: **Power Spectrum on Old Backward Curves of 4 Specific Months, Natural Gas (F1 $\sim$ F60, 2002 $\sim$ 2008)**

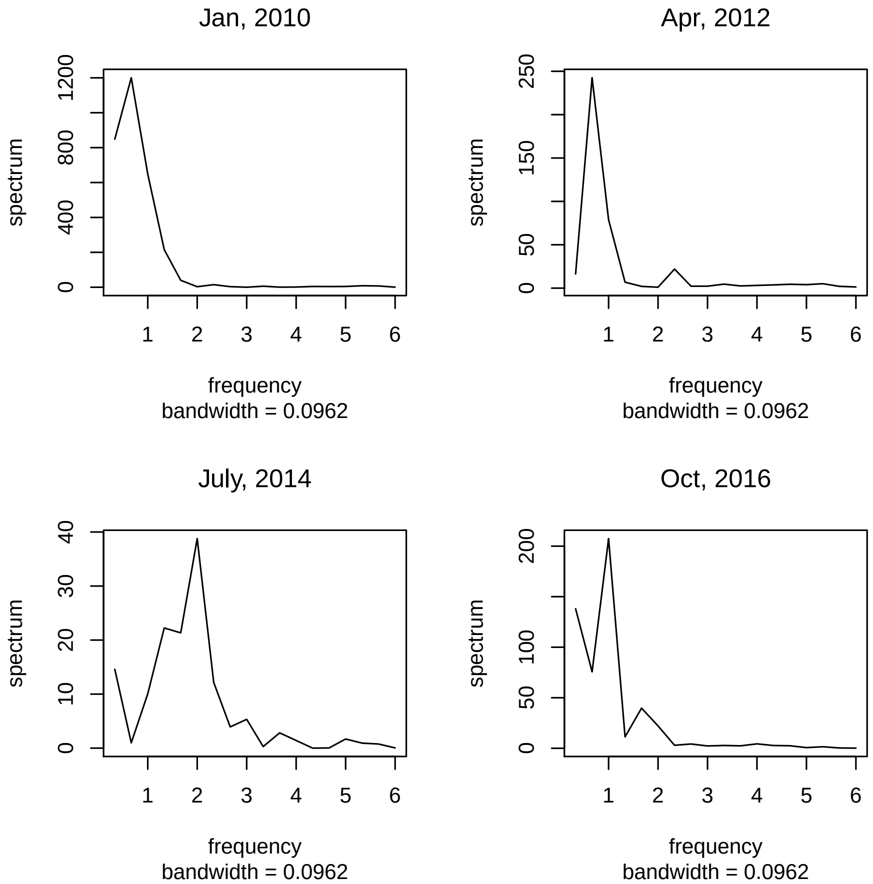

Figure 44: **Power Spectrum on Old Backward Curves of 4 Specific Months, Gasoline (F1 $\sim$ F36, 2007 $\sim$ 2017)**

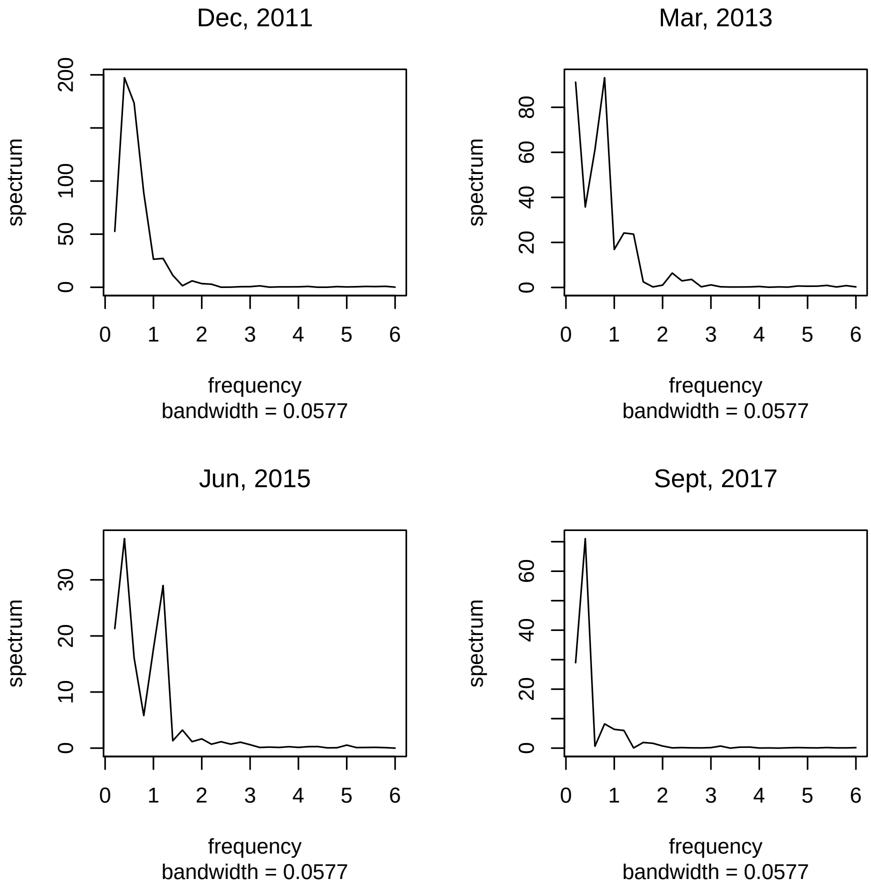

Figure 45: **Power Spectrum on Old Backward Curves of 4 Specific Months, Crude Oil (F1 $\sim$ F60, 2006 $\sim$ 2017)**

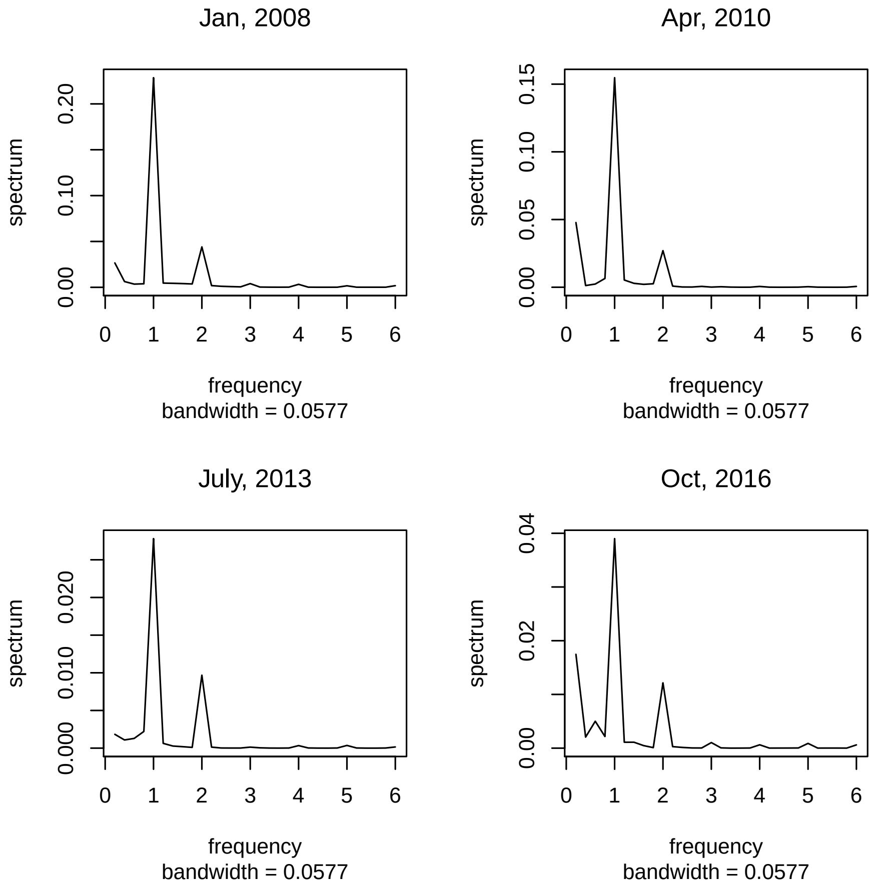

Figure 46: **Power Spectrum on Old Forward Curves of 4 Specific Months, Natural Gas (F1 $\sim$ F60, 2002 $\sim$ 2008)**

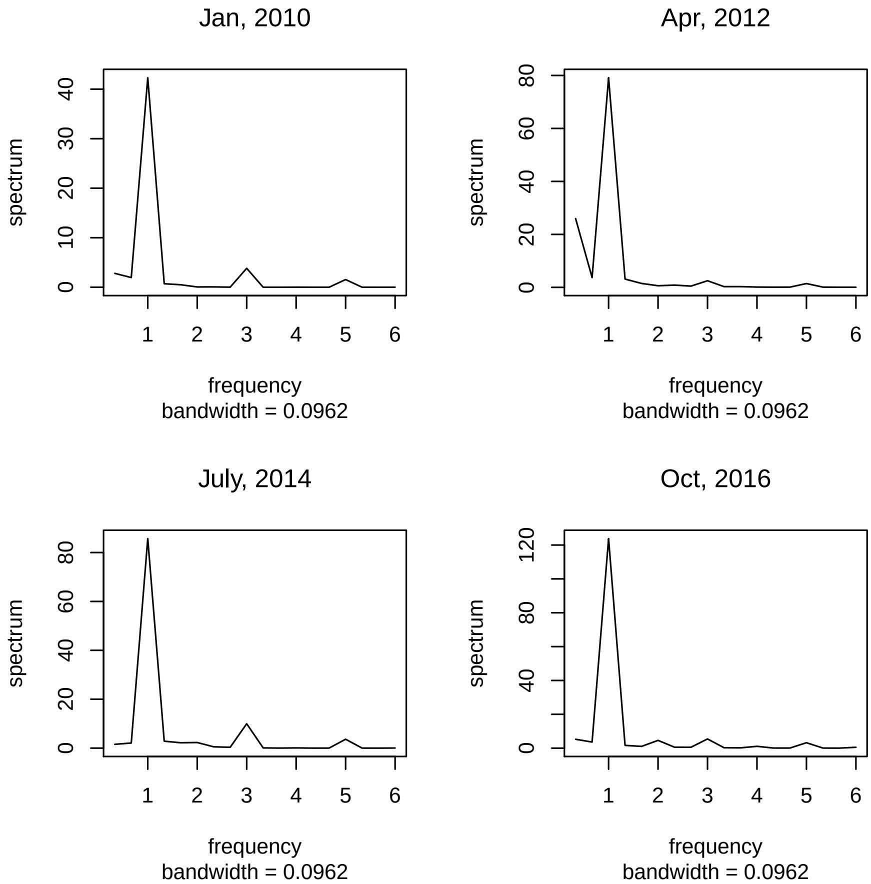

Figure 47: **Power Spectrum on Old Forward Curves of 4 Specific Months, Gasoline (F1 $\sim$ F36, 2007 $\sim$ 2017)**

