Improving the Price Forecast of Crude Palm Oil Futures Using Historical Return Variances

Azmi Omar* and Shamsul Majid**

ABSTRACT

This paper attempts to investigate the price relationship between the spot and futures prices of crude palm oil contracts traded in the Malaysian Derivatives Exchange. Using historical variances of spot and futures price returns, we applied a model for approximating the convenience yield developed by Longstaff (1995). This model was then tested on the actual crude palm oil prices for the spot-month and three-month contracts from January 1988 to December 2002, to determine the forecasting accuracy of this model relative to the simple cost-of-carry model. There was some improvement in the predictive ability of this approach which could be attributed to the inclusion of historical return variances. The improvement was also evident when tested against out-of-sample data, suggesting that the inclusion of historical return variances in the form of the convenience yield could provide a better forecast of crude palm oil futures prices.

INTRODUCTION

In a perfect world, the value of futures markets arises from their ability to predict spot prices at a specified future date, thus providing market participants with a means of managing the risks associated with trading a commodity. It is therefore important for players to forecast the *correct* futures price at any given time, as any deviation from this *correct* price will enable the players to either short or long their positions in such a way as to benefit from the mispricing of the commodity.

A model that is frequently used to link futures and spot contracts uses the concept of cost of storage. In this model, the difference between the spot and futures contracts is the cost of carrying the commodity into the future, specified by the maturity date of the futures contract. This difference, or the cost of storage, may either be negative or positive, depending on the inventory level of the underlying asset. However, in the case of crude palm oil (CPO)

* Faculty of Economics and Management Science, International Islamic University, Jalan Gombak, 53100 Kuala Lumpur, Malaysia.

** Faculty of Business and Accounting, University of Malaya, Lembah Pantai, 50603 Kuala Lumpur, Malaysia.

futures, the determination of price of storage using the inventory level is hampered by the time lag in the publication of end-month closing stocks of CPO.

Recognizing the problem associated with end-month closing stocks, this paper seeks to employ historical prices and investigate whether past variances in the returns of both spot and futures contracts can be used to explain the link between them. The basic model of price of storage was used as a basis for this investigation. Using 15-year data of spot month and three-month CPO futures contracts (between January 1988 and December 2002), the study found that the input of historical variances of spot and futures prices can be used to forecast prices of futures contracts. Furthermore, the forecast accuracy was not only limited to in-sample data, but also for out-of-sample data as demonstrated by applying the model to CPO futures contracts at the end of all 12 months in 2003.

LITERATURE REVIEW AND THEORETICAL BACKGROUND

In discussing the relationship between the spot price and futures price, early writers described the latter as a function of the former and price of storage. Brennan (1958) used a model that fixed the supply schedule and argued that the changes in demand determine the relationship between the spot and futures prices. Three components of the storage cost were identified: physical storage costs, risk premium for holding stocks and convenience yield. Physical storage costs are those related to the storage of the commodity, *e.g.* rental of warehouse and insurance. Risk premium is associated with the uncertainty about price changes between the time the commodity is stored and the time when it is delivered to the buyer. The last element, which will be the focus of this paper, is known as the convenience yield, representing the opportunity to gain by holding the physical inventory. This benefit arises from a possible increase in demand (or decrease in supply), providing the owner with the convenience of not having to purchase the commodity at a high price, or the convenience of not having to readjust his production schedule.

In this study, we attempt to test the relationship between the observed spot price and the corresponding futures price as explained by the cost-of-carry model. The cost-of-carry model describes the relationship between the futures price, spot price and the cost of carrying the asset until maturity using the equation:

$$F_{itT} = S_{it} \cdot exp(r_{tT} + w_{iT} - c_{tT}) \quad (1)$$

Here, F_{itT} represents the futures price of asset i observed at time t maturing at time T, S_{it} is the spot price of the same asset i observed at time t, *exp* is the exponential function, r_{tT} represents the risk-free rate applicable between time t to T and w_{iT} is the storage cost of asset i until time T. The next parameter c_{tT} represents the convenience yield for the time period t to T.

Since both w_{iT} and c_{tT} are not easy to estimate, frequently the relationship between spot and futures contracts is described using the simpler cost-of-carry model, *i.e.*, the one without w_{iT} and c_{tT}. In the case for CPO contracts, empirical data seem to show that the link between the spot and futures contracts can be roughly described by the simpler model.

