Cointegration Between Palm Oil Price and Soybean Oil Price: A Study on Market Integration

Mohammad Haji Alias
Jamal Othman

ABSTRACT

The main objective of the paper is to present a cointegration approach to ascertain whether there exists a long run relationship between palm oil price and soybean oil price. A related objective is to investigate causality patterns between the two price series using the Granger causality test. The study established that the time series on palm oil and soybean oil prices are cointegrated even though separately, each time series is non-stationary. This suggests there exists a long run equilibrium relationship between the two variables. Bidirectional causality is established at the 5 per cent level of significance for the F-test, however, at the 1 per cent level of significance, a unidirectional causality from soybean oil price to palm oil price is established.
INTRODUCTION

The main objective of the paper is to report the application of cointegration techniques to see whether there exists a long run relationship between palm oil price and soybean oil price. Palm oil and soybean oil are two of the main vegetable oils traded in the world market. In 1994 soybean oil accounted for 27.2 per cent and 21.49 per cent of world production and exports of vegetable oils, respectively. The corresponding figures for palm oil were 20.57 per cent for production and 44.5 per cent for exports.

Soybean oil and soybean meal are joint products. Therefore the markets for both products are closely inter-twined, with repercussions from one market being transmitted to the other (see Jamal et al. 1998). Soybean meal accounts for more than 60 per cent of total soybean product value. In the case of palm oil, its joint products are palm kernel oil and palm kernel meal.

Vegetable oil prices tend to move closely with one another. Price differentials play a significant role among vegetable oils in world markets. A slight differential in price among vegetable oils is sufficient to switch manufacturers’ preferences in many international markets. Soybean oil is palm oil’s most important competitor and vice versa. Palm oil often sells at a discount to soybean oil (Figure 1). Price competition especially between palm and soybean oils in some international markets was so stiff to the extent that some producer and allied interest groups launched a widespread negative promotion campaign targeting palm and coconut oils. A related objective of the paper is therefore to investigate causality patterns test between palm oil and soybean oil price using Granger causality (Granger 1996). It is thought that soybean oil is enjoying a “perceived” better health image over palm oil (see Jamal et. al. 1993), therefore, it is hypothesized that there exists a unidirectional causality from soybean oil price to palm oil price.

Earlier specification and estimation of price equations for palm oil proceeded along standard lines. Using annual data for the period 1970-1985, Mohammed Yusof (1988) hypothesised a linear relationship between the price of palm oil and the price of its substitute, soybean world production (proxied by the world production index) and Malaysian production of palm oil. The price equation is embedded in a multiple equation model representing the Malaysian palm oil industry. Non-linear Two Stage Least Squares (2SLS) is used to estimate
FIGURE 1: Palm Oil and Soybean Oil Prices
the model. Mohammad et al. (1997) specified and estimated a palm oil price equation with palm oil price postulated as a function of soybean oil price. The price equation is also part of a multi-equation model, estimated by non-linear 2SLS.

In both cases, tests for the presence of unit roots in variables were not performed. Standard techniques applied to the estimation of the price equations may result in spurious regressions. Granger and Newbold (1974) coined the phrase "spurious regressions" to describe regression results that are good in the sense of having high $R^2$ and significant t-statistics, but in fact do not have real meaning. Such regression results may be obtained in a regression involving economic time series $y_t$ (the dependent variable) and $x_t$ (the explanatory variable) that are trended or non-stationary random processes (see also discussions in Griffiths et al. (1993: 696-702). In this paper we pay special attention to the problem of non-stationarity in variables. Techniques of cointegration and error-correction will be used to help in model specification. Cointegration analysis addresses the issue of spurious regression and identifies the conditions under which relationships are not spurious (Engle and Granger 1987).

In section 2 we briefly outline the concept of cointegration and error-correction modelling strategy. In section 3 we specify the price equation, investigate the integration properties of variables involved and test whether the variables are cointegrated or not. In section 4 we investigate causality patterns between palm oil price and soybean oil price. The final section gives a summary and implications of the study.

COINTEGRATION AND ERROR-CORRECTION MODELLING

Granger (1981) has established a link between non-stationary processes and the concept of long run equilibrium. The link is provided by the concept of cointegration. We will consider the concept of cointegration in its simplest form. If an economic time series $y_t$ follows a random walk, then its first differences would form a stationary process. In this case we say $y_t$ is an integrated process of order one, denoted $I(1)$. If $y_t$ is stationary it is integrated of order zero, denoted as $I(0)$. Granger and Newbold warned against regressing an $I(1)$ variable on another $I(1)$ variable calling the regression as "spurious" as least squares procedure would break down in this case. However Granger has identified the condition under which regressing one $I(1)$ variable on another $I(1)$ variable would not lead to a "spurious regression". The situation is
when the two variables are cointegrated. The application of least squares procedure in this case would work better.

If time-series $y_t$ and $x_t$ are I(1), then in general

$$y_t - \alpha - \beta x_t = \varepsilon_t$$

is also I(1). It is also possible for $\varepsilon_t$ to be stationary or I(0). For this to happen then trends in both $y_t$ and $x_t$ must cancel out when the linear combination $\varepsilon_t = y_t - \alpha - \beta x_t$ is formed. The variable $y_t$ and $x_t$ are then said to be cointegrated with $\beta$ as the cointegrating parameter. Suppose that the long run relationship between $y_t$ and $x_t$ is given by

$$y_t = \alpha + \beta x_t \text{ or } y_t - \alpha - \beta x_t = 0.$$ 

The error term $\varepsilon_t$ then measures how far away $y_t$ and $x_t$ deviates from the equilibrium. If $y_t$ and $x_t$ are both I(1) and $\varepsilon_t$ is stationary with mean zero, then

$$y_t = \alpha + \beta x_t + \varepsilon_t$$

As $\varepsilon_t \sim (0, \sigma^2)$ is stationary, then the variables $y_t$ and $x_t$ obey a stable long-run relationship.

Once the cointegrated variables have been identified, an error-correction model (ECM) can be constructed that describes the dynamic adjustment of variables involved. In an ECM, the first difference of $y_t$ is regressed on the lags of the first differences of $y_t$ and $x_t$ as well as the lagged error-correction term. The error-correction term is the residual $\varepsilon_t$ from the cointegrating regression.

**EMPIRICAL ANALYSIS**

In this paper we specify the following palm oil price equation as estimated in Mohammad et al. (1997).

$$\text{POP}_t = \alpha_0 + \alpha_1 \text{SBOP}_t + u_t$$

where

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>POP</td>
<td>palm oil price</td>
</tr>
<tr>
<td>SBPOP</td>
<td>soybean oil price</td>
</tr>
<tr>
<td>$u_t$</td>
<td>random error</td>
</tr>
</tbody>
</table>
\( \alpha \) is expected to be positive as soybean oil is a substitute for palm oil. The price equation is a statistical relationship with the prior belief that soybean oil price drives palm oil price.

Quarterly data for the period 1980(1) to 1995(4) are used. POP is palm oil price in RM/tonne. SBOP is soybean oil price in RM/tonne. The data used in this study were obtained from the *Oil World*.

TESTING FOR STATIONARITY OF THE SERIES

Figure 1 shows the plot of palm oil and soybean oil prices over time. Both the series seem to be trending somewhat upward. The sense is that both series are not stationary.

Table 1 reports the Dickey-Fuller (1979) (DF) and augmented Dickey-Fuller (ADF) tests for stationarity of both palm oil and soybean oil prices. The Dickey-Fuller test is based on the following regression:

\[
\Delta x_t = \alpha + \beta \Delta x_{t-1} + \varepsilon_t
\]

where \( \Delta x_t = x_t - x_{t-1} \) and \( \varepsilon_t \) is a stationary random error. The null hypothesis that the variable \( x_t \) is non-stationary is rejected when \( \beta \) is significantly negative. The critical value of the Dickey-Fuller (DF) test statistic at the 5 per cent level of significance is about -2.93 for the case where trend is excluded and -3.5 for the case trend is included.

<table>
<thead>
<tr>
<th>Variable</th>
<th>DF test</th>
<th>Augmented DF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t-value</td>
<td>Trend</td>
</tr>
<tr>
<td>POP( _t )</td>
<td>-2.124</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>-2.237</td>
<td>Yes</td>
</tr>
<tr>
<td>SBOP( _t )</td>
<td>-1.847</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>-1.985</td>
<td>Yes</td>
</tr>
</tbody>
</table>

In all cases the null hypothesis of non-stationarity is not rejected. We next proceed to take the first differences of the price series and test for stationarity. The results are given in Table 2.
TABLE 2. Tests for stationarity (first difference form)

<table>
<thead>
<tr>
<th>Variable</th>
<th>F test</th>
<th>Augmented DF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t-value</td>
<td>Trend</td>
</tr>
<tr>
<td>Palm oil price</td>
<td>-6.308</td>
<td>No</td>
</tr>
<tr>
<td>(POP,)</td>
<td>-6.264</td>
<td>Yes</td>
</tr>
<tr>
<td>Soybean oil price</td>
<td>-7.163</td>
<td>No</td>
</tr>
<tr>
<td>(SBOP,)</td>
<td>-7.216</td>
<td>Yes</td>
</tr>
</tbody>
</table>

In all cases the null hypothesis of non-stationarity is rejected. Thus the two time series are stationary after being differenced one. We say that the original (random walk) series are integrated of order one, or I(1).

We next consider the possible cointegratedness of the two price series. As outlined in Section 2, we estimate the cointegrating regression (3) to obtain the residuals

\[
\hat{u}_t = \text{POP}_t - \hat{\alpha}_0 - \hat{\alpha}_1 \text{SBOP}_t
\]  

We then estimate the Dickey-Fuller regression

\[
\Delta \hat{u}_t = (\rho - 1) \hat{u}_{t-1} + \nu_t
\]

In this AR(1) model, \(u_t\) is stationary if \(|\rho| < 1\). If \(\rho = 1\), then the errors are non-stationary.

A simple way to verify whether POP\(_t\) and SBOP\(_t\) series are cointegrated or not is to use the cointegrating regression Durbin-Watson (CRDW) test. The critical values for this test were first provided by Sargan and Bhargava (1983). The CRDW test uses the residuals obtained from the cointegrating regression as a basis to calculate the D.W. statistic. In this case the calculated D.W. statistic is 0.818 (Table 3). The null hypothesis to be tested is D.W. = 0 compared to the
standard D.W. = 2. The critical values to test the null hypothesis that the true D.W. = 0 are 0.511 and 0.386 at the 1 per cent and 5 per cent levels of significance respectively (see also Gujarati 1995: 728). As the computed D.W. is above the critical value, suggests that POP_t and SBOP_t are cointegrated.

The test statistic to test the significance of ρ* of equation (6) is based on the t-statistic for ρ*. The calculated t-ratio is compared to the critical values provided by Engle and Granger (1987). The null hypothesis is Ho: ρ = 1 or ρ* = 0. Ho is rejected based on a one-sided t-test, if t ≤ tc*. The critical value for tc*, α = 0.05 for a sample size equal to 50 is -3.67.

Since t = -3.949 ≤ -3.67 = tc* (α = 0.05) (Table 3), we reject the null hypothesis that the least squares residuals are not stationary; and hence we do not reject the null hypothesis that POP_t and SBOP_t are cointegrated. The use of the augmented Engle-Granger (AEG) test also rejects the null hypothesis that u from the cointegrating equation is non-stationary. The use of the AEG test is recommended for its high power. The estimates of the parameters of the co-integrating equation are also given in Table 3. We accept that 1.976 is a valid estimate of the long run response of POP to a unit change in SBOP.

**TABLE 3. Cointegration tests**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Intercept</th>
<th>SBOP_t</th>
<th>R^2</th>
<th>D.W.</th>
<th>EG</th>
<th>AEG</th>
</tr>
</thead>
<tbody>
<tr>
<td>POP_t</td>
<td>-21.968</td>
<td>1.976</td>
<td>0.629</td>
<td>0.818</td>
<td>-3.949</td>
<td>-4.42</td>
</tr>
<tr>
<td></td>
<td>(-0.223)</td>
<td>(10.26)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: Figures in parentheses are t-ratios.*

CRDW, EG and AEG tests confirm that POP_t and SBOP_t are cointegrated. Even though each variable individually is non-stationary, there appears to be a stable long-run equilibrium relationship between them. In other words, there is no tendency for the two variables to drift wide apart.

**COINTEGRATION AND ERROR-CORRECTION MECHANISM**

We have established earlier that POP_t and SBOP_t are cointegrated i.e. there exists a long-run equilibrium relationship between the two variables.
There may be disequilibrium in the short run. The error term in (4) may be regarded as the "equilibrium error". The error-correction-mechanism (ECM) corrects for the disequilibrium. In this paper we report estimates of the following ECM model:

$$\Delta \text{POP}_t = a_0 + a_1 \Delta \text{SBOP}_t + a_2 u_{t-1} + \varepsilon_t$$  \hspace{1cm} (7)

where $\Delta$ is the first difference operator; $u_{t-1}$ is the one-period lagged value of the residuals from the cointegrating regression; and $\varepsilon_t$ is a random error with classical properties.

The estimates of equation (7) are as follows:

$$\Delta \text{POP}_t = 5.73 + 1.65 \Delta \text{SBOP}_t - 0.38 u_{t-1}$$  \hspace{1cm} (8)

(0.319) \hspace{1cm} (4.79) \hspace{1cm} (-3.424)

$$R^2 = 0.31 \hspace{1cm} \text{D.W.} = 1.66$$

where figures in parentheses are t-ratios.

In equation (7) the change in POP$_t$ is related to the change in SBOP$_t$ and the "equilibrating" error. The variable (SBOP$_t$ captures the short term disturbance in soyabean price, while the lagged u term captures the adjustment of POP to its long run equilibrium value. From equation (8), it can be seen that a short term disturbance in SBOP has a significant positive impact on POP. The coefficient of $u_{t-1}$ is highly significant and negative. About 0.38 of the discrepancy between the actual and long run equilibrium value of POP is eliminated each quarter.

CAUSALITY TESTS

In this section we report tests of causality between palm oil price and soybean oil price using Granger test. We want to test in the Granger sense whether SBOP 'causes' POP (SBOP→POP) or POP 'causes' SBOP (POP→SBOP) or is there feedback between the two (SBOP→POP and POP→SBOP).

To implement the Granger test we estimated the following regressions:
The error terms \( u_{it} \) and \( u_{2t} \) are assumed to be uncorrelated.

The Granger test is based on the sub-set F-test. The application of the Granger test is described clearly for example in Gujarati (1995: 620-622). Table 4 gives a summary of the results of the causality test. The results suggest that the direction of causality is from SBOP → POP and POP → SBOP since the estimated F values are significant at the 5 per cent level. Thus we have feedback or bilateral causality between the two variables. However if we take a smaller level of significance at 1 per cent level, the direction of causality is only from SBOP → POP.

<table>
<thead>
<tr>
<th>Direction of causality</th>
<th>F value</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBOP → POP</td>
<td>3.96</td>
<td>( Fm, n-k, \alpha = 0.01 )</td>
</tr>
<tr>
<td>POP → SBOP</td>
<td>2.83</td>
<td>( Fm, n-k, \alpha = 0.01 )</td>
</tr>
<tr>
<td></td>
<td>2.83</td>
<td>( Fm, n-k, \alpha = 0.05 )</td>
</tr>
</tbody>
</table>

Granger (1988) pointed out that when series are cointegrated, direct Granger test of causality will result in invalid causal inferences. This is because the error correction terms from the cointegrating regression are not included in the tests. As a result the model will be misspecified.

SUMMARY AND CONCLUSION

The study established that the time series on palm oil and soybean oil prices are cointegrated even though separately, each time series is non-stationary. This means that there exists a long run equilibrium relationship between the two variables. There is therefore no tendency for the two variables to drift wide apart in the long run.
In an analysis of causality patterns between the two variables, feedback or bidirectional causality is established at the 5 per cent level of significance. This finding is interesting as it suggests that in the palm oil price determination process, we cannot ignore information on soybean oil price trends, and vice versa. We have some evidence of market integration. However, if we use the 1 per cent level of significance for the F-test, a unidirectional causality from soybean oil price to palm oil price is established. This means that disturbance from the soybean meal subsector would first impact on the soybean oil subsector, because of the joint product relationship, and then impact on the palm oil price. The direction of influence is unidirectional; no feedback impact from palm oil price to the soybean meal subsector.

NOTE

1 The discussion that follows is based heavily on Griffiths et al. (1993: 700-01)

REFERENCES


Faculty of Economics
Universiti Kebangsaan Malaysia
43600 UKM Bangi
Selangor Darul Ehsan