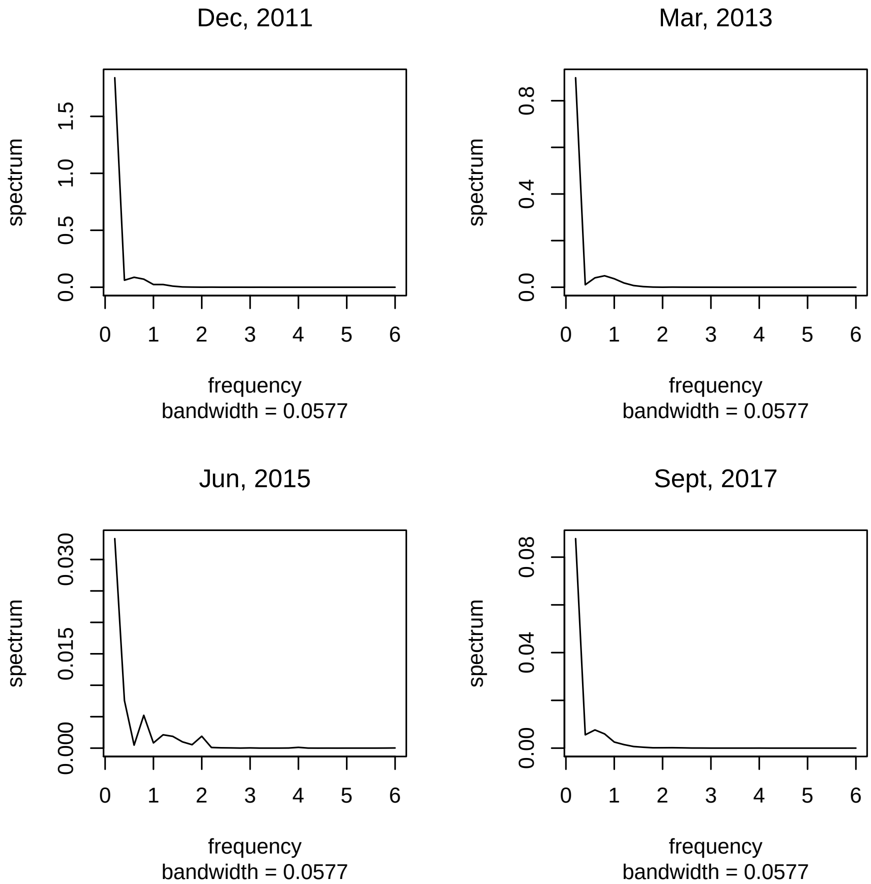

Figure 48: **Power Spectrum on Old Forward Curves of 4 Specific Months, Crude Oil (F1 $\sim$ F60, 2006 $\sim$ 2017)**

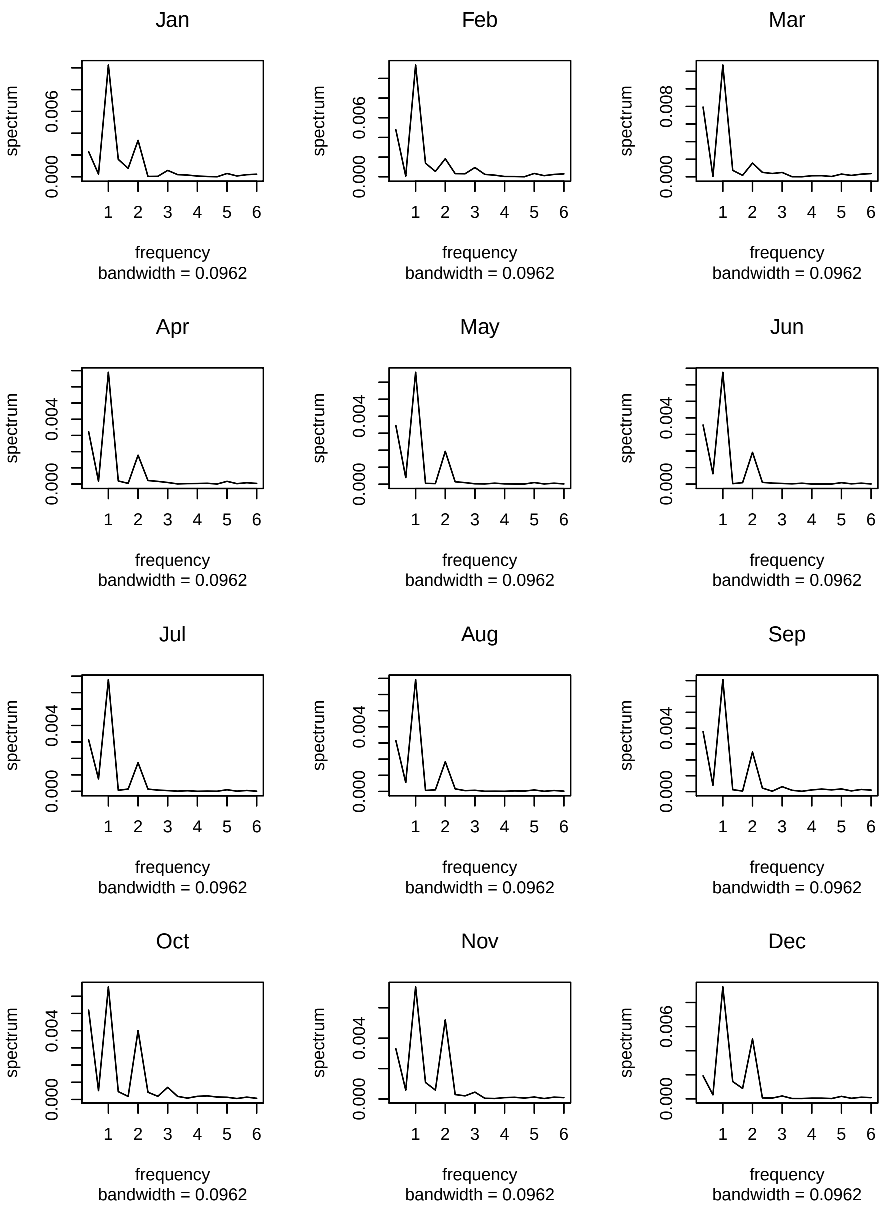

Figure 49: **Power Spectrum on New Backward Curve, Natural Gas, Group 1 (F1 ∼ F36, 1997 ∼ 2017)**

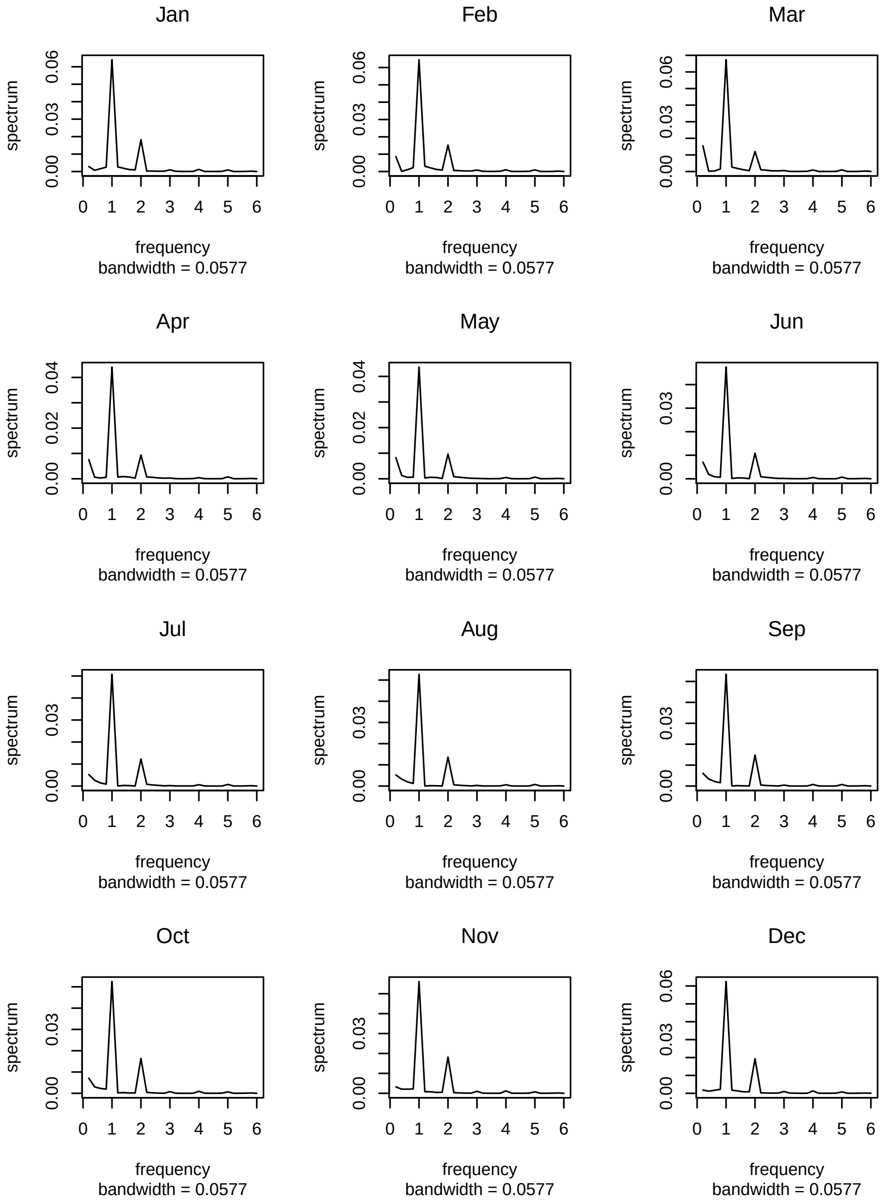

Figure 50: **Power Spectrum on New Backward Curve, Natural Gas, Group 2 (F1 $\sim$ F60, 2002 $\sim$ 2017)**

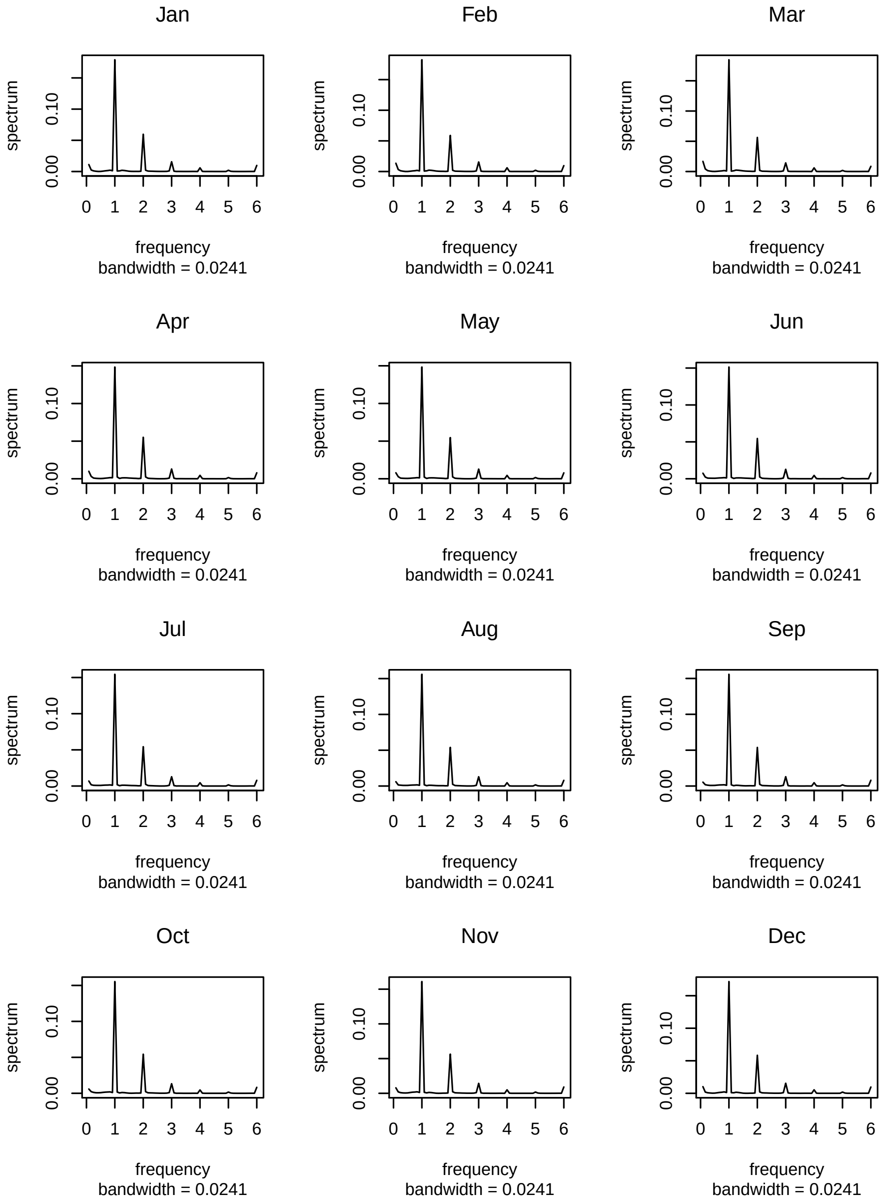

Figure 51: **Power Spectrum on New Backward Curve, Natural Gas, Group 3 (F1 $\sim$ F144, 2008 $\sim$ 2017)**

