# MARKET TIMING DIEBOLD-MARIANO TEST AND LEADING INDICATORS FOR PREDICTION MODEL

**Suppanunta Romprasert**
Martin de Tours School of Management, Assumption University

## ABSTRACT

*The forecasting accuracy of econometric models in this paper is examined using daily futures price of rubber commodity (RSS3). The criteria investigated include market timing criterion and Diebold-Mariano test. The results suggest that judging by the Confusion Rate (CR) values, it is interesting to note that most of the models are actually quite accurate, correctly predict the direction of price changes of time. In other words, most of the models are useful for predicting the direction of RSS3 price changes. Furthermore, Diebold-Mariano test shows that each pair of model-RW and RWD, RW and RWDT – equally accurate in terms of prediction. It was also found that the quantity of imports natural rubber (China) was able to be the leading indicator of rubber futures price for the next few months.*

## 1. INTRODUCTION

This paper examines the forecast performance of five models using daily futures prices for RSS3 futures market. The econometric models used include a random walk without drift (RW), a random walk with drift (RWD), a random walk with drift and trend (RWDT), a vector autoregressive model with time trend (VAR), and autoregressive (AR). To allow the term structure of prices to evolve over time, the paper estimates all models dynamically, using 310 days, and construct forecasts based on forecast horizons of 1-step-ahead forecasts. The data used to forecast prices is updated daily as new observation becomes available, and ex ante forecasts are constructed. The results are then compared with true values, and out-of-sample forecasting errors are generated. Then, a number of model selection criteria based on these errors are applied and analyzed. Such an approach, often called a model selection approach, has advantages over the more traditional hypothesis testing approach. One reason is that the approach allows focusing on out-of-sample forecasting performance without worrying about the specification of a correct model.

This paper includes market timing criteria and Diebold-Mariano test. The two market timing criteria include confusion matrix (CM) and confusion rate (CR). Furthermore, the other test is the asymptotic loss differential test of Diebold and Mariano (1995), which examines whether two models, are equally accurate based on predictive ability.

By adopting a model selection approach to commodity prices in a real time forecasting scenario, paper attempts to shed light on the usefulness of econometric forecasting and the empirical relevance of modeling theoretical relationships between futures and spot prices when constructing forecasting model. The rest of paper is organized as follows. Section 2 discusses data, while section 3 outlines the forecasting models examined. Section 4 describes estimation strategies, and section 5 introduces the model selection criteria used. Section 6 summarizes the results and conclusion.

## 2. DATA

Daily settlement prices for ribbed smoked rubber sheet No.3 (RSS3) futures markets are employed. All price data are obtained from Agricultural Futures Exchange of Thailand (AFET), Rubber Research Institute of Thailand (RRIT), and The Thai Rubber Association (TTRA) database. The sample period starts on 1 August 2007 - 31 October 2008. The out-of-sample period used is 4 June

2007 to 31 October 2008. Thus, the first forecast for 4 June 2007 is constructed based on in-sample estimation using the period 1 August 2007 to 3 June 2007. However, the researcher does not expect that the results will be affected. In Swanson, Zeng, and Kocagil (1996), they used the data with 3-month, 6-month and one-year in sample sizes and did not find any effects from choosing different in-sample sizes. Futures prices are ordered according to their maturity and are called nearby prices, with nearby one corresponding to the nearest maturity, and nearby two referring to contract with the second nearest maturity, etc. According to Bessembinder, Coughenour, Seguin, and Smoller (1995); Bailey and Chan (1993); and Fama and French (1987), the prices of first nearby futures contracts can be used to proxy for the spot price. Therefore, the analysis based on the futures prices can also be extended to basis movements, or spreads between futures prices and spot prices. The advantage of using futures prices is that it avoids problems which arise when overlapping contracts are used, as well as problems associated with the volatility near the delivery periods. Furthermore, along with demand-supply variables in the market forecasting, because of limitation on available data, the monthly prices are used for examining the leading indicator. The sample period starts on May 2004 to December 2009. The out-of-sample period used is June 2007 and ends on December 2009. Thus, the first forecast for June 2007 is constructed based on in-sample estimation using the period May 2004 to May 2007.

## 3. FORECASTING MODELS

The first model paper considers is a simple random walk without drift. Let $F_t$ be a N*1 price vector at time t for a given RSS3 commodity, and let N be the number of nearby contracts this commodity has. Thus, $F_t = (F_{1,t}, ..., F_{N,t})'$, where $F_{j,t}$ is the $j^{th}$ nearby contract price. Then Random Walk (RW) is $F_t = F_{t-1} + e_t$ where $e_t$ is a N*1 vector of white noise errors.