Inclusion of the convenience yield in the cost-of-carry model should theoretically improve the forecast accuracy of futures prices. As mentioned earlier, the difficulty in determining the convenience yield lies in the delay of publication of the end-month closing stock data. It is therefore necessary to find another source of information to estimate the convenience yield. We decided to use a model proposed by Longstaff (1995), who originally developed it for approximating the upper bound benefit that can be enjoyed by owners of non-marketable securities if trading restrictions on those securities are lifted. This benefit can be described as the convenience benefit enjoyed by the holders of marketable securities who are in the position to trade their securities whenever it is profitable to do so. From the perspective of a commodity trader, this can be likened to the convenience yield, except that the asset is a commodity rather than a security. Longstaff provided the solution to approximate the convenience benefit for holding marketable securities, which we now modify to include parameters that are relevant to commodity futures. Essentially, the equation describes the convenience yield as a function of the spot and futures price variances prior to the observation date, represented by:

$$c_{tT} = \ln\left\{\left[2+\frac{\sigma_S^2(T-t)}{2}\right]N\left[\frac{\sqrt{\sigma_S^2(T-t)}}{2}\right]+\sqrt{\frac{\sigma_S^2(T-t)}{2\pi}}\exp\left[-\frac{\sigma_S^2(T-t)}{8}\right]\right\} - \ln\left\{\left[2+\frac{\sigma_F^2(T-t)}{2}\right]N\left[\frac{\sqrt{\sigma_F^2(T-t)}}{2}\right]+\sqrt{\frac{\sigma_F^2(T-t)}{2\pi}}\exp\left[-\frac{\sigma_F^2(T-t)}{8}\right]\right\} \quad (2)$$

where $(T - t)$ is the maturity period, σ_S^2 and σ_F^2 represent the spot and futures return variances for the 15-trading-day period immediately preceding the end of the month when the spot price, futures price and interest rate are observed. *N(.)* and *ln* represent the cumulative normal distribution and the natural log, respectively.

THE DATA

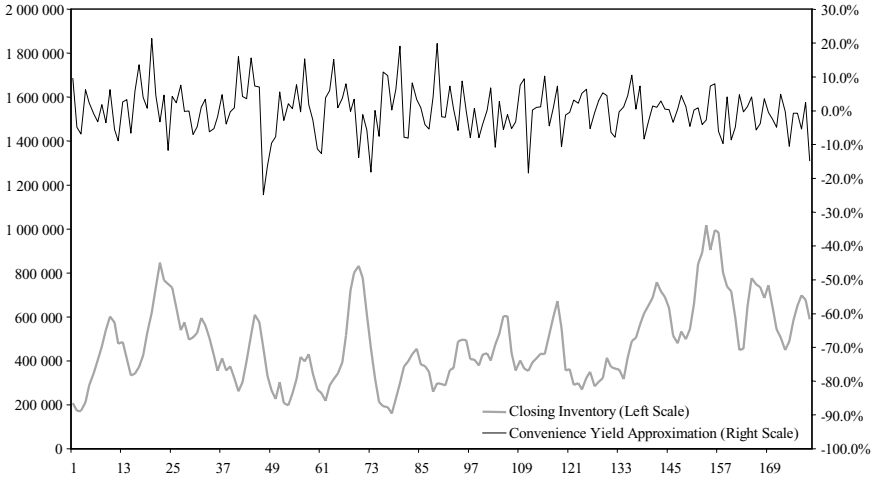
End-of-month observations from the Malaysian Derivatives Exchange on the spot-month and three-month futures prices for the CPO contract were used. These data spanned a 15-year period, from January 1988 to December 2002. The spot-month contract prices were used as a proxy for the current cash price. Since CPO contracts are traded in ringgit, the Malaysian Government's Treasury Bill three-month mid-rate was chosen as a proxy for the risk-free rate, available from Bank Negara's monthly statistical bulletin. The data on monthly closing stocks for CPO were taken from the Malaysian Palm Oil Board (MPOB).

It is worth noting here that there are two markets for palm oil in Malaysia. The first is the physical market whose prices are compiled and released by MPOB, which includes immediate delivery and up to three-month forward prices. This is the market where farmers, refiners and traders trade their need for palm oil. However, the players are strictly those registered with MPOB. The other is the paper market which is managed and run by the Malaysian Derivatives Exchange. This market is open to all players without the need to register with MPOB.

As it is often argued that these two markets serve different needs of the industry, we decided to source our price data from one market for the sake of consistency. Furthermore, the price correlation between the spot-month contract in the Malaysian Derivatives Exchange and the immediate month delivery in the MPOB physical market was found to be close to unity.

This study is limited to the three-month futures contract because of poor liquidity in the longer-dated contracts. Although the study could have benefited from analyses of the six-month, nine-month and 12-month contracts, we were not able to undertake them as these contracts were relatively illiquid.