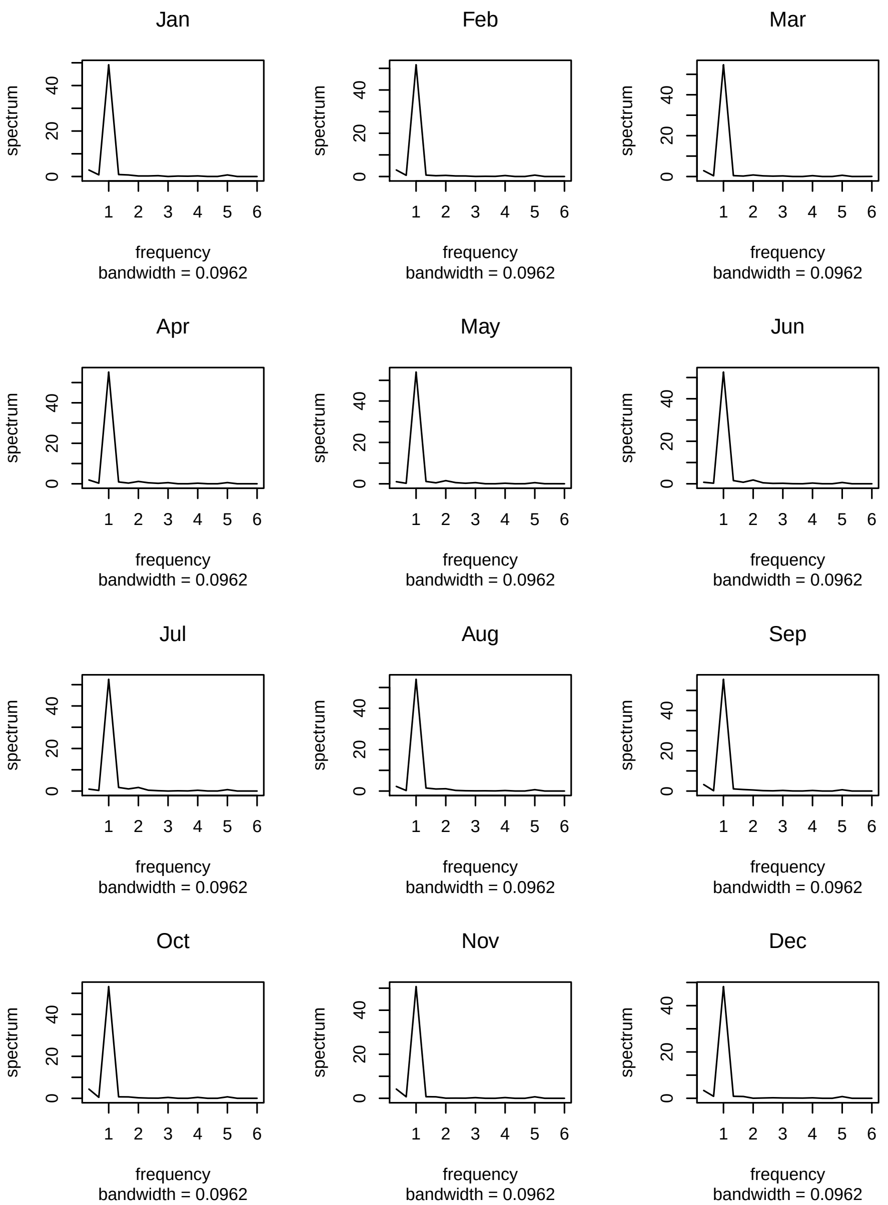

Figure 52: **Power Spectrum on New Backward Curve, Gasoline (F1 $\sim$ F36, 2007 $\sim$ 2017)**

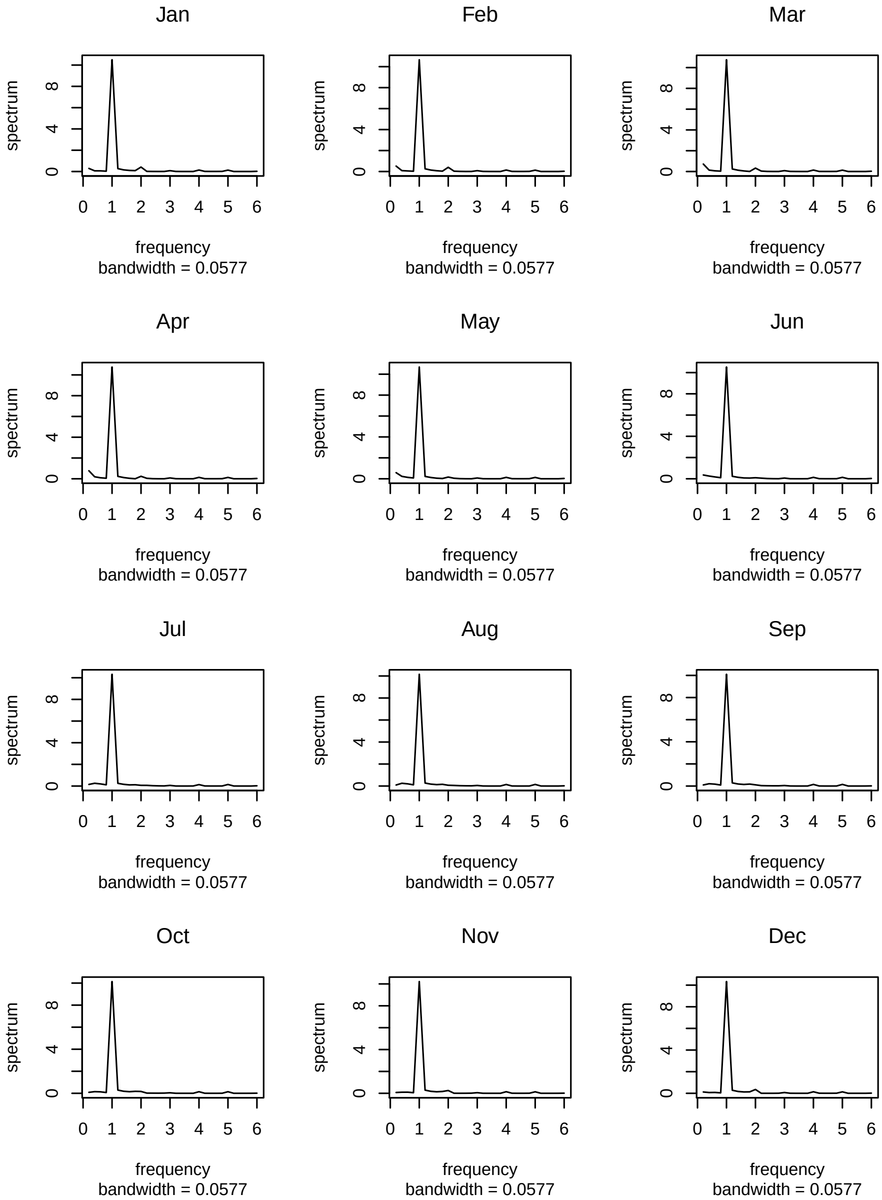

Figure 53: **Power Spectrum on New Backward Curve, Crude Oil (F1 $\sim$ F60, 2006 $\sim$ 2017)**

An Arbitrage Opportunity Based on Seasonality of the Trading Date in the Commodity Market

Abstract

In the last paper, we find solid empirical evidence to suggest the existence of seasonality that relates to the trading date of a future contract in the commodity market, which indicates a potential arbitrage opportunity. In light of this, we decide to construct a trading strategy that is designed specifically to profit from the new seasonal pattern in three commodity markets, namely natural gas, gasoline, and crude oil. The results show promising profit over the long run for all three commodities, with relatively low risks. Then, we establish a model based on the Sorensen (2002) model, with the introduction of an arbitrage factor to capture the trading-date seasonality. We calibrate the model using Kalman filter in the state space form, and the results suggest that the vast majority of the parameters are highly statistically significant in explaining the movement of the futures prices in the three commodity markets.

1 Introduction

The global commodity market has become one of the largest, most important and most active financial market in the world. Billions of dollars worth of commodity futures and derivatives are traded everyday by the market participants. And the market is still growing at incredible speed, to fulfil the never-ending needs of the exploding global population and to build the modern world. Given such a significant role that commodities are playing these days, it is of extreme importance for both the traders and the academics to understand the market and its behaviour over time. With the huge volume of trade that happens everyday, any trading strategy that may lead to reliable and consistent profit could potentially carry a significant value.

To design a profitable trading strategy, it is very important to study and understand the market behaviour and be able to identify its features or patterns first. Of all the features of the commodity market, seasonality is a very common and noticeable one. It has widely been confirmed that seasonality exists in multiple energy commodities such as natural gas and gasoline. The conventional idea believes that the seasonal pattern is closely related to the maturity dates. In other words, the futures prices of a seasonal commodity may significantly be influenced by when it matures. For example, the futures prices of natural gas are higher if the contracts mature in the winter time. This is due to the higher demand for natural gas in the winter for heating purposes. However, the seasonal pattern for gasoline appears to be almost the opposite to natural gas, which indicates a higher price in the summer time, and a lower price in the winter time. The reason is because people seem to prefer travelling during the hot season, which significantly increases the demand for gasoline as fuel. However, there are also other commodities such as crude oil, that presents no seasonality in its price.