The second model considered is also a random walk, but with a drift term. Random Walk with Drift (RWD) is $F_t = \alpha + F_{t-1} + e_t$ where $\alpha$ is a N*1 vector of intercepts. Random walk models explicitly impose a unit root on the system, and often perform well relative to a wide class of more complex models, in practice, and are thus useful benchmarks.

The third model examined is a linear vector autoregressive model defined as follows: VAR is $F_t = \alpha_0 + \alpha_1 t + \sum_{i=1}^{p} A_i F_{t-1} + e_t$ where $\alpha_0$, $\alpha_1$ are N*1 vectors, and $A_i (i = 1..., p)$ are N*N coefficient matrices. Notice that the forecasts are based on price levels rather than differences. Level VARs may outperform differenced VAR empirically, even though the variables are non-stationary. One reason why this may be the case is that differencing could result in a loss of information.

The fourth model examined is also a random walk, but with a drift and trend term. Random Walk with Drift and Trend (RWDT) is $F_t = \alpha + F_{t-1} + \beta t + e_t$.

The fifth model paper considered is an autoregressive. The autoregressive (AR) model of an order p can be written as AR(p) and is defined as $X_t = \alpha_1 X_{t-1} + ... + \alpha_{t-p} X_{t-p} + Z_t$ where $Z_t$ is a purely random process and $E(Z_t) = 0$, $Var(Z_t) = \sigma^2$. The parameters $\alpha_1, ..., \alpha_p$ are called the AR coefficients. The name "autoregressive" comes from the fact that $X_t$ is regressed on the past values of itself.

## 4. ESTIMATION STRATEGIES

The estimations of random walk models are implemented directly. VAR parametrs are estimated by using least square. The lag-length is selected by minimizing the Schwarz Information Criteria (SIC) calculated as follows: $SIC = \ln|\Sigma| + n * 1n\, 310/310$ where $|\Sigma|$ is the determinant of the covariance matrix based on the in-sample regression residuals and n is the total number of parameters estimated in all equations. The researcher strategy is to choose an optimal lag length, by beginning with a maximum p = 12 lags, and decreasing p until SIC is minimized. The maximum lag chosen really does not matter in these cases, as the final lag length is often one or two.

Note that seemingly unrelated regression does not improve the efficiency of the least square estimators since all regressions have identical right-hand-side variables and the error terms are assumed to be serially uncorrelated with constant variance. The

excellent discussions on applying VAR models can be found in Sim (1980), Enders (1995), Hoffman and Rasche (1996), and Clements and Hendry (1996). Using levels VARs instead of differenced VARs can be justified by noting that differencing might lead to a loss of information with respect to comovements among variables. The advantage of applying levels VARs is that paper may better mimic the true data generating process.

Finally, it is estimated that the parameters of all regressions at each point of time using a fixed sample size and then forecast prices based on these estimated parameters. At each day, all the estimators and models are updated as the fixed 310 days sample moves forward one period. The forecasting horizon examined is one-step-ahead.

## 5. MODEL SELECTION CRITERIA

Paper employs a number of out-of-sample model selection criteria to evaluate the predictive performance of the five models considered, across RSS3 commodity and one forecast horizon. These criteria can be classified into two categories: criteria for market timing and Diebold-Mariano Test.

All criteria are calculated using forecasting errors based on forecast horizon. The researcher also constructs Diebold and Mariano predictive accuracy test for pair-wise model comparison, as well as market timing test based on confusion matrices, and associated chi-square tests of independence.

### 5.1 Market Timing Criteria

An alternative model selection criterion is the market timing criterion suggested by Henriksson and Merton (1981); Schnader and Stekler (1990: 99-107); Pesaran and Timmermann (1994: 1-7); and Stekler (1994), which can be used to forecast economic turning point. The confusion rate calculated in the paper is retrieved from a 2*2 contingency table, called confusion matrix (CM). The following is the definition of a CM.

| | | Actual Price Movement | |
|---|---|---|---|
| | | up | down |
| Predicted Price Movement | up | $n_{11}$ | $n_{12}$ |
| | down | $n_{21}$ | $n_{22}$ |

where $n_{11}$ = number of cases correctly predicted up; $n_{21}$ = number of cases wrongly predicted down; $n_{12}$ = number of cases wrongly predicted up; $n_{22}$ = number of cases correctly predicted down. The confusion rate is then computed as the frequency of off-diagonal elements, or $CR = (n_{12} + n_{21})/T$

where $T = n_{11} + n_{12} + n_{21} + n_{22}$. The best model according to CR is the least confused one (the one is with the smallest value of CR). The researcher examines the standard chi-squared test of independence. The null hypothesis is independence between the actual and the predicted directions. Thus, rejecting the null hypothesis provides direct evidence that the model is useful as a predictor of the sign of change in the prices. The chi-squared test statistics is calculated as $\chi^2 = \sum_{i=1}^{2}\sum_{j=1}^{2}\frac{(n_{ij} - ne_{ij})^2}{ne_{ij}}$