Our data show that the prices of spot and three-month futures over the 15-year period tracked each other very closely. The inverse relationship between the convenience yield (estimated using historical return variances) and stock level is depicted in *Figure 1*, where the former tends to increase when the latter reduces. Conversely, the convenience yield tends to decrease when the stock level increases. This is in line with the theory of storage which suggests that the convenience yield will lower the cost-of-carry during periods of low inventories, thus making it possible for the spot price to be higher than the futures price.



Notes: The left vertical axis represents the closing inventory in tonnes, and the right vertical axis the convenience yield estimate in %. The horizontal axis represents the month (month 1: January 1988, month 169: January 2002)

Figure 1. Plot of monthly closing inventory and convenience yield estimate (January 1988 to December 2002).

THE MODEL AND RESULTS OF ANALYSIS

The basic model that we used to analyse our data is the transformed cost-of-carry model as described in Equation (1):

$$lgp_T = w_T - c_T \text{ where}$$

$$\text{where} \quad lgp_T = ln\left(\frac{F_T}{S_t}\right) - r_T \qquad (3)$$

In Equation (1), the suffix *tT* (which means from time *t* to *T*) is now changed to *T* to simplify the notations of parameters in the subsequent equations. The equation is then transformed into a linear equation that can be regressed using actual data to test the characteristics of parameters implicit in the cost-of-carry model:

$$lgp_T = \alpha + \beta c_T + \varepsilon_T \qquad (4)$$

Here, *a* represents the storage cost and *b* is the coefficient factor for the convenience yield, which was calculated using Equation (2). Initially we employed the ordinary least squares regression method, but the residuals showed a significant positive serial correlation. We then transformed the above equation into its first-order autoregressive or AR(1) form:

$$lgp_T = \alpha + \rho \cdot lgp_{T-1} + \beta\ (c_T - \rho \cdot c_{T-1}) + u_T \qquad (5)$$

where *p* represents the AR(1)

coefficient and the error $u_T = \varepsilon_T - \rho\varepsilon_{T-1}$. The suitability of AR(1) was tested and confirmed using Wald's test, and this result were further strengthened by the lack of residual serial correlation as indicated by the chi-square test and Durbin-Watson statistic.

The results of the AR(1) regression using the Cochrane-Orcutt iterative method are summarized in *Table 1*. The regression intercept was -0.0137, not significant at $a = 5\%$. However, the slope factor was 0.1052 and statistically significant. No time-dependent variance in the residuals was found using the ARCH(1) test.

The test for structural stability of the model was conducted using Chow's test by dissecting the whole period into two sub-periods of equal lengths. The F-value for this test was 1.798 (p-value 0.15), indicating lack of evidence for significant structural change in the model.

APPLICATION

The importance of the model shown in Equation (5) can only be appreciated by comparing the observed futures price and the theoretical futures price calculated using the two variants of the model - the first by the simple cost-of-carry model as shown in Equation (1), assumed to contain only the risk free-rate (with storage cost w_T and convenience yield $c_T = 0$), and the other by including the convenience yield as shown by the equation below:

$$F_{itT} = S_{it} \cdot exp(r_{tT} + lgp_T) \qquad (6)$$

The means and standard deviations of the absolute and percentage price differences between the observed futures price and the theoretical futures price are reported in *Table 2*, indicating that the model which includes the convenience yield provides a better prediction. The average pricing error fell from 1.86% to 0.03%, with the standard deviation of error dropping from 4.90% to 2.35%.

The model was also tested using out-of-sample data by applying a recursive regression that included all the data until the preceding month for predicting the three-month CPO futures price for the following month. The out-of-sample data tested for this model were the prices of three-month futures contracts for all the months in the year 2003, and the results are tabulated in *Table 3*. Again, there was a significant improvement in prediction by the model including the convenience yield. On average, the price difference decreased from 4.87% to 0.73%, with the standard deviation of error declining from 3.32% to 1.86%. The results indicate that better prediction can be achieved by incorporating the convenience yield in the cost-of-carry model, where the parameter can be approximated using the historical return variance model described in Equation (2).

The accuracy of the forecast was also demonstrated using the prices for the first three months of 2004. The three-month CPO futures contract prices calculated using the model at the end of January, February and March 2004 were RM 1756, RM 1920 and RM 1944 /t, respectively, as compared to actual settlement prices of RM 1760, RM 1935 and RM 1910 for the three end-of-month observations. These represent very small pricing errors of -0.22%, -0.79% and 1.78%, respectively.

However, take note that this model forecasts the prices of three-month futures contracts. In order to apply this model, the spot-month price of CPO contracts must be known as only then can a price relationship be applied to determine the corresponding three-month futures contract. Readers must have noticed that nowhere in this paper have we

TABLE 1. AR(1) REGRESSION OF lgp_T VERSUS c_T USING THE COCHRANE-ORCUTT ITERATIVE PROCEDURE[1] (January 1988 to December 2002)

Parameter	Value
Constant	-0.0137 (-1.0770)
Coefficient C_t	0.1052 (5.2797)*
Coefficient AR(1)	0.8612 (22.9215)*
R-square	0.7558
F-statistic	272.41*
Durbin-Watson statistic	1.984
Residual c^2	11.687
Glejser's test	0.0035 (0.2026)
White's heterocedasticity test	1.9737 (0.3000)**
Park's test	0.0005 (0.4294)
ARCH(1) test	0.0786 (1.048)
Chow's stability F-test value[2]	1.798

Notes: *Significant at 5% level of significance.**p-value corresponding to the c^2 value of the test[1] The AR(1) model was chosen after testing for common factor restriction via Wald's Test. We could not reject the restriction at 5% significance and this provided a valid argument for maintaining the model. [2]Data were split into two sub-periods of 90 months, from January 1988 to June 1996, and the second from July 1996 to December 2002. [3]Analysis carried out using Microfit 4.1

TABLE 2. IMPACT OF CONVENIENCE YIELD ESTIMATES ON FUTURES PRICE (calculated versus actual) FOR IN-SAMPLE DATA (January 1988 to December 2002)

Model	Simple cost-of-carry	Cost-of- carry with convenience yield adjustment
Futures Price Absolute Difference (RM/t)		
Maximum	200.16	149.14
Minimum	-84.12	-106.02
Mean	27.75	1.62
Standard deviation	60.09	29.61
Futures Price Percentage Difference		
Maximum	19.01%	8.97%
Minimum	-10.28%	-8.30%
Mean	1.86%	0.03%
Standard deviation	4.90%	2.35%

proposed a model for predicting the spot-month prices of CPO which, by definition, should be governed by the random-walk model. What it does say is that if the price of spot-month CPO futures is known, we can apply Equations (2) and (5) to calculate the corresponding price for the three-month futures contract. The actual price may vary, but the forecast error should not be more than 2.6% from the observed settlement price.

CONCLUSION

This paper attempts to investigate the price relationship between the spot and futures prices of CPO contracts traded in the Malaysian Derivatives Exchange. In particular, the effort is targeted towards estimating the convenience yield implicit in the prices of futures contracts using historical return variances. This parameter is then included in the cost-of-carry model for determining whether or not any improvement in forecasting ability can be achieved. The results, based on 15-year from 1988 to 2002, support the claim that the model provides a better prediction than the simple cost-of-carry model, both for in-sample and out-of-sample data. This strengthens the view that historical return variances play a role in the pricing of futures contracts for commodities, and can be used to provide a relatively good forecast of futures prices.

ACKNOWLEDGEMENT

TABLE 3. IMPACT OF CONVENIENCE YIELD ESTIMATES ON PREDICTION OF FUTURES PRICE FOR OUT-OF-SAMPLE DATA (calculated versus actual)

Month	Actual three-month futures price (RM/t)	Simple cost of carry estimated		Cost-of-carry with convenience yield estimate	
		Calculated price	Pricing error (%)	Calculated price	Pricing error (%)
Jan-03	1 618	1 617.28	-0.04	1 622.43	0.27
Feb-03	1 592	1 601.15	0.57	1 588.23	-0.24
Mar-03	1 428	1 448.07	1.41	1 439.63	0.81
Apr-03	1 346	1 405.77	4.44	1 379.66	2.50
May-03	1 400	1 476.23	5.45	1 424.69	1.76
Jun-03	1 387	1 474.18	6.29	1 408.77	1.57
Jul-03	1 269	1 394.63	9.90	1 314.45	3.58
Aug-03	1 328	1 409.89	6.17	1 325.05	-0.22
Sep-03	1 497	1 507.59	4.48	1 443.48	0.03
Oct-03	1 869	1 881.98	9.61	1 759.17	2.46
Nov-03	1 828	1 840.73	7.77	1 704.65	-0.20
Dec-03	1 781	1 793.39	2.36	1 689.29	-3.58
Average error		71.91	4.87	9.29	0.73
Std Dev of error		51.02	3.32	28.84	1.86

Note: *All prices are settlement prices of the last trading day of each month.

The authors would like to acknowledge the financial grant received from the University of Malaya under Vot-F Account No. F0359/2003B.

REFERENCES

BRENNAN (1958). The supply of storage. *The American Economic Review*, *48*: 50-72.

LONGSTAFF (1995). How much can marketability affect security values? *Journal of Finance*, *50*: 1767-1774.