A large amount of literature has been devoted to the study of seasonality in the commodity market. Some of the early work include Fama and French (1987), which confirms that seasonality exists in the convenience yield that is closely related to the inventory level of the specific commodity, which is usually subject to seasonal changes of demand or supply. Kramer (1994), on the other hand, studies the interesting January effect in the stock market, and argue that the source could be a seasonality that relates to macroeconomy. In recent years, the commodity market has inspired various pricing models in order to quantitatively measure the futures prices. Based on some early efforts such as Schwartz (1997) and Schwartz and Smith

(2000) that are highly successful but embeds no seasonality, Sorensen (2002) develops a pricing model with a seasonal factor of Fourier form to capture seasonality in the agriculture market. Lucia and Schwartz (2002) examines the electricity prices in the Nordic Power Exchange, and also includes a seasonal component in the form of either some seasonal dummies or a sinusoidal function in their model. Moreover, unlike most of the previous models with deterministic seasonality factor, Mirantes *et al* (2012) and (2013) develop stochastic seasonal components in their model for commodities based on their empirical tests on the data. Furthermore, seasonality has not only been embedded in the prices, but also in other factors. For example, Suenaga and Smith (2011) examines the dynamics of volatility in the energy market and also found significant evidence of seasonality. Shao *et al* (2015) studies risk premium in the natural gas market that is both time-varying and seasonal. Jin *et al* (2012) finds seasonal pattern in the long-term structure of commodity future contracts. Back *et al* (2013) develops a model with seasonal volatility to price the commodity options. The large body of previous literature proves that seasonality plays a significant role in the prices of the commodity market.

One of the application of studying the behaviour of prices in the financial market is to find a trading strategy for market participants to hedge their risk exposures or to profit from it. There has been a large volume of literature about this as well, involving different kinds of trading strategies. One of the most popular ones is based on the mispricing of the same or very similar underlying assets due to their geographical differences. Earlier studies include Brown (1997) that study the arbitrage activities between the London and US silver markets to unravel the relation between sterling and US dollar interest rate. More recent works include Brown and Yucel (2009), which investigate the arbitrage opportunities in the natural gas market in between Europe and America, given the integration between the natural gas price and the crude oil price. Another type of arbitrage involves the asymmetry of new information coming to the market. An example could be Edmans *et al* (2015), where the authors examine the motivation and behaviour of the market participants who trade on information. Their analysis reveals asymmetric effect on the trading behaviour from good and bad news. There are also arbitrage strategies that are designed for a specific type of financial instruments. For example, Duarte *et al* (2007) studies the extremely popular fixed-income strategy in the market that constitutes a large proportion of the trading volume in the market. Five various strategies are discussed in the paper, and their results suggest that the strategies involving more “intellectual capital” seem to generate more alpha-type profit in the market. In summary, combining the knowledge of the financial market and a well-designed and well-executed arbitrage strategy, one could generate considerable and consistent profit from the financial market, at least until the particular strategy becomes gradually known and used by other market players and thus eliminating the remaining profit margin eventually.

However, it is extremely hard to construct a trading strategy based on seasonality of the commodity futures. This is because the maturity dates of all the future contracts are already determined and known to the public when they are introduced to the market to trade. The price of a contract that matures on a specific date may fluctuate due to some changes of other factors during its life time till maturity, such as new information coming or unexpected demand shift, but not for the seasonality factor. Nonetheless, based on the results from our last paper, we believe to have found a new arbitrage opportunity to exploit profits by trading according to seasonality in the commodity market. However, it is not the seasonality that relates to the maturity dates of the future contracts that lays the foundation of our strategy, but instead the seasonality that relates to the trading dates. To be more specific, futures prices are usually higher when being traded in some months of a year, and lower in others, independent of the maturity dates. This, in theory, violates the rule of no-arbitrage in the classic pricing model, and provides a real opportunity of arbitrage in the market. To be specific, we can sell a contract in the months of higher prices, and then buy one with the same maturity date in the months of lower prices, to make a profit. If the two contracts that we have traded in two different months mature at the same time, their prices must converge to the same spot price in the maturity month.

Hence, in this paper, we decide to test our theory by constructing a trading strategy based on the trading-date seasonality in the first part of this paper, and then quantitatively model the arbitrage opportunity in the commodity market in the second part. First, to confirm the practicality of our theory in reality, we build a simple trading strategy of "buy low sell high" in three commodities, namely natural gas, gasoline, and crude oil. In particular, we buy the future contracts in the month of their lowest price, and sell those with the same maturity dates in the month of their highest price. We can profit from the price differences between the contracts of the two trade, if the trading-date seasonality stands. The results of our strategy show both positive and negative performances in different years, but the final outcomes of running the strategy over a long period of time are universally positive across all three commodities. This confirms not only the existence of the trading-date seasonality from the perspective of a market participant, but also the practicality of an arbitrage trading strategy that based on the new seasonality. Secondly, we build a quantitative model to capture the trading date seasonality and calibrate the model with the historical prices of the three commodities. Our model is based on Sorensen (2002) of long-term and short-term stochastic factors as well as a deterministic seasonal component, and embeds our own seasonality factor that relates to the trading date. The empirical results show that the new seasonality factors are mostly significant in explaining the market prices of the three commodities.

The rest of the paper is organized as follows. Section 2 briefly describes the data that we are going to use in this paper. The arbitrage trading strategy that we design to prove the feasibility of profiting from the trading-date seasonality is shown in Section 3. Then, we build a quantitative model to include the new seasonality factor in Section 4, and discuss the methodology to calibrate the model in Section 5. In Section 6, we demonstrate and analyse the results from the empirical study. Section 7 draws the conclusion for our research.

2 The Data

We decide to explore the arbitrage opportunities in three of the commodity markets, namely, Henry Hub natural gas, gasoline, and generic crude oil. All the data are daily observed, obtained from Bloomberg Terminal, and the specific price we use are called "closing price one day ahead". In the majority part of our paper, we denote all the future contracts as F1, F2, F3, ..., where F1 indicates the future contract that is closest to its maturity date, F2 the second closest, so on and so forth. In the natural gas market, from 1997 till 2002, only futures contracts with maturity dates expanding to the next 36 months from the trading dates (F1 $\sim$ F36) were traded. Since 2002, longer term contracts were added. Accordingly, we decide to divide the natural gas data into two groups. The first group includes all the data since 1997 till early 2017, from F1 to F36. The second group includes all the data from 2002 to early 2017, from F1 to F60. As for gasoline and crude oil, the data we use are from 2007 and 2006 respectively, to 2017, with maturity dates up to 36 and 60 months ahead, respectively. A brief description of the data is illustrated in Table 1, where the mean and standard deviation of some selected contracts are shown.

3 A Simple Arbitrage Trading Strategy

To explore the possibility of arbitrage based on the trading-day seasonality in the three commodities, we decide to construct a very simple trading strategy in their future markets. As our previous paper suggests, there is clear evidence that, on average, the futures prices of the three commodities are traded at a higher price in some months of a year, while lower in some other months, regardless of the maturity dates. Accordingly, we can buy a future contract in the "low price" month, sell one with the same maturity date in the "high price" month to make a risk-free profit in theory. The key to the strategy is to pair the contracts that we buy and sell by their maturity dates. To demonstrate the underlying mechanism, let the futures price be denoted here as $F_i(t,T)$, where t and T indicates the trading date and maturity date. A long trade is marked as $i=b$, and a short one as $i=s$. Also, $P(T)$ represents the spot price at maturity. If we assume that the position is closed on the maturity date by a trade of the opposite direction, the payoff of the long position at the closing date, denoted by $V_b(T)$, can be written as $V_b(T) = P(T) - F_b(t_1,T)$, while the payoff of the short order can be written as $V_s(T) = F_s(t_2,T) - P(T)$. Therefore, the collective payoff of both positions will be $V(T) = V_b(T) + V_s(T) = F_s(t_2,T) - F_b(t_1,T)$, which is secured as soon as the second trade of the opposite direction is fulfilled in the market. Hence, if our theory of the trading-date seasonality stands, or $F_s(t_2,T) > F_b(t_1,T)$ consistently in the long run, we are almost guaranteed to make a profit.

The implementation of the strategy in reality is described as follows. First, to simplify the strategy, we decide to trade in only two months of each calender year. We buy in the month of the lowest price, and sell short in the month of the highest price, in each year. As a result, we view one year as a single period during which we conduct two trades of opposite directions. According to our analysis in the last study, the three commodities present different months of highest and lowest prices. Consequently, for natural gas futures, we buy in February and then sell in June. In the cases of both gasoline and crude oil, we short first in July, and then take long position in December. Each contract traded in one month must match another one in the other trading month with the same maturity date.

In either of the trading month in a single period, we do not buy or sell those contracts traded in the market that cannot be paired. To explain in detail, assume that we sell short gasoline futures in July, and go long in December of the same year. According to the data description in the last section, there are 36 contracts that we can trade everyday, from F1 to F36. Hence, when we sell in July, the F1 to

F5 contracts will mature before we can engage in the opposite trade in December. Therefore, they cannot be paired with any contract in December that has the same maturity date. Also, when we buy the gasoline futures in December, the F32 to F36 contracts will not have any short contracts sold back in July to match with the same maturity date. In other words, in the case of gasoline, only contracts of F6 to F36 are sold in every July, and F1 to F31 are bought in every December, from 2007 to 2017. For crude oil, F6 to F60 are sold in every July, and F1 to F55 are bought in every December, from 2006 to 2017. As for natural gas, we have two groups of its futures prices. In group 1, we buy F5 to F36 in February, and sell F1 to F32 in June, from 1998 to 2016. In group 2, F5 to F60 are bought in February, and F1 to F56 are sold in June, from 2002 to 2016. The details of traded contracts for each commodity in one period are listed in Table 2.

To explain the payoff of each year's operation in mathematical formulas, let the futures prices be denoted as ${F_i}^y(t,T) = {F_i}^y(t_j, t_j + K)$ here, where i here remains to represent a buy or sell order, y indicates the specific year of the trade, and $j = 1$ means the first trade of that year, and $j = 2$ the second one of the same period (remember that for natural gas, we buy first and sell later in a single period, while for gasoline and crude oil, we sell first and buy later), and K is the time till maturity. Hence, the payoff of year y for each commodity group can be written as follows:

$$NG1 : V(y) = N * [\sum_{K=1}^{32} {F_s}^y(t_2, t_2 + K) - \sum_{K=5}^{36} {F_b}^y(t_1, t_1 + K)] \tag{1}$$

$$NG2 : V(y) = N * [\sum_{K=1}^{56} {F_s}^y(t_2, t_2 + K) - \sum_{K=5}^{60} {F_b}^y(t_1, t_1 + K)] \tag{2}$$

$$XB : V(y) = N * [\sum_{K=6}^{36} {F_s}^y(t_1, t_1 + K) - \sum_{K=1}^{31} {F_b}^y(t_2, t_2 + K)] \tag{3}$$

$$CL : V(y) = N * [\sum_{K=6}^{60} {F_s}^y(t_1, t_1 + K) - \sum_{K=1}^{55} {F_b}^y(t_2, t_2 + K)]\ , \tag{4}$$

where N represents the equal trading volume of each contract of either direction in every period (in our demonstration below, we set $N = 100$).

Table 3 illustrates the results of our trading strategy by the year of operation for the three commodities. To simplify the discussion, we assume no transaction cost for our entire operation. One could easily apply the relevant expenses according to real world scenario. The numbers above the lowest column of Total represent the payoff of all the contracts traded in the two months of that specific year. First, it is easy to see that the strategy does not work in every year during our observation period. It appears to generate profit more consistently for natural gas, since there have clearly been more profitable years than negative ones. Out of the 19 years of trading for natural gas group 1 and 15 years for group 2, only 7 and 4 years end up with a loss, respectively. However, the numbers seem to reveal less promising annual outcomes for gasoline and crude oil based on the performance of individual year, where only 4 and 6 out of the 11 and 12 years of the observation period lead to positive results. Nevertheless, the total payoffs of the entire operation period for all three commodities result in profits, which means that the earnings from the profitable years can more than compensate for all the losses in the losing years.