The value of $ne_{11}$, for example, can be computed by following these four steps. First, compute the probability of actual up, which is $(n_{11} + n_{21})/T$. Second, calculate the probability of predicted up, $(n_{11} + n_{12})/T$. Third, compute the joint probability of actual up and predicted up as the product of the above two, or $[(n_{11} + n_{21})/T][(n_{11} + n_{12})/T]$. Then $ne_{11} = (n_{11} + n_{12})(n_{11} + n_{21})/T$. Similarly, $ne_{22} = (n_{12} + n_{22})(n_{21} + n_{22})/T$, $ne_{12} = (n_{11} + n_{12})(n_{12} + n_{22})/T$ and $ne_{21} = (n_{11} + n_{21})(n_{21} + n_{22})/T$.

### 5.2 Diebold-Mariano Test

Harvey, Leybourne, and Newbold (1998) originally proposed a modification of the Diebold-Mariano test for differences in MSE to account for non-normal distributions of the forecast error series. The researcher also constructs the asymptotic loss differential test proposed in Diebold and Marino (1995). Using only the loss differential series and the assumption that the loss differential series is covariance stationary and short memory, the DM test has a null hypothesis that both forecasting models are equally accurate. Let $\{d_t\}_{t=1}^{T}$ be the loss differential series, then the test statistics is $DM = \bar{d} / \sqrt{2f(0)/T}$, $N(0,1)$ where $\bar{d} = (1/T)\sum_{t=1}^{T} d_t$ is the sample mean loss differential, $f(0) = (1/2\pi)\sum_{\tau=-\infty}^{\infty}\delta(\tau)$ is the spectral density of the loss differential at frequency

zero, $\delta(\tau) = E(d_t d_{t-\tau})$ the autocovariance of the loss differential at displacement $\tau$. f(0) is estimated in the usual way as a two-sided weighted sum of sample autocovariances. The loss differential series used in the analyses are $d_t = (fe_{i,t})^2 - (fe_{i,t})^2$, for the test based on MSE; $d_t = |fe_{i,t}| - |fe_{i,t}|$ for the test based on MAD; and $d_t = \frac{|fe_{i,t} - fe_{j,t}|}{FE_t}$ for the MAPE test, where $fe_{i,t}$ and $fe_{j,t}$ correspond to the forecast error sequences from two forecast models i and j, which are being compared and FEt is $fe_{i,t} + fe_{j,t}$.

### 5.3 Creating Determinants of Prediction Model

Studying the variables affect on rubber price by regression for monthly time-series, the data start from May 2004 through December 2009. Those variables are: RSS3 futures price at time t-1, oil price, exchange rate (Baht per Dollar US.), quantity of consuming natural rubber in the world, quantity of imports natural rubber (Japan), and quantity of imports natural rubber (China).

Method of building the model of RSS3 futures price by multiple regression:

Step 1: Checking the relationship between dependent and independent variables by considering from coefficient in each pair of dependent and independent variable.

Step 2: Finding the best regression by constructing the possible forecasting model between dependent and independent variables, after that bring each method to compare with MSE value and pick the lowest MSE value.

Step 3: Testing hypothesis: Does dependent and independent variable in the model have the relationship?

$H_0$: $\beta_1 = \beta_2 = \beta_3 = \ldots = \beta_j = 0, j = 1, 2, 3, \ldots$ and $H_1$: at least one beta is not equal to zero by using F-test. Also, test that any $\hat{\beta}_j$ is zero by testing on hypothesis $H_0$: $\beta_j = 0$ and $H_1$: $\beta_j \neq 0$. This hypothesis tests the relationship of independent variables that "should $X_i$ be inside the model?" by using t-statistic.

Step 4: Analyze on error term ($\varepsilon_i$): Using histogram tests on the normal distribution, Testing mean and variance of error term on graph between error and time, Testing correlation itself is on ACF and PACF.

Method of constructing model is as: Normalisation: it uses to adjust each variable being in the same measuring unit: Calculating percentage of monthly data of each variable by using $C_{it} = \frac{100(X_{it} - X_{it-1})}{(X_{it} + X_{it-1})/2}$ where $X_{it}$ = data of variable i at time t, $X_{it-1}$ = data of variable i at time t-1, $C_{it}$ = percentage changing of $X_{it}$, Creating the reference graph by using monthly RSS3 is the reference graph on rubber price, Plotting histogram between price and independent variables, considering the leading variable.