Secondly, it is also worth noting that the annual payoff of all three commodities seem to fluctuate significantly. It is interesting to note that the most profitable year for all three commodities is 2008, when the payoff appears to dwarf that of other periods, contributing to a very large percentage of the total payoff at the end of the observed period. On the other hand, it is rare to see a single year of extremely large losses that is comparable to the unusually large profit in 2008 in any of the three commodities. Furthermore, it is also very interesting to calculate the total payoff minus the profit from 2008. The result is still positive for both of the natural gas groups and crude oil, but not for gasoline, which suffers a fairly small loss. Consequently, we argue that our trading strategy of buying and selling according to the seasonality of trading dates can result in very promising outcomes in the long run.

It is also worth studying if the different contracts that are traded each year may have any different performances according to our trading strategy. To do so, we calculate the average profit generated from each pair of the traded contracts over the entire observation period. The calculation is conducted as follows. For each commodity, in a single period, there are a total of M contracts being bought or sold, or M pairs of contracts matched by their maturity date, respectively ($M = 32$ for natural gas group 1, $M = 56$ for natural gas group 2, $M = 31$ for gasoline, and $M = 55$ for crude oil). Each pair of matched contracts has a unique time till maturity since the second trade of every year, denoted as $m,\ m = 1, 2, \ldots, M$. If we denote each pair in year y as $PR(m, y)$, we would like to see how the value of $\frac{\sum_{y=1}^{Y} PR(m,y)}{Y}$ changes with respective to m, where Y is the latest year of our observation period.

We present the results in Figure 1, where the x-axis represents m, or time till maturity since the second trade of every year. First, it is easy to observe a general trend of decreasing payoff with respective to the time till maturity for natural gas, and especially for crude oil. In other words, the contracts of longer term to maturity appear to generate lower profit, compared to those that mature sooner in the future. Nevertheless, it is not entirely true for natural gas, as the contracts with the shortest term to maturity perform rather poorly. On the other hand, such trend is hard to confirm for gasoline.

Next, an obvious seasonal pattern can be seen in both graphs of natural gas. However, although the x-axis indicates the time till maturity for different pairs of contracts since the second trade in a period, the peaks and troughs of the two graphs of natural gas in Figure 1 are mostly inconsistent with the seasonality of the futures prices that relates to the maturity dates. In other words, the maturity month of the pair of contracts that generates the highest profit in our trading strategy is not in the winter time when the natural gas prices are usually higher. Regardless, from the perspective of the speculators, it would still be logical to spend larger amount of money to invest in the natural gas contracts according to the wave pattern shown in the upper two graphs of Figure 1, and also avoid the contracts with the shortest time till maturity.

The graph of gasoline shows a different pattern, which carries only two obvious peaks of all the 31 contracts. Hence, the peaks and troughs of the payoff from different pairs of gasoline contracts seem to have no relationship with the maturity-date seasonality of its futures prices. Still, traders are able to earn higher profit according to the graph of gasoline in Figure 1, when larger proportion of the investment should be put in those contracts with higher payoff. The graph of crude oil in Figure 1 shows an unusually smooth curve. Since it also demonstrates an obvious trend of diminishing return with respective to time till maturity, it is definitely more profitable for traders to engage in those contracts with closer maturity dates.

After the analysis of the potential profit from our strategy, the next step is to study its risks. One noticeable feature of our strategy is that although the long and short positions of each year are taken with the same quantity of future contracts that share the same maturity date, the two trades do not take place at the same time, but several months apart. This may imply a certain level of risk, since for a short window of time (4 months for natural gas, and 5 months for gasoline and crude oil), our initial position after the first trading month in each period is not hedged by any trade of the opposite direction. However, we believe that the risk exposure during such short period of time is fairly limited. To investigate in depth, we decide to calculate the balance in the margin account during the risky period. As a unique feature of the future market, the floating gain or loss of a position of a future contract is marked to market everyday. Remember that in the case of natural gas, we buy first in February, and sell then in June of each year. For gasoline and crude oil, we sell first in July, and buy then in December of each period. Denote $\Delta B_i(t^*)$ as the change of balance in the margin account at current time t^*, $t_1 < t^* < t_2$, where $i = b$ to indicate a buying month as the first trading month of a year (such as the case of natural gas), and $i = s$ a selling month first (such as the cases of gasoline and crude oil). Assume that we have M open positions from time t_1 to t^*, then

$$\Delta B_b(t^*) = N * \sum_{j=1}^{M} [F_b^y(t^*, t_1 + K) - F_b^y(t_1, t_1 + K)] \,, \tag{5}$$

and

$$\Delta B_s(t^*) = N * \sum_{j=1}^{M} [F_s^y(t_1, t_1 + K) - F_s^y(t^*, t_1 + K)] \,. \tag{6}$$

Table 4 illustrates the lowest balance in the margin account during the risky period between the first trade and the second trade of each period. The initial balance of the margin account before the first trade is assumed to be zero, which means that we either withdraw the profit or compensate any loss in the margin account during the no-risk period when all the open positions are closed. We would like to point out two observations from Table 4. First, the floating losses in the margin account only occurs in about half of the years during our trading operations. The situation is slightly more severe for gasoline and crude oil. On the other hand, it also means that the margin account never drops below zero in the other years during the risky period. Second of all, even the biggest floating losses that we ever have to bear during these risky months, shown in the bottom column of Table 4, is relatively small, comparing to the average payoff from running the trading strategy for the entire observation period in Table 2. Hence, we argue that the amount of risk to execute our trading strategy during our sample time is fairly limited.

4 A Quantitative Model to Measure Arbitrage

In the last section, we are able to use a simple and straightforward trading strategy to confirm that the arbitrage opportunity based on the trading-date seasonality exists and can generate rather considerable profits with limited risk. However, we would like to take a step further to quantitatively measure the arbitrage opportunity. Hence, we decide to establish a mathematical model based on the Sorenson (2002) model to help identify and analyse how the new seasonality factor affects the prices in the future markets of the three commodities. The model of Sorensen (2002) has widely been cited in a large amount of literature, which derives the closed-form formula for futures prices based on the spot price dynamics, which consists of two stochastic factors and a seasonality factor. When the seasonality factor equals to zero, the model reduces to the model of Schwartz and Smith (2000), which is a variation of Schwartz (1997) 2-factor model with stochastic spot price and convenience yields.

According to Sorensen (2002) model, there are three factors that constitute the logarithm of the spot price at time t, denoted as $p(t) = log(P(t))$, namely, the two stochastic state variables of a long-term and a short-term factor, and a seasonality factor. The idea of the long-term and short-term factors are borrowed from Schwartz and Smith (2000), where the long term factor represents the level of the equilibrium price in the long run, while the short-term factor captures the unexpected deviation from the equilibrium price in the short run. The seasonality factor allows the spot price to be able to exhibit the seasonal pattern that is present in a large number of commodities, including energy and agricultural commodities. The mathematical model to represent the spot price and the three factors are as follows:

$$p(t) = x(t) + y(t) + S(t) \ , \tag{7}$$

where

$$dx = (\mu - \frac{1}{2}\sigma_x^2)dt + \sigma_x dW_1 \ ,$$
$$dy = -\kappa y dt + \sigma_y dW_2 \ ,$$

and

$$S(t) = \sum_{k=1}^{K}(\gamma_k cos(2\pi kt) + \gamma_k^* sin(2\pi kt)) \ .$$

Equation (7) describe the stochastic processes of the long term effect on the spot price, where e^x is expected to grow at a constant rate of μ with a volatility of σ_x. y represents the short-term deviation from the equilibrium price, which follows a Urnstein-Uhlenbeck process with zero mean and κ as the speed of mean-reversion. W_1 and W_2 are two standard Brownian motions correlated with coefficient ρ. $S(t)$ indicates the seasonal component in the spot price, with γ_k and γ_k^* as the coefficients. Since the seasonality of the three commodities in our paper is on an annual basis, we set K here to be 2, which is the same as in Sorensen (2002). If γ_k and γ_k^* are 0, the model reduces to Schwartz and Smith (2000). In this paper, the interest rate is assumed to be constant. However, a stochastic interest rate is fairly easy to embed.

To price the corresponding futures prices, we introduce the risk-neutral measure $\mathbb{Q}^*$, under which both of the stochastic factors must carry risk premiums, denoted as λ_x and λ_y. This is because neither of them are tradable assets that can be hedged to diversity their specific risks. If we let $\mu^* = \mu - \lambda_x$, the two state variables can be written as:

$$dx = (\mu^* - \frac{1}{2}\sigma_x^2)dt + \sigma_x d\widetilde{W}_1 \ , \tag{8}$$

$$dy = -(\lambda_y + \kappa y)dt + \sigma_y d\widetilde{W}_2 \ , \tag{9}$$

where the two Brownian motion $\widetilde{W}_1$ and $\widetilde{W}_2$ under the risk neutral measure $\mathbb{Q}^*$ are correlated with coefficient ρ. The seasonal factor, $S(t)$, remains the same.

Under the rules of risk-neutral pricing, the futures prices are calculated as the expectation of the spot price at maturity date, conditional on all the information at current time, under $\mathbb{Q}^*$. Hence, let $F(t,T)$ represent the future price at current time t that will mature at time $T > t$, then $F(t,T) = \mathbb{E}_t^Q(P(T))$, where $\mathbb{E}_t^Q$ is the operator of conditional expectation under risk neutral measure $\mathbb{Q}^*$, conditional upon all the information until time t. It is easy to see that $P(t)$ here is log-normal, which lead to the following solution of $F(t,T)$:

$$F(t,T) = exp[S(T) + x(t) + y(t)e^{-\kappa(T-t)} + A(T-t)] \ , \tag{10}$$

where

$$A(T-t) = \mu^*(T-t) + \frac{\lambda_y - \rho\sigma_x\sigma_y}{\kappa}(1 - e^{-\kappa(T-t)}) + \frac{\sigma_y^2}{4\kappa}(1 - e^{-2\kappa(T-t)}) \ .$$

It is very easy to observe how the maturity-date seasonality, $S(T)$, affects the futures price.

Nevertheless, the above model of Sorensen (2002) is based on no-arbitrage pricing, which is not entirely true in the three commodity markets that we have studied in the last section. The reality is that the existence of the trading-date seasonality obviously violates the rule of no arbitrage in the classic pricing model of financial instruments. Hence, we would like to build a model based on Sorensen (2002), that also acknowledges and encompasses the arbitrage opportunity in the real market that is generated by the trading-date seasonality. However, one of the complications to include the trading-date seasonality is that the spot price must only include one seasonality factor, since its trading date and maturity date is exactly the same by definition. Hence, the futures prices derived as the conditional expectation of the spot price at the maturity date cannot include the second seasonality in any way. As a result, to include the trading-date seasonality in our model, we decide to introduce an arbitrage factor that influence the futures prices derived from risk-neutral pricing, or

$$F^*(t,T) = F(t,T) * e^{\pi(t)} \ , \tag{11}$$

where the arbitrage factor $e^{\pi(t)}$ follows

$$\pi(t) = \sum_{k=1}^{K} (\widehat{\gamma_k} cos(2\pi kt) + \widehat{\gamma_k^*} sin(2\pi kt)) \ . \tag{12}$$

It is easy to see that $\pi(t)$ is a sum of trigonometric functions of the trading date t, independent of the maturity date. Both $\widehat{\gamma_k}$ and $\widehat{\gamma_k^*}$ represents the coefficients for the trading-date seasonal factor. Since the second seasonality is also repetitive on an annual basis, we set K in equation (12) as 2. Therefore, $F^*(t,T)$ embeds both the seasonality of maturity dates and trading dates, where either is independent of another. When both $\widehat{\gamma_k}$ and $\widehat{\gamma_k^*}$ are zero, the model reduces to the arbitrage-free model of Sorensen (2002).

5 Calibration Using the State Space Approach and Kalman Filter

It is easy to see that the spot price is an affine function of the three components, namely the long-term factor, the short-term factor, and the maturity-date seasonality factor, and that the logarithm of the futures price is also linear to all of them plus the trading-date seasonality factor. As a result, it is appropriate to calibrate the model with the state space approach, where the state variables are the two unobservable stochastic factors, $x(t)$ and $y(t)$, that are generated by two standard Brownian motions. Then, after the model can be transformed into the state space formulas, the calibration can be conducted by using the Kalman filter to estimate the unknown parameters in the model based on maximum likelihood.

We first briefly review and justify our use of Kalman filter in the calibration process. The detailed description of the method can be found in Harvey (1989). Kalman filter was introduced by Dr. R.E. Kalman in 1960 in an attempt to solve the problem of linear filtering of time series in engineering. It is a recursive method that has the capacity to include all the available information at the current time for the estimation of the state variables. As long as the equations of the state variables are linear and the residuals can be assumed to be Gaussian, which our model follows, Kalman filter can be used to generate the optimal outcomes.

To implement the Kalman filter, the state space form includes two equations, namely the transition equation and the measurement equation. The transition equation represents the evolution of the unknown state variables following their predetermined stochastic processes in discrete time, or equation (7) in our model. Let $\Delta t = t_n - t_{n-1}$ be the time distance between two observations, where $n = 1, 2, \ldots, N$, with t_N being the maturity date. Then, the transition equation can be written as:

$$X_t = b + BX_{t-1} + w_t \ , \tag{13}$$

where

$$X_t = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}, \quad b = \begin{pmatrix} \mu - \frac{1}{2}\sigma_x^2 \\ 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & e^{-\kappa\Delta t} \end{pmatrix} \ ,$$

$$\eta = \begin{pmatrix} \sigma_x^2 \Delta t & \frac{\sigma_x \sigma_y \rho}{\kappa}(1 - e^{-\kappa\Delta t}) \\ \frac{\sigma_x \sigma_y \rho}{\kappa}(1 - e^{-\kappa\Delta t}) & \frac{\sigma_y^2}{2\kappa}(1 - e^{-2\kappa\Delta t}) \end{pmatrix} \ , \tag{14}$$

where η here is the variance-covariance matrix of the normally distributed, serially uncorrelated residual, w_t. It is also worth noting that $\mathbb{E}(w_t) = 0$.

However, the transition equation only constitutes half of the state space form. Since both of the state variables are unobservable, the transition equation could only simulate the time series of the state variables. Without anything observable references to confirm their values from time to time, it could only lead to unreliable results. Hence, the measure equation is required to complete the state space form, which governs how the state variables are related to the observable element in the model. Since the futures prices can be directly observed in the three commodity market, and we have established the relationship between the futures price, $F^*(t, T)$, and the two state variables, $x(t)$ and $y(t)$, we can form the measurement equation according to equation (10) to (12):

$$Y_t = c_t + C_t Y_t + v_t \ , \tag{15}$$

where

$$Y_t = \begin{pmatrix} log(F(t, T_1)) \\ \vdots \\ log(F(t, T_N)) \end{pmatrix} \ , \quad c_t = \begin{pmatrix} S(T_1) + A(T_1 - t) + \pi(t) \\ \vdots \\ S(T_N) + A(T_N - t) + \pi(t) \end{pmatrix} \ ,$$

$$C_t = \begin{pmatrix} 1 & e^{-\kappa(T_1-t)} \\ \vdots & \vdots \\ 1 & e^{-\kappa(T_N-t)} \end{pmatrix} \quad , \quad \epsilon(t) = \sigma_v^2 I_{(t*t)} \ .$$

Here, $F(t, T_n)$, $n = 1, 2, \ldots, N$ indicates the prices of the different future contracts that are traded on any day in the market. v_t is the normal uncorrelated errors to represent the disturbance for the measurement equation, with zero mean and $\epsilon(t)$ as the covariance matrix, where $I_{(t*t)}$ is the identity matrix of dimension $t * t$. σ_v is the standard deviation, or the measurement disturbances. The simple structure of ϵ is exactly the same as that in Sorensen (2002) and Schwartz and Smith (2000).

By running the Kalman filter based on a given set of parameters, we are able to not only obtain the optimal estimates of the unobservable state variable, but also the distribution of the state variables. Denoting $P_{t_1|t_2}$ and $Q_{t_1|t_2}$ as the prediction of the mean and covariance of the state variables of t_1 from t_2 ($t_1 \geq t_2$), we have the following updates for the distribution of the state variables at each time step:

$$P_{t|t-1} = \mathbb{E}[(x_t, y_t)'|t-1] = b + BP_{t-1|t-1} \ , \tag{16}$$

$$Q_{t|t-1} = \mathbb{V}[(x_t, y_t)'|t-1] = BQ_{t-1|t-1}B' + \eta \ , \tag{17}$$

where $'$ indicates the transpose operator of any vector or matrix. We can also obtain the predicted mean and covariance of the futures prices at time t_1, given all the information at time t_2:

$$U_{t|t-1} = \mathbb{E}[(F(t, T_1), \ldots, F(t, T_N))'|t-1] = c_t + C_t P_{t|t-1} \ , \tag{18}$$

$$V_{t|t-1} = \mathbb{V}[(F(t, T_1), \ldots, F(t, T_N))'|t-1] = C_t P_{t|t-1} C_t' + \epsilon \ . \tag{19}$$

Then, by constructing a so-called Kalman gain of the form $K_t = P_{t|t-1} C_t V_{t|t-1}^{-1}$ as a correction factor to $P_{t|t-1}$, we can have the mean and variance of the two state variables based on all the information available at time t:

$$P_{t|t} = \mathbb{E}[(x_t, y_t)'|t] = P_{t|t-1} + K_t(Y_t - U_{t|t-1}) \ , \tag{20}$$

$$Q_{t|t} = \mathbb{V}[(x_t, y_t)'|t] = Q_{t|t-1} - K_t V_{t|t-1} K_t' \ . \tag{21}$$

See Harvey (1989) for detailed derivation and explanation of equations (16) - (21). Since the state variables as well as the logarithm of the futures prices follow normal distribution, and now we know the expression for the mean and variance of the futures prices, it is fairly easy to obtain the likelihood function, and then find the optimal estimation for the parameters that maximize the likelihood function.

6 Empirical Results

We calibrate our model on the data of the four groups of three commodities. $\Delta t = 1/260$, which means that we assume to have 260 trading days in a year. The interest rate is set at 3%. In each group, we only use some selected contracts to calibrate our model. For both natural gas group 1 (1997 $\sim$ 2017) and gasoline (2007 $\sim$ 2017), we use a total of 6 contracts, namely F1, F7, F13, F19, F25 and F31. For both natural gas group 2 (2002 $\sim$ 2017) and crude oil (2006 $\sim$ 2017), we use 9 contracts in total, namely F1, F8, F15, F22, F29, F36, F43, F50 and F57. A brief description of their statistics are listed in Table 5, in which it can be seen that the standard deviation of the future contracts seem to decrease with the length of time till maturity, which is consistent with the Samuelson effect. In the calibration process, the same set of parameters are estimated for each group of the commodities by the methodology described in the above section, with the exception of crude oil, where $\gamma_k = 0$ and $\gamma_k^* = 0$, $k = 1, 2$ from equation (7). This is because crude oil does not carry conventional seasonality, as proved in our last paper.

Table 6 to 9 shows the calibrated results of our model. We list the parameters that are calibrated for each group of commodity, their calibrated values, the corresponding standard deviation and the t value. We also show the level of significance in the tables. It can be seen that in each case, the majority of the parameters are significant, though a limited some are not significantly different from 0. First of all, in all cases, κ, σ_x and σ_y are highly significant. According to the model, κ indicates the speed of mean reversion. Our result is thus consistent with previous papers such as Schwartz and Smith (2000) and Sorensen (2002) that embed the feature of mean reversion in the futures prices. However, the values of κ differ among different commodities. The price of gasoline has the highest speed of mean reversion, while that of crude oil is much lower. It is also interesting to note that κ presents different values for natural gas group 1 and 2, given the different sample period and lengths of time till maturity of the contracts in the two groups.

Volatilities for both of the stochastic factors are highly significant for all three commodities as well. In the case of natural gas, the results from both groups obviously show a much larger volatility of the short-term factor than the long-term equilibrium factor, or $\sigma_y > \sigma_x$. This is consistent with Sorensen (2002) of the study of agriculture commodities. For both gasoline and crude oil, however, the difference between the volatilities of the two factors is very small.

μ_* captures the growth rate of the long-term factor. The value is positive for both groups of natural gas and crude oil, but negative for gasoline. This indicates a decline of price in the long run for gasoline during the observation period. Also, given the solution of the futures prices of equation (10), negative μ^* could mean a negative relationship between the futures price and time till maturity, $(T-t)$. This is consistent with the statistics shown in Table 1 and Table 5, where gasoline futures prices obviously drops with the time to maturity, while no other two of the three commodities show such trend.

λ_y is the risk premium for the short-term disturbances in price. It is negative as well as significant for natural gas group 1, group 2 and crude oil. The value is also very close to each other for the two commodities. However, we fail to reject the null hypothesis of $\lambda_y = 0$ for gasoline. ρ indicates the relation between the two Brownian motions of the two stochastic factors. They are both negative, around -0.3, for the two natural gas groups, but positive in both cases of gasoline and crude oil. In the past papers of commodities, it is easy to see that both risk premium and coefficients of the Brownian motions could vary significantly in different commodities or different sample periods (see Schwartz (1997), Sorensen(2002), among others).

The four seasonality factors that relate to the maturity dates, or γ_1, γ_1^*, γ_2, and γ_2^*, are mostly highly significant for the two natural gas groups and gasoline. This is very strong evidence to suggest the maturity-date seasonality, which is consistent with the past papers. Also, the majority of all four parameters that links to the trading-date seasonality, or $\widehat{\gamma_1}$, $\widehat{\gamma_1^*}$, $\widehat{\gamma_2}$ and $\widehat{\gamma_2^*}$, are significant. They also share similar values to the seasonality parameters of maturity dates, which indicates that the trading-date seasonality influences the futures prices to a very similar extent as the maturity-date seasonality. As a result, the new seasonality component as an arbitrage factor that we introduce in our model is able to join the other factors from the previous model to play its own independent and important role in explaining the price behaviour in the commodity market. Given that our model could capture both seasonality factors while inheriting the merits of previous studies, we would argue that it could largely improve the prediction of market prices, as well as our understanding of the complicated commodity market, and thus help market participants engage in both hedging and speculating in the commodity markets with extra confidence.