The researcher will select the graph that is similar to the reference graph by using visual examination.

## 6. FORECAST PERFORMANCE

**Table 1: Diebold-Mariano Statistics of Predictive Accuracy**

| | Pairwise | Model | Predictive | Accuracy |
|---|---|---|---|---|
| h=1 | 12 | 13 | 14 | 15 |
| RMSE | 0.195 | 0.169 | 0.179 | 0.210 |
| MAE | 0.190 | 0.131 | 0.155 | 0.180 |
| MAPE | 0.001793 | 0.001222 | 0.001448 | 0.001680 |

**Note:** 12 is comparing between RW and RWD. 13 is comparing between RW and VAR. 14 is comparing between RW and AR(1). 15 is comparing between RW and RWDT. Also, h=1 means 1 step ahead.

**Table 2: Diebold-Mariano Test (5% level)**

| Diebold Mariano Test | 5% level | Reject or Unable to reject |
|---|---|---|
| 1 step ahead | Absolute S > 1.96 | Reject the null hypothesis |
| S of 12 | -0.167540167 | Unable to reject |
| S of 13 | 3.153187852 | Reject |
| S of 14 | 2.19018364 | Reject |
| S of 15 | 0.99519309 | Unable to reject |

**Table 3: Model Ranking By Market Timing Criterion**

| | Market | Timing | Criterion | | |
|---|---|---|---|---|---|
| CRITERION | RW | RWD | VAR | AR(1) | RWDT |
| CM | -597 | -510 | -834 | -1,096 | -569 |
| CR | 0.596154 (3) | 0.567308 (1) | 0.644231 (4) | 0.692308 (5) | 0.586538 (2) |
| CHI-SQUARE | 2.918402 | 0.838801 | 0.723896 | 0.621158 | 0.865951 |

The results of the Diebold-Mariano test in Table 1 and 2 show the results that RW & RWD and RW & RWDT are unable to reject the null hypothesis of equal predictive accuracy. Statistically from Table 1, the test also shows that RW & VAR does differ in terms of their squared forecast errors, especially VAR is more accurate than RW model forecasting.

Table 3 reports that judging by the confusion rate (CR) values, it is interesting to note that most of models are actually quite accurate, correctly predict the direction of price changes of time. While RSS3 on RWD model has the lowest CR values. At 5% significance level, all of the chi-square values suggest rejecting the null hypothesis of statistical independence. In other words, most of models are useful for predicting the direction of price changes.

Lastly, finding the factors affects on the monthly rubber futures price movement by multiple regression by collecting all data of all variables from May 2004 through May 2009, totally 61 months. Defining the variables is as following: Dependent Variable: futures = monthly RSS3 futures price at AFET at time t; Independent Variables: futures(-1) = monthly RSS3 futures price at AFET at time t-1, oil = crude oil price, fx1 = exchange rate (Baht per Dollar US.), fx2 = (Yen per Dollar US.), wnc = quantity of consuming natural rubber in the world, wsc = quantity of consuming synthetic rubber in the world, imnj = quantity of imports natural rubber (Japan), imsj = quantity of imports synthetic rubber (Japan), imnc = quantity of imports natural rubber (China), imsc = quantity of imports synthetic rubber (China).

The result from bringing all independent variables analyzed by stepwise shows that exchange rate (Baht/$U.S) and crude oil price affect the monthly rubber futures price. The relationship is shown in the following model:

$$\text{dlog(futures)} = -0.013895 + \underset{(2.341095)^{*}}{3.907749}\,\text{dlog(fx1)} + \underset{(2.967697)^{**}}{0.456859}\,\text{dlog(oil)}$$

$R^2$ (R-Squared) = 0.309811 $R^2$ Adjust = 0.244079

Durbin Watson = 1.663676 F statistic = 4.713232

**Note:** the number in parenthesis is t-Statistic. ** is significant at level 0.05 and * is significant at level 0.25.

Comparing true value with forecasting value from above regression model:

| Month/Year | True Value | Forecasting Value |
|---|---|---|
| January/2009 | 51.42 | 78.80 |
| February/2009 | 51.28 | 56.02 |
| March/2009 | 50.65 | 48.22 |
| April/2009 | 57.20 | 52.91 |
| May/2009 | 59.56 | 56.87 |

The paper finds that exchange rate and crude oil price can be used as leading indicator, especially crude oil price. Also, both of leading indicators show a positive relationship with futures price in AFET. According to forecasting value compares with true value, it expresses that this regression model can be used to show the direction of movement for futures true value.

## SUMMARY

The results suggest that random walk with drift model performs better in shorter forecast horizons, when model is compared based on quadratic loss measures and confusion matrices. Further, by pairwise matching between VAR and randon walk, and between AR(1) and random walk, the results show that both of forecasting models are not equally accurate. We can say that VAR is more accurate than RW model because of least forecast error series based on RMSE, MAD, and MAPE. In addition, the study on the variable that can be the leading indicator for analyzing the trend of future RSS3 futures price found that when both exchange rate (Baht/$U.S.) and crude oil price increase by change in percentage rate, the RSS3 futures price will affect in the same direction.

The positive results of this study could contribute to the skills of agents in the futures market. They would help the traders of the commodities better decisions. Also, this study aims to fill the gap with a guide to understand how the fundamental analysis method works.

Further analysis should combine market integration of all commodities in futures market.

## REFERENCE


Abken, P. A.: *An Analysis of Intra-Market Spreads in Heating Oil Futures*. Journal of Futures Market, Vol. 9, No. 1, pp. 77-86, 1989.

Bailey, W. and Chan, K. C.: *Macroeconomic Influences and the Variability of the Commodity Futures Basis*. Journal of Finance, Vol. 48, pp. 555-573, 1993.

Bessembinder, H., Coughenour, J. F., Seguin, P. J. and Smoller, M. M.: *Mean Reversion in Equilibrium Asset Prices: Evidence from the Futures Term Structure*. Journal of Finance, Vol. 50, pp. 361-375, 1995.

Cecchetti, S., Lam, P. and Mark, N.: *Mean Reversion in Equilibrium Asset Prices*. The American Economic Review, Vol. 80, pp. 399-418, 1990.

Clements, M. P. and Hendry, D. F.: *Forecasting in cointegrated system*. Journal of Applied Econometrics, Vol. 10, pp. 127-146, 1996.

Diebold, F. X. and Mariano, R. S.: *Comparing Predictive Accuracy*. Journal of Business and Economic Statistics, Vol. 13, pp. 253-263, 1995.

Enders, W.: *Applied Econometric Time Series*. John Wiley & Sons, Inc., 1995.

Fama, E. F. and French, K. R.: *Commodity Futures Prices: Some Evidence on Forecast Power*, Premiums, and the Theory of Storage. Journal of Business, Vol. 60, No. 1, pp. 55-73, 1987.

Fama, E. F. and French, K. R.: *Business Cycles and the Behavior of Metals Prices*. Journal of Finance, Vol. 43, pp. 1075-1093, 1988.

Harvey, D. I., Leybourne, S. J. and Newbold, P.: *Tests for forecast Encompassing*. Journal of Business and Economic Statistics, Vol. 16, pp. 254-259, 1998.

Henriksson, R. D. and Merton, R. C.: *On Market Timing and Investment Performance II: Statistical Procedures for Evaluating Forecast Skills*. Journal of Business, Vol. 54, pp. 513-533,1981.

Hoffman, D. and Rasche, R.: *Assessing forecast performance in a cointegration system*. Journal of Applied Econometrics, Vol. 11, pp. 459-517, 1996.

Kim, M. J., Nelson, C. R. and Startz, R.: *Mean Reversion in Stock Prices? A Reappraisal of the Empirical Evidence*. Review of Economic Studies, Vol. 58, pp. 515-528, 1991.

Leitch, G. and Tanner, J. E.: *Economic Forecast Evaluation: Profits Versus the Conventional Error Measures*. American Economic Review, Vol. 81, pp. 580-590, 1991.

Miller, M. H., Muthuswamy, J. and Whaley, R. E.: *Mean Reversion of Standard & Poor's 500 Index Based Changes: Arbitrage-induced or Statistical Illusion*. Journal of Finance, Vol. 49, pp. 479-513, 1994.

Sim, C.: *Macroeconomics and Reality. Econometrica*, Vol. 48, No. 1, pp. 1-48, 1980.

Stekler, H. O.: *Are Economic Forecasts Valuable?*. Journal of Forecasting, Vol. 13, pp. 495-505, 1994.

Swanson, N. R., Zeng, T. and Kocagil, A. E.: *The Probability of Mean Reversion in Equilibrium Asset Prices and Returns*. Working Paper, Penn State University, 1996.

**About the Author:**

Suppanunta Romprasert is a full-time lecturer in the Martin de Tours School of Management. She can be reached at thailandsuppanunta@yahoo.com