7 Conclusion

In this paper, we further confirm the existence of the new seasonality pattern that we find in the last paper. We first build a trading strategy based on the new seasonal pattern, which leads to positive total returns across all our three commodities during our observation years. Then, a quantitative model is established to capture the new seasonality factor. The model is based on Sorensen (2002), and embeds a new deterministic seasonality factor as an arbitrage factor that solely depends on the trading date. We calibrate the model using the state space approach, with Kalman filter to project the unobservable state variables. The calibrated parameters are then presented, with most of them prove to be statistically significant in explaining the futures prices of the commodities. Thus, we argue that our research not only confirms the practicality of using the newly discovered seasonal pattern to profit from the market by a simple but effective trading strategy, but also quantitatively measures its impact on the futures prices.

However, our paper is not without its own shortages. Firstly, our trading strategy of "buy low sell high" is arguably very rudimentary. Although it proves the existence of the trading date seasonality from a practical point of view, it faces several potential issues, such as the transaction cost and liquidity risk. This is especially true when a large amount of future contracts are being traded and need to be hedged later in the same year. The transaction fees to conduct such an operation on a large scale could be potentially huge. Also, since our trading strategy requires a second trade to perfectly hedge the position of the first trade, liquidity in the market at the second trade could pose a real threat to our operation. However, the shortage of our trading strategy can also be improved in different ways, one of which is to buy and sell specific contracts according to their performance in terms of profitability as illustrated in Figure 1. Since there are some contracts that yield higher return than the rest over our observation period, we can choose to trade only these contracts instead of all of them, which could significantly lower the transaction costs as well as the potential liquidity risk.

Furthermore, we do not have a satisfyingly abundant sample of data to play with at the time of this paper. The historical prices of two of the commodities in this paper, gasoline and crude oil, only dates back slightly over 10 years. Hence, we do not know how the trading strategy will perform in a longer run. Nevertheless, we are eager to find out how the trading strategy as well as the arbitrage model perform when data of longer time becomes available. We are also curious if the same seasonality of trading date also exists in other commodity markets. Hence, we appeal for further investigation to explore this phenomenon.

References

Back, J., Prokopczuk, M. and Rudolf, M. (2013) Seasonality and the Valuation of Commodity Options. *Journal of Banking and Finance*, 37 (2), 273 - 290.

Brown, B.D. (1977) The Forward Sterling Market and Its Relation to Arbitrage between the Silver Market in London, Chicago, and New York. *Oxford Economic Papers, New Series*, 29 (2), 292 - 311.

Brown, S.P.A. and Yucel, M.K. (2009) Market Arbitrage: European and North American Natural Gas Prices. *The Energy Journal*, 30, Special Issue: World Natural Gas Markets and Trade: A Multi-Modeling Perspective, 167 - 185.

Duarte, J., Longstaff, F.A. and Yu, F. (2007) Risk and Return in Fixed-Income Arbitrage: Nickels in Front of a Steamroller? *The Review of Financial Studies*, 20 (3), 769 - 811.

Edmans, A., Goldstein, I. and Jiang, W. (2015) Feedback Effects, Asymmetric Trading, and the Limits to Arbitrage. *The American Economic Review*, 105 (12), 3766 - 3797.

Fama, E.F. and French, K.R. (1987) Commodity Futures Prices: Some Evidence on Forecast Power, Premiums, and the Theory of Storage. *The Journal of Business*, 60 (1), 55 - 73.

Harvey, A.C. (1989) *Forecasting, Structural Time Series Model and the Kalman Filter.* Cambridge, England: Cambridge University Press.

Jin, N., Lence, S., Hart, C., and Hayes, D. (2012) The Long-Term Structure of Commodity Futures. *American Journal of Agricultural Economics*, 94 (3), 718 - 735.

Kramer, C. (1994) Macroeconomic Seasonality and the January Effect. *The Journal of Finance*, 49 (5), 1883 - 1891.

Lucia, J.J. and Schwartz, E.S. (2002) Electricity Prices and Power Derivatives: Evidence from the Nordic Power Exchange. *Review of Derivatives Research*, 5 (1), 5 - 50.

Mirantes, A.G., Poblacion, J., and Serna, G. (2012) The Stochastic Seasonal Behaviour of Natural Gas Prices. *European Financial Management*, 18 (3), 410 - 443.

Mirantes, A.G., Poblacion, J., and Serna, G. (2013) The Stochastic Seasonal Behavior of Energy Commodity Convenience Yields. *Energy Economics*, 40, 155-166.

Schwartz, E.S. (1997) The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging. *The Journal of Finance*, 52 (3), 923-973.

Schwartz, E. and Smith, J. (2000) Short-Term Variations and Long-Term Dynamics in Commodity Prices. *Management Science*, 46 (7), 893 - 911.

Shao, C., Bhar, R. and Colwell, D.B. (2015) A Multi-Factor Model with Time-Varying and Seasonality Risk Premium for the Natural Gas Market. Energy Economics, 50 (2015), 207 - 214.

Sorensen, C. (2002) Modeling Seasonality in Agricultural Commodity Futures. *Journal of Futures Markets*, 22 (5), 393 - 426.

Suenaga, H. and Smith, A. (2011) Volatility Dynamics and Seasonality in Energy Prices: Implications for Crack-Spread Price Risk. *The Energy Journal*, 32 (3), 27 - 58.

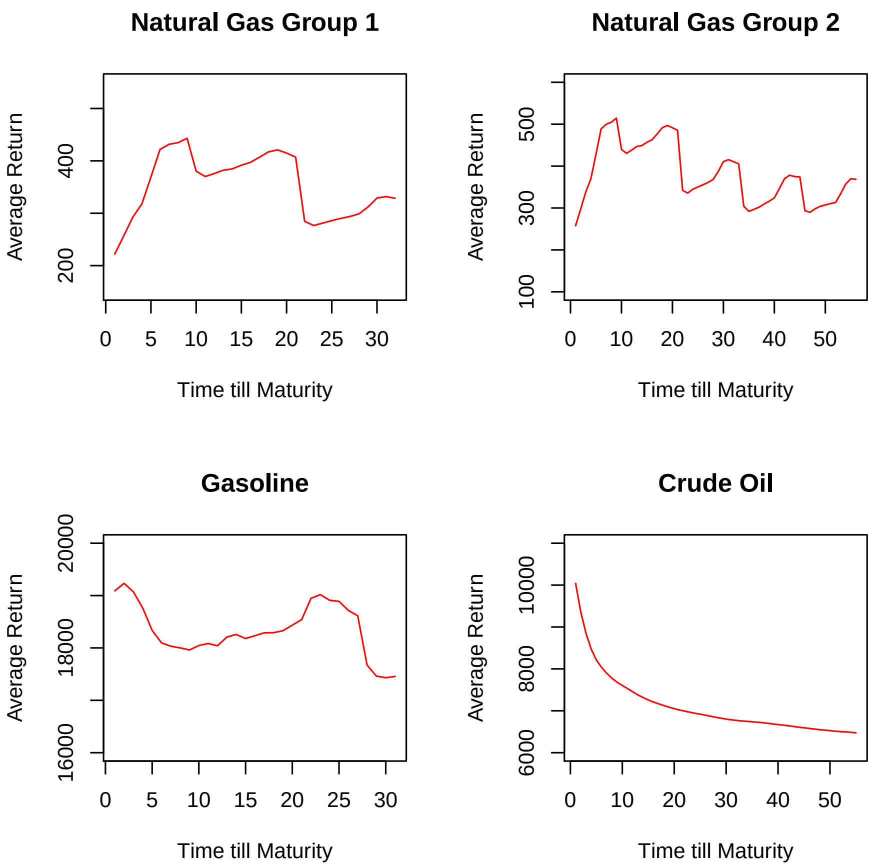

Figure 1: **Results of the Trading Strategy by Contracts**

Table 1

Average Prices and Standard Deviations of Some Selected Contracts

Commodity	F1	F4	F10	F18	F30	F42	F60
Natural Gas, Group 1	3.1393	3.1702	3.1201	3.0475	2.9899		
(4/1997 ∼ 3/2017)	(1.3457)	(1.0512)	(0.9712)	(0.8127)	(0.7316)		
Natural Gas, Group 2	6.7186	7.1273	7.2300	7.0764	6.7427	6.4944	6.1947
(1/2002 ∼ 4/2017)	(2.3399)	(2.4185)	(2.3261)	(2.3360)	(2.1826)	(2.0572)	(1.1770)
Gasoline	221.3155	219.4478	217.5467	215.4578	213.4606		
(3/2007 ∼ 12/2017)	(62.9581)	(58.8160)	(55.2958)	(49.2234)	(45.0019)		
Crude Oil	81.8602	83.1176	83.7683	83.5904	83.0481	82.7449	82.8280
(3/2006 ∼ 12/2017)	(20.4189)	(19.0312)	(17.2767)	(15.6588)	(14.3093)	(13.7424)	(13.5869)

* Numbers in the brackets refer to observation period in the first column, and standard deviation in the rest columns.

Table 2

Details of Trades in Each Period

Commodity	First Trading Month			Second Trading Month		
	Position	Month	Contracts	Position	Month	Contracts
Natural Gas Group 1	Long	Feb.	F5 $\sim$ F36	Short	June	F1 $\sim$ F32
Natural Gas Group 2	Long	Feb.	F5 $\sim$ F60	Short	June	F1 $\sim$ F56
Gasoline	Short	July	F6 $\sim$ F36	Long	Dec.	F1 $\sim$ F31
Crude Oil	Short	July	F6 $\sim$ F60	Long	Dec.	F1 $\sim$ F55

Table 3

Results of the Trading Strategy by Year

	Natural Gas Group 1	Natural Gas Group 2	Gasoline	Crude Oil
1998	-19.72			
1999	726.05			
2000	3,057.05			
2001	-2,407.90			
2002	2,364.81	3,527.96		
2003	2,550.75	4,496.03		
2004	3,028.70	3,722.90		
2005	3,791.72	6,478.78		
2006	-1,265.03	-1,620.85		36,511.50
2007	1,936.83	3,769.18	-99,591.41	-78,992.24
2008	8,987.58	13,583.32	626,767.64	386,418.32
2009	-223.83	220.03	-70,150.09	-45,482.64
2010	-1,939.06	-3,103.6	-77,012.36	-40,981.73
2011	404.86	626.26	105,766.85	47,199.94
2012	-956.23	-1,761.59	-49,281.72	-4,891.46
2013	243.51	175.06	-19,417.11	1,570.51
2014	353.62	1,250.82	248,858.73	130,811.41
2015	-475.55	-1,115.26	90,851.23	63,189.32
2016	1,220.33	1,684.66	-54,871.65	-9,504.89
2017			-76,412.15	-16,328.40
Total	21,378.49	31,933.68	625,507.95	469,519.63

Table 4

Lowest Balance in the Marginal Account

	Natural Gas Group 1	Natural Gas Group 2	Gasoline	Crude Oil
1998	-63.457			
1999	-32.566			
2000	543.784			
2001	-1,053.780			
2002	1,484.738	2,368.544		
2003	349.136	365.844		
2004	905.900	1,237.587		
2005	2,001.440	3,113.090		
2006	-319.788	-761.083		-930.713
2007	272.133	923.362	-89,142.995	-66,425.381
2008	1,904.830	2,992.595	118,242.050	94,399.794
2009	-1,436.768	-2,509.826	-86,482.05	-53,553.409
2010	-2,119.508	-2,825.495	-49,055.143	-30,647.714
2011	196.861	616.119	59,284.620	51,739.657
2012	-1,112.485	-1,660.440	-67,734.983	-18,917.925
2013	427.620	467.649	-12,412.273	-7,647.955
2014	84.486	351.794	15,535.591	6,758.045
2015	-597.069	-1,090.911	48,351.089	28,322.606
2016	119.112	168.178	-34,546.121	-10,356.981
2017			-74,186.155	-11,030.229
Minimum	-2,119.508	-2,825.495	-89,142.995	-66,425.381

Table 5

Mean and Variance of Future Contracts for Calibration

Natural Gas Group 1			Gasoline		
Contracts	Mean	Std.	Contracts	Mean	Std.
F1	4.557	2.311	F1	221.506	63.169
F7	4.912	2.403	F7	218.220	56.243
F13	4.945	2.311	F13	217.081	54.040
F19	4.960	2.316	F19	215.224	49.180
F25	4.905	2.173	F25	214.567	48.463
F31	4.903	2.172	F31	213.307	45.246
Natural Gas Group 2			**Crude Oil**		
Contracts	Mean	Std.	Contracts	Mean	Std.
F1	5.002	2.339	F1	75.461	23.242
F8	5.512	2.398	F8	77.451	21.054
F15	5.574	2.276	F15	77.606	19.748
F22	5.556	2.146	F22	77.428	18.791
F29	5.539	2.082	F29	77.258	18.110
F36	5.485	1.928	F36	77.151	17.628
F43	5.499	1.934	F43	77.144	17.316
F50	5.474	1.793	F50	77.219	17.102
F57	5.496	1.776	F57	77.364	16.968

Table 6

Calibration Results of Natural Gas Group 1

Parameter	Value (Std.)	t statistics
κ	1.04474 (0.02578)***	40.51983
μ^*	0.05280 (0.02713)*	1.94599
λ_y	-0.19710 (0.02523)***	-7.81299
σ_x	0.12401 (0.00557)***	22.24739
σ_y	0.39136 (0.01363)***	28.72413
ρ	-0.32130 (0.04761)***	6.74851
γ_1	-0.00654 (0.00086)***	-7.62309
γ_1^*	-0.06178 (0.00092)***	-67.35968
γ_2	-0.00931 (0.00325)***	-2.86816
γ_2^*	-0.00878 (0.00327)***	-2.68360
$\widehat{\gamma_1}$	-0.00806 (0.00564)	-1.42879
$\widehat{\gamma_1^*}$	0.01434 (0.00603)**	2.37698
$\widehat{\gamma_2}$	-0.01648 (0.00440)***	-3.74341
$\widehat{\gamma_2^*}$	0.02190 (0.00444)***	4.92765

Notes: *a*. Maximum Likelihood: 38,219

b. No. of Observations each contract: 5,010

c. * 10% significance, ** 5% significance, *** 1% significance

Table 7

Calibration Results of Natural Gas Group 2

Parameter	Value (Std.)	t statistics
κ	0.66677 (0.01425)***	46.79867
μ^*	0.11807 (0.03089)***	3.82219
λ_y	-0.17991 (0.01673)***	-10.75110
σ_x	0.11201 (0.00808)***	13.85976
σ_y	0.46863 (0.03224)***	14.53514
ρ	-0.30561 (0.10760)***	-2.84026
γ_1	0.06292 (0.00079)***	79.85361
γ_1^*	-0.00714 (0.00077)***	-9.32955
γ_2	0.02537 (0.00080)***	31.71094
γ_2^*	0.00444 (0.00079)***	5.57751
$\widehat{\gamma_1}$	-0.02206 (0.00618)***	-3.56876
$\widehat{\gamma_1^*}$	-0.00081 (0.00590)	-0.13675
$\widehat{\gamma_2}$	0.01186 (0.00325)***	3.64988
$\widehat{\gamma_2^*}$	-0.01061 (0.00318)***	-3.33955

Notes: *a*. Maximum Likelihood: 43,190

b. No. of Observations each contract: 3,842

c. * 10% significance, ** 5% significance, *** 1% significance

Table 8

Calibration Results of Gasoline

Parameter	Value (Std.)	t statistics
κ	1.35167 (0.143780)***	9.39974
μ^*	-0.19826 (0.05620)***	-3.52758
λ_y	0.02990 (0.03445)	0.86813
σ_x	0.17364 (0.00771)***	22.53572
σ_y	0.18919 (0.01512)***	12.51026
ρ	0.20732 (0.09215)**	2.24988
γ_1	-0.00592 (0.00117)***	-5.07942
γ_1^*	0.05346 (0.00125)***	42.68482
γ_2	-0.01553 (0.00500)***	-3.10483
γ_2^*	0.00997 (0.00452)**	2.20705
$\widehat{\gamma_1}$	-0.04228 (0.01118)***	-3.78296
$\widehat{\gamma_1^*}$	0.05699 (0.01215)***	4.69203
$\widehat{\gamma_2}$	0.00391 (0.00761)	0.51423
$\widehat{\gamma_2^*}$	-0.02549 (0.00772)***	-3.30328

Notes: *a*. Maximum Likelihood: 21,244

b. No. of Observations each contract: 2,732

c. * 10% significance, ** 5% significance, *** 1% significance

Table 9

Calibration Results of Crude Oil

Parameter	Value (Std.)	t statistics
κ	0.49439 (0.02085)***	23.71152
μ^*	0.29481 (0.06264)***	4.70658
λ_y	-0.02216 (0.01094)**	-2.02528
σ_x	0.15008 (0.00818)***	18.33839
σ_y	0.15175 (0.00695)***	21.85048
ρ	0.50442 (0.05535)***	9.11376
$\widehat{\gamma}_1$	-0.02466 (0.00866)***	-2.84681
$\widehat{\gamma}_1^*$	0.03807 (0.01019)***	3.73503
$\widehat{\gamma}_2$	-0.01120 (0.00491)**	-2.28410
$\widehat{\gamma}_2^*$	-0.00697 (0.00499)	-1.39582

Notes: *a*. Maximum Likelihood: 34,963

b. No. of Observations each contract: 2,982

c. * 10% significance, ** 5% significance, *** 1% significance

Thesis Conclusion

This thesis mainly explore the commodity market that constitutes a major part of the financial world nowadays, and contribute to the existing literature in various ways. The first part of the thesis attempts to price Asian options under the Schwartz (1997) model, where there are three stochastic state variables, namely the underlying spot price, the convenience yield, and the interest rate. Since there are no closed-form solutions for the arithmetic average Asian option, we first price the geometric average Asian option with a closed-form solution, and use it as a control variate to price the arithmetic counterpart by Monte Carlo simulation. We are able to significantly improve the accuracy of the simulation without sacrificing computational effort. Then, we show how the option prices respond to different values of the parameters. The next main contribution of the first chapter involves the introduction of a jump diffusion in the spot price. The challenge appears to be that even the price of the geometric average Asian option is no longer log-normal, because the jumping time is governed by Poisson Distribution, and that the Asian option is path-dependent. Hence, no analytical solutions exist for either type of the Asian options. However, we develop a new approach based on the fact that conditional on knowing the jumping time, the geometric average of the underlying price remains log-normal. As a result, there exists analytical solution for the geometric average Asian option, which can be used to price the arithmetic counterpart by Monte Carlo method. Our results show observable improvement in terms of simulation accuracy. However, one of the drawbacks of the first chapter is that the concept of pricing Asian options using Monte Carlo simulation has been arguably outdated in the academia. There have been numerous new and successful methods developed nowadays to accurately price Asian options. We believe that our technique in the first chapter is both innovative and practical among all the different approaches, and a comprehensive comparison between our method and the others could be very interesting.

The second chapter of the thesis focus on the energy market and their seasonality feature. In the past literature, seasonality of a commodity future contract is always related to the maturity date. Nevertheless, we find a new pattern of seasonality that relates to the trading date of their future contracts in three specific energy commodities, namely natural gas, gasoline and crude oil. It is particularly interesting in the case of crude oil, which is believed in the past literature to carry no seasonal feature. We find unique annual seasonal patterns for all three commodities, that the prices of the future contracts are relatively higher on average when traded in some months, and lower in another. We designed the so called backward curves, as opposed to the forward curves that have appeared and been studied in the past literature, to help visualize the new seasonality feature. We further observe an interesting change of behaviour in the natural gas price over the observation period, from normal backwardation to contango in around 2008. Next, we adopt different statistical tests, including the Kruskal-Wallis test, the autocorrelation function, and the power spectrum test, to further consolidate our finding of a new seasonality pattern. Nonetheless, one of the main shortage of the second chapter is the lack of an explanation for such phenomenon. Unlike the conventional maturity-date seasonality, of which the cause is fairly straightforward to understand, we fail to find any solid evidence to explain the trading-date seasonality either in the past literature or in the data that we have, including the daily trading volume data for different commodities. Hence, we believe that it is of significant value for academics to conduct further research on discovering the reasons behind the new seasonality feature in the commodity market.

We further extend our study of the new seasonality finding in the third chapter of the thesis. First, it appears that if there exists a trading date seasonality, it will surely violates the rule of no-arbitrage in the financial market. Therefore, we design a simple "buy low sell high" trading strategy according to our findings in the second chapter when we buy and sell all the available contracts of the three commodities in the corresponding lowest and highest months in a year across our observation period. The results show profit in the long run for all the three commodities, with limited risk exposure during the operation. Then, in light of the positive result from our trading strategy, we decide to build an quantitative pricing model that can capture the arbitrage opportunity to price the future prices of the three commodities. The model is based on Sorensen (2002) paper, with the introduction of a new arbitrage factor that captures the new trading-date seasonality pattern. We calibrate the model with the state space approach, using Kalman filter to estimate the unobservable variables. The results of the calibration show significant value for the vast majority of the parameters, includes the parameters, embedded in the new arbitrage factor, that captures the new seasonal feature related to the trading dates. This indicates that the new trading-date seasonality appear to carry significant statistical value in explaining the futures prices. Nevertheless, we also appeal for further research to extend the third chapter. Firstly, our trading strategy is arguably too rudimentary. Either high transaction cost or high liquidity risk could eliminate the profitability of the operation. However, we believe that this can be resolved if further research could be dedicated to find more sophisticated trading strategy with specific design to improve the return with lower fees and risk exposure.

Acknowledgement

I would like to first thank my main supervisor, Prof Ewald, for his help during the past four years. This paper can never take shape without his wisdom and patience. I was able to learn so much from him, both his knowledge and his dedication to academic research. He has taught me not only how to write a academic paper, but also how to be an excellent academic. I would also like to thank my second supervisor, Dr Minjoo Kim, for his input in this paper. During first three years, I was very lucky to be the Graduate Teaching Assistant for two wonderful teachers, Dr. Lombardi and Dr. Yoshimoto. Their lectures offered me a new world of economics and business, which I will never forget. I would also like to thank my colleagues here in our university for their support. Last but not least, I will dedicate this paper to my family. They are the reason why I have been able to push through so many difficulties during the past 4 years.

I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Printed Name: Yuexiang Wu

Signature: