Intra-Market Price Discovery in an Emerging Stock Market: Vector Fractionally-Integrated Error Correction Model and Toda-Yamamoto Level VAR Approaches

A. Mansur M. Masih* and Rumi Masih

Abstract

The study investigates the price discovery process by which markets attempt to find equilibrium prices among a system of disaggregate daily spot share price indices relating to the Malaysian Stock Exchange using a VAR. Specifically: (i) a vector fractionally-cointegrated error-correction model is proposed and estimated to investigate the short-run dynamics accounting for the long-run information via a fractionally-integrated error-correction term; and (ii) the Toda-Yamamoto (1995) [k + d(max)]th-order VAR procedure to specify a 'level' VAR containing integrated and cointegrated processes of arbitrary orders is adopted to uncover the long-run driving forces behind stock market linkages. The results are interpreted in the context of the price discovery process among spot stock prices. Our findings indicate consistently that the price discovery process was focused on the palm oil market in Malaysia in the sense that this market played, relatively, the leading role (both in the short- and long-term) being the most exogenous of all. In particular, we demonstrate that previous research, by using ordinary difference VARs either by intent or lack of support of cointegration in the standard sense, ignored an important component of linkages displayed over the long run.

Keywords: Intra-Market Price Discovery; Granger Causality; Mean-Reversion; Fractional Cointegration.
JEL classification: G15, C52

---

* Corresponding author. A. Mansur M. Masih is the Saudi Aramco Chair Professor of Energy Economics at the Department of Finance and Economics, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia, 31261, e-mail: masih@kfupm.edu.sa. Rumi Masih is a Managing Director at the JP Morgan Asset Management, New York, NY 10167. The authors would like to express their gratitude to the Chief Editor of the Journal and the two anonymous referees for their helpful comments. We thank Jim Stock and Fallaw Sowell for their contributions. We also thank The King Fahd University of Petroleum and Minerals for their generous support. The views expressed in this paper are not necessarily attributable to JP Morgan or any of its affiliated agencies.

1. Introduction

Most of the studies on price discoveries, particularly in emerging economies, test the relationship between financial markets rather than within a particular financial market such as stocks, interest rates, bonds and foreign exchange. There is obviously a need for more studies on the process of intra-market price discovery in a financial market.Moreover, a major deficiency in present studies, particularly those relating to the methodological aspect, is that there is a relative dearth of studies dealing with the appropriate modelling of the low frequency dynamics inherent in the temporalcharacteristics of a number of financial variables.[1] This issue is of pertinent importance, particularly in the empirical analysis of high frequency financial time-series.

In many cases while the standard cointegration tests may fail to give any robust evidence of a cointegrating relationship, the fractional analysis may show up cointegration. Intuitively, it implies that although deviation from equilibrium may exhibit substantial persistence in the short run, it will still be mean-reverting in the long run. This is because the standard cointegration tests capture only those cases where the deviation from equilibrium is *short* and the restoration of equilibrium is *quick*. But the fractional cointegration tests can capture cases where the deviation from equilibrium is *prolonged* and the restoration of equilibrium is *slow*. Unlike the standard cointegration analysis, the fractional cointegration captures very slow mean-reverting processes. Intuitively, if two markets are fully cointegrated, any shock to one of these markets (resulting in deviation from equilibrium) is quickly dissipated by arbitrage activities and equilibrium is restored quickly, but if the two markets are fractionally cointegrated, any shock to one of these markets is not quickly dissipated by arbitrage activities and the deviation from equilibrium persists for a long time before the equilibrium is finally restored (see, Masih and Masih, 2004 for further details).

With the above motivation in mind, our purpose in this paper is firstly, to investigate the process of intra-market price discovery in an emerging economy. In other words, we want to examine whether the different stocks in a stock exchange are in equilibrium or not in the long run, and, if they are found to be in equilibrium, then since there is a deviation from equilibrium in the short-run, we want to examine which stocks lead and which follow

[1] See, by example, studies using fractional differencing: Diebold and Rudebusch (1989) in a study with aggregate output, Lo (1991) using financial time series, Cheung (1993) in exchange rate data, Masih and Masih (1995, 1998, 2004) in studies of PPP and spot and forward exchange rates, Booth et al (1982) in analysis of foreign exchange rates in two separate monetary regimes, Shea (1991) in an investigation of the term structure of interest rates, Sowell (1992a, 1992b), Haubrich (1993) who calculate the stochastic properties of consumption and income as a fractionally-differenced process.

in order to bring about the equilibrium in the long-run. In other words, we want to investigate the process as to which stock is the initial receptor of an exogenous shock and as a result which stocks bear the brunt of short-run adjustment to bring about the long-run equilibrium. We use the Malaysian stock exchange as a case study.

Secondly, to demonstrate that the standard tests of cointegration may pose too strict a test. In order to do so, the paper applies a relatively new but generalised concept of fractional cointegration and applies it to shed light on the relationship between a system of six major disaggregate Malaysian daily asset prices. Essentially, this approach relaxes the strict assumption in standard tests of cointegration that the residual from the cointegration equation must be an integer [more specifically $I(0)$] to allow for rejection of the hypothesis of non-cointegration. In other words, the standard cointegration tests assume a discrete hypothesis, i.e., the deviation from equilibrium or the residual in the cointegrating equation is $I(1)$ against an alternative $I(0)$ but the fractional cointegration tests assume that the residual in the cointegrating equation is $I(1)$ against an alternative $I(\mathrm{d})$ where 'd' is less than unity. That means in fractional cointegration the alternative is broader than the ADF test. If 'd' is less than unity, then the deviation may be large and may persist over a long period but eventually the process is mean-reverting and the process will get back to equilibrium. In other words, in a fractional cointegration the deviation from equilibrium or the residual is a long-memory process (unlike the residual in a standard cointegration which is a short-memory process). Given the work on fractionally integrated processes as cited above, a number of researchers have questioned the strict nature of the integer choice of differencing, in that it appears too restrictive leading to invalid conclusions. An illustration of one such problem, is in reference to the equilibrium error since, in terms of a mean-reverting process, according to Granger and Joyeux (1980) and Hosking (1981), the process does not need to be strictly $I(0)$. Moreover, we also illustrate the potential of this approach to provide a wide range of cases of mean-reversion with processes that exhibit fractionally cointegrated behaviour.

Thirdly, in order to investigate the propagation mechanism of share price fluctuations across the market we then specify a vector fractionally-integrated error-correction model (VFIECM) where the error-correction term (ECT) in this special case is no longer $I(0)$ but $I(d)$ where $0<d<1$. This allows us to conduct short- and long-run causal inference (in the Granger sense) through lagged difference terms and the fractionally integrated ECT which incorporates the highly persistent, low frequency dynamics that manifests over time to reach a long-run equilibrium. The methodology employed also holds theoretical and empirical favour over much of the previous literature which has attempted to investigate stock market linkages using ordinary bivariate or multivariate first-difference VARs [see Mathur and Subrahmanyam (1990) on linkages between Nordic and US stock markets, Chowdhury (1994) in analysing linkages of NIC markets, and Brocato (1994) in an investigation of global market adjustments].

Fourthly, in a statistically acceptable and compatible manner, we also apply a convenient method prescribed in the econometric literature by Toda and Yamamoto (1995) that permits statistical (causal) inference to be made in VARs which may be stationary around a deterministic trend, integrated or cointegrated of an arbitrary order. This novel approach specifies the VAR in which variables appear in *levels* but it incorporates the asymptotic theory that validates statistical (causal) inference as in normal causality testing in VARs. In this respect, purely the long-run linkages may be deciphered which is of critical concern to financial analysts.

Finally, given the generality of the techniques introduced and utilised in the analysis, we illustrate how such methodological rigour may be packaged in a compatible and complementary fashion to offer highly intuitive information that would not have been gauged from the methodology that has been used in this area. Such an investigation fills a void in the literature, to date, on studies involving the fast growing emerging market. To this extent, the paper may also be seen as a primer for the wealth of applied financial econometric literature relying upon tests of long-run equilibrium relationships and dynamic causal linkages.

This paper is organised in the following manner. The methodology is presented and discussed briefly in Section 2. We then turn to the presentation and discussion of results in Section 3. The paper concludes with implications for policy, and future research in financial econometrics in Section 4.

2. Data and Methodology

The variables used are the daily stock indices of the Malaysian industrial (*IND*), financial (*FIN*), property (*PRO*), tin (*TIN*), palm oil (*POL*) and rubber (*RUB*). The study is based on unpublished data set of 1727 daily observations for each series covering the period from 2000 to 2005 collected through personal correspondence.

2.1. Long-Run Causal Inference in Level VARs Including Integrated and Cointegrated Processes of Arbitrary Orders

The VFIECM discussed in the previous section permits inference on both short-run and, to a certain extent, long-run linkages. While this has been proposed in the recent literature [see King, Plosser, Stock and Watson (1991), Mosconi and Giannini (1992), Toda and Phillips (1993)] and exhibits highly desirable properties both from a theoretical and empirical point of view, this formulation does have its drawbacks in that it is implicitly dependent upon pre-tests of integration and cointegration. Furthermore, there is also a covert, yet heavy reliance on such tests providing an accurate report of the cointegration ranks and, as we have already noted, even in the case where the error-correction term is appropriately derived, there is an additional step to augment this term in the original VAR formulation. In addition, the VECM formulation involves implicit non-linear restrictions on parameter vectors, which may be problematical as tests for Granger non-causality may involve size distortions due to rank deficiency. The basic necessity of this additional step of embedding an error-correction term arises from the need to re-capture the long-run information lost through differencing the variables entering the VAR. One way to circumvent this problem is to posit a VAR in which variables appear purely in their level form. However, while Sims, Stock and Watson (1990) do propose the asymptotic theory to validate appropriate Granger causality tests in level VARs, here there is also a pre-requisite that the system be cointegrated. Such tests, in particular the most popular being due to Johansen (1991), do tend to be sensitive to nuisance parameters [see Toda (1995)] and suffer from finite-sample biases.

Toda and Yamamoto (1995) proposed a complementary procedure that allows causal inference to be conducted in level VARs that may contain integrated processes but does not involve rigorous attention and strict reliance upon integration and cointegration properties of any or all variables in the system. In essence, this procedure circumvents some of the pre-test biases that practitioners may be confronted with in VECM and other modelling formulations involving unit root and cointegration pre-testing. Furthermore, the Toda-Yamamoto procedure is simple and convenient to apply and permits linear as well as non-linear tests of restrictions. These restrictions themselves would then imply long-run causal inference since, unlike ordinary difference VARs, this formulation involves only variables in levels. Unlike the Johansen procedure, the Toda-Yamamoto procedure is applicable regardless of whether the series is $I(0)$, $I(1)$, $I(2)$, non-cointegrated or cointegrated of any arbitrary order. Its applications can be seen, among others, by Wolde-Rufael (2005), and Rousseau and Vuthipadadorn (2005).

Toda and Yamamoto (1995) propose estimation of a levels' VAR of the form:

$$y_t = \gamma_0 + \gamma_1 t + \ldots + \gamma_q t^q + \theta_1 y_{t-1} + \ldots + \theta_k y_{t-k} + \ldots + \theta_p y_{t-p} + \zeta_t \qquad (1)$$

by OLS, where $t = 1,...,T$, and p ³ $(k+d)$ consisting of y_t's that are $I(d)$ which may be $CI(d,b)$. The q's are coefficient matrices but hypothesis testing of restrictions will preclude the terms $\theta_{K+1},...,\theta_p$ which are assumed to be zero. In matrix notation, this may be written as:

$$Y' = \Gamma\Lambda + \Phi X' + \Psi Z' + E' \qquad (2)$$

where $\mathbf{\Gamma} = [\gamma_0 ,... ,\gamma_q]$, $\mathbf{\Lambda} = [\tau_1 ..., \tau_T]$ with $_t = (1, t ,..., t^q)'$, $\Phi = [\theta_1 ..., \theta_k]$ $\mathbf{X}_t = [x_1 ,..., x_T]$ with $x_t = (y'_{t-1} ,..., y'_{t-k})'$, $\mathbf{\Psi} = [\theta_{k+1} ,..., \theta_p]$ and $\mathbf{Z} = [z_1 ..., z_T]$ with $z_t = (y'_{t-k-1} ,..., y'_{t-p})'$. Toda and Yamamoto (1995) then show that the test of hypothesis H_0: $f(\phi)$ where f = vec(Φ) is a parameter vector, may be tested by a Wald statistic which is asymptotically chi-square with *m* degrees of freedom, subject to p^3k+d. The statistic is given by:

$$\omega = f(\phi)'[F(\phi)\{T^{-1}E'E \otimes (X'QX)^{-1}\}F(\phi)']^{-1} f(\phi) \tag{3}$$

where , and $Q_\tau = I_T - \Lambda(\Lambda'\Lambda)^{-1}\Lambda'$,

where I_t is a $T'T$ identity matrix. Effectively, this implies that all one needs to do is to determine the maximal order of integration [d(max)] that we may believe the model to incorporate and ascertain the lag structure, and then to construct a VAR with variables appearing in their levels with a total of $p=k+d$(max) lags. However, at the inference stage, linear or non-linear restrictions should only be tested on the first k lags since the p-k lags are assumed zero and ignored. The over-parametrization of the model is done intentionally, and the procedure will be valid as long as the order of integration does not exceed the true lag-length (k) of the model.[2] According to Toda and Yamamoto (1995), the Wald statistic is shown to be valid in a wide variety of cases, regardless of any component in Y being stationary (around a linear trend), $I(1)$, $I(2)$, non-cointegrated or cointegrated of an arbitrary order.

While this method does circumvent some of the problems associated with pre-test bias from tests required for the VECM or Sims, Stock and Watson (1990) procedure, the issue of over-fitting the model does itself entail a loss of power. Since the procedure is valid asymptotically, efficiency will also be affected in particular cases where the true lag-length may be as small as one, and augmenting additional lags in a small-sample VAR may prove to be costly in terms of parsimony. This, though, will not occur frequently in applied work using data with a frequency greater than annual and orders of integration not exceeding two.

It is with these caveats in mind that in this study the procedure outlined will be applied as a complement to the VECM already proposed. In adopting the methodology as a complementary tool, the model employing a VAR in *levels* will add an extra dimension to the analysis in addition to providing another facet to assess the general robustness of the results generated from the VECM. Moreover, as Toda and Yamamoto (1995, p. 246) remark "...we are *not* suggesting that our method should totally replace the conventional hypothesis testing that is conditional on the estimation of unit roots and cointegrating ranks. It should rather be regarded as complementing the pretesting method that may suffer serious biases in some cases." In this analysis, we are employing the Toda-Yamamoto methodology sharing a similar motivation.

[2] Note, however, that this is most unlikely to be the case in most empirical works. If $d=1$, then the procedure will always be valid since $k^3 1$. The one exception is in the very special case where $d=2$ and the true lag-length $k=1$.

2.2. Variance Decompositions (VDCs) and Causal Relativities

The VFIECM, *F*- and *t*-tests may be interpreted as within-sample causality tests. They can only indicate the Granger-exogeneity or endogeneity of the dependent variable within the sample period. They do not provide an indication of the dynamic properties of the system, nor do they allow us to gauge the relative strength of the Granger-causal chain or degree of exogeneity amongst the variables beyond the sample period. VDCs, which may be termed as out-of-sample causality tests, by partitioning the variance of the forecast error of a certain variable (say, the industrial share prices) into proportions attributable to innovations (or shocks) in each variable in the system including its own, can provide an indication of these relativities. Placed under an alternative context, VDCs provide a literal breakdown of the change in the value of the variable in a given period arising from changes in the same variable in addition to other variables in previous periods. A variable that is optimally forecast from its own lagged values will have all its forecast error variance accounted for by its own disturbances [Sims (1982)][3].

3. Application, Estimation and Discussion of Results

3.1. Univariate Integration and Multivariate Cointegration

3.1.1. Tests for Univariate Integration

In order to verify to what degree the time series variables share univariate integration properties, we perform both unit root tests and mean stationary tests.

The DF type tests and the non-parametric Phillips-Perron (PP) type tests developed by Dickey and Fuller (1979, 1981), Phillips (1987), Phillips and Perron (1988), and Perron (1988) are convenient testing procedures, both based on the null hypothesis that a unit root exists in the autoregressive representation of the time series. DF tests attempt to account for temporally dependent and heterogeneously distributed errors by including lagged

[3] The results based on VARs & VDCs are generally found to be sensitive to the lag length used and the ordering of the variables. A considerable time has been spent in selecting the lag structure through FPE criterion. FPE method is based on an explicit optimality criterion of minimizing the mean squared prediction error. The criterion tries to balance the risk due to bias when a low order is selected, and the risk due to increase in the variance when a higher order is selected. By construction, the errors in any equation in a VAR are usually serially uncorrelated. However, there could be contemporaneous correlations across errors of different equations. These errors were orthogonalised through Choleski decomposition. In order to orthogonalise the innovations, a pre-determined triangular ordering of the six variables had to be made. The innovations were orthogonalised in the following order: [*IND, FIN, PRO, TIN, POL, RUB*]. Alternative orderings did not alter the results to any substantial degree. This is possibly due to the variance-covariance matrix of residuals being near diagonal.

sequences of first differences of the variable in its set of regressors. The PP tests try to account for dependent and IID processes through adopting a non-parametric adjustment thereby eliminating any nuisance parameters. Recently these tests have been shown [see Schwert (1987) and DeJong et al (1992)] to suffer from lack of power as they often tend to accept the null of a unit root too frequently against a stationary alternative. Moreover, the Phillips-Perron statistics have been shown to perform poorly over small samples.

These studies have also implied that it would be worthwhile to conduct tests of the null hypothesis of mean stationarity in order to determine whether variables are stationary or integrated. Mean stationarity tests are performed with a test proposed by Kwiatkowski, Phillips, Schmidt and Shinn (1992). This test [abbreviated as KPSS] is based on the statistic:

$$\eta(u) = (1/T^2)\sum_{t=1}^{T} S_t^2 / \sigma_k^2 \text{ where } S_t = \sum_{i=1}^{t} v_i, t = 1, \ldots T \tag{4}$$

with v_t being the residual term from a regression of y_t on a intercept, and σ_k^2 is a consistent long-run variance estimate of y_t, and T represents the sample size. Kwiatkowski et al (1992) show that the statistic $h(u)$ has a non-standard distribution and critical values have been provided therein. If the calculated value of $h(u)$ is large then the null of stationarity for the KPSS test is rejected. Since we entertain both the ADF tests and the KPSS test in this exercise, we consider a variable to contain a unit root or be unit-root non-stationary if the null hypothesis of non-stationarity is not rejected by the ADF tests but the null hypothesis that the variable is mean stationary is rejected by the KPSS test.[4]

For presentation purposes we only report results of the ADF test for a unit root and KPSS tests for mean stationarity. Both sets of results for each spot price indexes are presented in Table 1. ADF tests indicate the presence of a unit root in each series since for no series can the null of non-stationarity be rejected. To allow us to measure how persistent the unit root in the process is, we also calculate a confidence interval due to Stock (1991) who proposes a method to derive 90% confidence intervals for the largest autoregressive root. The confidence interval estimates, not surprisingly, tend to suggest that the unit root is highly persistent with all lower bounds quite clearly above 90%. Considering evidence of mean stationarity, the KPSS tests reveal that the null of mean-stationarity is rejected quite convincingly for all series.

---

[4] This guidline in considering the stochastic properties of univariate time series is also used in an empirical analysis containing error-correction modelling by Mehra (1994).

Table 1. KPSS Tests for Mean-Stationarity, Augmented Dickey-Fuller Tests and Stock's (1991) Confidence Intervals for the Largest Autoregressive Root

	KPSS (1992) Mean-Stationarity [$\eta(u)$]	ADF	Stock's (1991) 90% Confidence Interval
IND	-0.795**	-2.690	(0.923,1.013)
FIN	-0.894**	-2.537	(0.927,1.043)
PRO	-1.003**	-2.192	(0.938,1.027)
TIN	-0.954**	-2.220	(0.927,1.057)
POL	-0.921**	-1.978	(0.910,1.066)
RUB	-1.147**	-2.752	(0.924,1.011)

Notes: Variable definitions: Malaysian industrial (*IND*), financial (*FIN*), property (*PRO*), tin (*TIN*), palm oil (*POL*) and rubber (*RUB*). The full sample of Malaysian composite and disaggregate share indices consisted of 1727 daily observations. The optimal lag used for conducting the Augmented Dickey-Fuller test statistic was selected based on an optimal criteria (Akaike's Final Prediction Error), using a range of lags. ** represents significance (rejection of null) at 5 per cent level.

Given this unambiguous set of results from both tests, we conclude that these daily asset prices are integrated at most of order one. This provides a requisite for the forthcoming multiple cointegration analysis.

3.1.2. Tests for Multivariate Cointegration

In this analysis we employ the Johansen and Juselius (JJ) procedure of testing for the presence of multiple cointegrating vectors. It is demonstrated in Johansen (1991) that the procedure involves the identification of rank of the *m* by *m* matrix P in the specification given by:

$$\Delta X_t = \delta + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Pi X_{t-k} + \varepsilon_t \qquad (5)$$

where, $\mathbf{X}_t$ is a column vector of the *m* variables, G and P represent coefficient matrices, □ is a difference operator, *k* denotes the lag length, and □ is a constant. If P has zero rank, no stationary linear combination can be identified. In other words, the variables in $\mathbf{X}_t$ are non-cointegrated. If the rank *r* of P is greater than zero, however, there will exist *r* possible stationary linear combinations and P may be decomposed into two matrices □ and □, (each *m* x *r*) such that P =□□′. In this representation b contains the coefficients of the r distinct cointegrating vectors that render □ $'X_t$ stationary, even though $\mathbf{X}_t$ is itself non-stationary, and □contains the speed-of-adjustment coefficients for the equation.

The JJ test results for both the maximum eigenvalue (L) and trace tests are presented in Table 2. By using an optimally determined lag length for the six-dimensional VAR, evidence of cointegration amongst this set of asset prices is, according to standard critical values, non-existent, with r=0 not being able to be rejected by either the L or trace tests. We also tried alternative numbers of lags, but results seemed insensitive, at least across lag-length. This would seem to imply that these prices, as a system, do not share any long-term cointegrating relationships. Intuitively, the finding of non-cointegration would seem to suggest that, at least in the long-run, fluctuations from one asset price may not be used optimally to forecast the future movements of another price.

However, an examination of the stochastic properties of these prices reveals that, while the system is not cointegrated in their logs, they appear to be fractionally cointegrated if we allow for mean-reverting processes that are $CI(1,d)$ with $0<d<1$. The paper demonstrates that relaxing the condition that the residual from the cointegration equation must be a $I(0)$ process, captures a much wider class of mean-reversion and low-frequency dynamic behaviour even in high frequency data.

Table 2. Johansen-Juselius MLE Tests for Multiple Cointegrating Vectors

	Hypotheses		Test Statistics	
Vector	H_0:	H_1:	Max Eigenvalue	Trace
[*IND, FIN, PRO, TIN,*				
POL, RUB]	$r = 0$	$r > 0$	23.742	82.524
	r £ 1	$r > 1$	22.641	58.782
	r £ 2	$r > 2$	18.620	36.141
	r £ 3	$r > 3$	12.520	17.52
	r £ 4	$r > 4$	4.661	5.002
	r £ 5	$r = 6$	0.341	0.341

Notes: The optimal lag structure for each of the VAR models was selected by minimising the Akaike's FPE criteria. Critical values are sourced from Johansen and Juselius (1990). ** indicates rejection at the 95% critical values.

The evidence of cointegration rules out the possibility of spurious relationship among the variables but cointegration cannot tell us which variable is the leader and which variable is the follower. For that we have to go to the error-correction model. Hence, we engage in an analysis of these low-frequency dynamics by employing recently proposed econometric procedures to investigate the propagation mechanism driving these short-run dynamics and propelling the persistent but equilibrium-tending long-run relationship.

Specifically a vector fractionally-cointegrated error-correction model is estimated to investigate the short-run dynamics accounting for the long-run information via a fractionally-integrated error-correction term (Table 3). The results are interpreted in the context of the price discovery process among the spot stock prices. Our findings indicate consistently that the price discovery process was focused in the palm oil market in Malaysia in the sense that this market played, relatively, the leading role (both in the short- and long-term) being the most exogenous of all. In other words, it was the initial receptor of any exogenous shocks to the long run equilibrium resulting in disequilibrium in the short run and all other markets (which are bound together through cointegration) had to bear the burden of short-run adjustment endogenously in different proportions to bring about long-run equilibrium. That implies that the palm oil market picks up the new information first and all other markets follow suit.

Table 3. Summary of Temporal Causality Results Based on Vector Fractionally-Integrated Error-Correction Model (VFIECM)

	Short-Run Lagged Differences						*Error-Correction Term*
	ΔIND	ΔFIN	ΔPRO	ΔTIN	ΔPOL	ΔRUB	**ECT**[$\xi_{i,t-1}$]
Dep Variable			*F*-statistics				*t*-statistic
ΔIND	–	1.45	0.46	2.18**	0.78	1.71	-2.700**
ΔFIN	1.65	–	0.70	1.87*	0.74	1.18	3.328**
ΔPRO	2.10*	0.95	–	1.46	0.86	1.33	-2.242*
ΔTIN	1.64	1.30	0.41	–	0.59	1.83*	-2.619**
ΔPOL	1.46	1.30	1.03	1.50	–	1.15	-1.697
ΔRUB	2.31**	1.59	0.63	1.18	2.63**	–	-1.734

Notes: The ECTs [x_{it-1}] were derived by normalising the cointegrating vector on *IND*, with the residual fractionally-integrated. While other normalisations were tried, they did not significantly alter the findings summarised above. Figures presented in the final column are estimated *t*-statistics testing the null that the lagged fractionally-integrated ECT is statistically insignificant. All other estimates are asymptotic Granger *F*-statistics. The VFIECM was estimated including an optimally determined [Akaike's FPE] lag structure of 12 for all lagged-difference terms and a constant. ** and * indicate significance at the 1%, 5% levels.

3.2. Toda-Yamamoto [k+d(max)]th-Order Levels VAR: Long-Run Statistical Causal Inference

Given the non-standard nature of the above cointegration results, we now proceed to apply the Toda-Yamamoto (1995) [k + d(max)]th-order VAR procedure to specify a 'level' VAR containing integrated and cointegrated processes of arbitrary orders in order to uncover the long-run driving forces behind stock market linkages in Malaysia. Although common univariate integrational properties were verified by standard unit root tests, this procedure provides a method by which to test for statistical causal inference and circumvents, to a certain degree, the problems associated with variables integrated and cointegrated of alternative orders, including those which are stationary around a linear trend, and the pre-test biases involved with reaching the condition required for a standard VECM.

Governed by the Wald statistic in determining the order of the lag-length [Toda and Yamamoto(1995), p. 243-245)], the summary results of restrictions upon k=9 lags from fitting and estimating a (k+1=10)th order VAR in levels, is presented in Table 4 with these lead-lag linkages being summarised in a flow diagram appearing in Panel B of Table 5. These results also suggest, quite clearly, the palm-oil market as econometrically exogenous in the long-run as no other prices appear to be causally active in the POL equation. In other words, consistent with our earlier VFIECM results, the palm oil market is the driver in the long-run. There also appears to be quite a significant degree of influence from this market to the rubber market (RUB).

The utility of this method as a complementary augmenting tool to conduct causal inference is illustrated in the results presented in Table 5, where the two sets of flow diagrams describing short- and long-run causal flows, derived from the VFIECM and the Toda-Yamamoto level VAR procedure, show up to be broadly similar.

3.3. Variance Decompositions: Causal Relativities

Although the VFIECM (as noted earlier), provides us with a rich, dynamic framework for which temporal causality may be tested, they are strictly within-sample tests. The error-correction model can tell us which variable is exogenous (i.e., leader) and which variable is endogenous (i.e., follower), but the error-correction model cannot tell us the relative exogeneity or endogeneity of the variables. In order to gauge the relative exogeneity / endogeneity of the variables or to 'quantify our temporal causality results', we now shock the system of stock markets and partition the forecast error variance of each of the markets. The decomposition results are presented in Table 6 up to sixty days following the shock. Looking through the main diagonal, we may ascertain the extent to which a variable is exogenous since this represents how much of a market's own variance is being explained by movements in its own shock over the forecast horizon. In a statistical

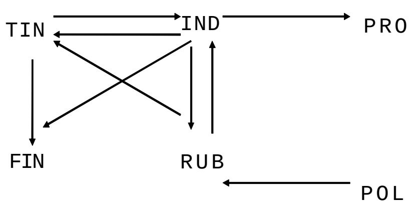
Table 4. Summary of Long-Run Causality Results Based on Toda-Yamamoto [k+d(max)]th-Order 'Level' VAR Procedure

	IND	*FIN*	*PRO*	*TIN*	*POL*	*RUB*
Dep Variable	*Sig Levels Associated with* ω					
IND	–	0.480	0.948	0.133	0.823	0.074*
FIN	0.027**	–	0.741	0.062*	0.833	0.327
PRO	0.003***	0.556	–	0.301	0.669	0.107
TIN	0.026**	0.276	0.973	–	0.980	0.013**
POL	0.180	0.182	0.740	0.206	–	0.114
RUB	0.007***	0.245	0.810	0.768	0.006***	–

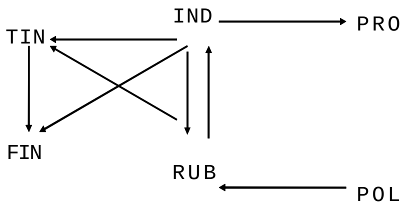
Notes: The VAR was estimated in levels with d(max) = 1 as evidence indicated the maximum order of integration was unambiguously equivalent to one. The selection of the lag-length (k = 9) was determined via a Wald statistic as prescribed by Toda and Yamamoto [(1995), p. 243-245]. Reported above are significance levels associated with asymptotic Wald statistic w [see text] for testing exclusion restrictions. *** , **, and * indicates significance at the 1%, 5% and 10% levels.

Table 5. Lead-Lag Linkages Summarised From VFIECM and Toda Yamamoto Level VAR Procedure

Panel A: Vector Fractionally-Integrated Error-Correction Model (VFIECM)



Panel B: Toda-Yamamoto Level VAR Procedure



Notes: X ® Y indicates that changes in X contains leading information for changes in Y (ie. changes in X Granger causes changes in Y, or changes in Y lags or is influenced by changes in X). X ¬ Y implies the reverse.

sense, if a variable explains most of its own shock, it then does not allow variances of other variables to contribute to it being explained and is therefore said to be relatively exogenous. The important issue here is that the decomposition analysis is a relative exercise and results from it should be interpreted with this in mind.

A novelty of variance decompositions and its usefulness in quantifying causal linkages is brought to light in examining the industrial share price decomposition. The decomposition analysis quite clearly brings out the substantial contribution of the industrial market to all other markets, and consistent with the results derived from the VFIECM analysis, the industrial prices make the least explanatory contribution on the shock to the palm-oil prices. From the main diagonal, the two prices which are by far the most exogenous appear to be those representing palm-oil and industrial shares with about 78 and 75 per cent of their variances being explained by their own innovations. By far the most explained markets in terms of its relative variance being explained by other prices, are those of the financial and property markets where approximately 70 per cent of their variances are explained by innovations in other prices. In this respect, this analysis also sheds some additional insight into which prices are more endogenous/ exogenous relative to each other, out of sample.

From the point of view of price discovery, the out-of-sample variance decompositions results are consistent with those of the within-sample VFIECM in that the palm oil market tends to drive all other markets in the context of the Malaysian stock market.

4. Conclusions and Policy Implications

In this paper, we have suggested and applied recently developed methods that allow for such analysis employing a multivariate, dynamic framework allowing for both short and long run relationships to manifest over time. In this respect some of the methodological deficiencies existing in previous studies in the literature, have to some extent been addressed.

Acknowledging the warnings of several empiricists in the literature, this paper has attempted to illuminate a relatively new but generalised notion of fractional cointegration, which effectively relaxes the knife-edged, stringent distinction in much of the previous works involving cointegration, where the equilibrium error must be either $I(0)$ or $I(1)$. Although previously suggested by Granger (1986), there has been relatively little application of this concept to any issues of interest in the literature. In illustrating the technique's potential, we apply it to shed some light on price discovery using disaggregate daily stock prices of Malaysia as a case study. The results indicate the attractiveness of fractional cointegration to detect mean reversion behaviour for this system of prices.

Table 6. Decomposition of Variance

Percentage of Forecast Variance Explained by Innovations in:

		D*IND*	D*FIN*	D*PRO*	D*TIN*	D*POL*	D*RUB*
Mths	Relative Variance in:						
1	D*IND*	100.00	0.00	0.00	0.00	0.00	0.00
3		98.54	0.10	0.05	0.10	0.19	0.47
6		76.28	1.10	0.36	21.02	0.43	0.81
12		75.22	1.33	0.40	21.46	0.65	0.95
24		75.22	1.33	0.40	21.46	0.65	0.95
Mths	Relative Variance in:						
1	D*FIN*	68.86	31.13	0.03	0.10	0.00	0.00
3		68.16	30.85	0.19	0.34	0.11	0.18
6		67.03	30.59	0.41	1.06	0.35	0.56
12		66.49	30.20	0.54	1.33	0.67	0.75
24		66.49	30.20	0.55	1.33	0.67	0.76
Mths	Relative Variance in:						
1	D*PRO*	66.33	1.77	31.90	0.00	0.00	0.00
3		65.90	2.09	31.38	0.11	0.10	0.42
6		64.87	2.34	30.97	0.79	0.28	0.75
12		64.55	2.37	30.58	1.02	0.55	0.93
24		64.55	2.37	30.57	1.02	0.55	0.93
Mths	Relative Variance in:						
1	D*TIN*	42.75	1.24	1.58	54.43	0.00	0.00
3		43.38	1.32	1.58	52.90	0.16	0.66
6		43.56	1.84	1.67	51.68	0.30	0.94
12		43.18	2.02	1.75	50.98	0.82	1.25
24		43.18	2.02	1.75	50.98	0.82	1.25
Mths	Relative Variance in:						
1	D*POL*	17.72	0.69	0.36	1.15	80.08	0.00
3		17.64	0.85	0.45	1.60	79.09	0.36
6		16.78	1.44	0.74	2.03	78.18	0.82
12		16.48	1.46	1.05	2.33	77.83	0.85
24		16.48	1.46	1.05	2.33	77.83	0.85
Mths	Relative Variance in:						
1	D*RUB*	49.72	0.96	0.22	1.15	7.65	40.30
3		50.46	1.08	0.26	1.21	8.19	38.80
6		49.45	1.62	0.62	1.54	8.67	38.10
12		49.32	1.83	0.77	1.83	8.81	37.43
24		49.31	1.83	0.77	1.85	8.81	37.43

Notes: Figures in the first column refer to horizons (ie. number of days). All other figures are estimates rounded to two decimal places - rounding errors may prevent a perfect percentage decomposition in some cases. Several alternative orderings of these variables were also tried such alterations, however, did not alter the results to any substantial degree. This is possibly due to the variance-covariance matrix of residuals being near diagonal, arrived at through Choleski decomposition in order to orthogonalise the innovations across equations.

In order to analyse these low-frequency dynamics, a vector error-correction model incorporating an error-correction term, which is fractionally integrated, was formulated and estimated. As this formulation is a subset of the larger class of VECMs, these models mechanically specify a formulation of dynamic linkages from a broad set of possibilities [see Tegene and Kuchler (1994)], they permit the data to express its driving forces without imposing any constraint by way of a specific structure. In this regard, VECMs provide a rich variety of procedures to be applied in testing temporal causal hypotheses [see King, Plosser, Stock and Watson (1991), Toda and Phillips (1993)]. In this analysis an extension of the VECM formulation was made by embedding an error-correction term which was suggestive of being fractionally integrated. Extensions to this model were also made in reference to assessing our within-sample (VECM) tests by decomposing the forecast error variance from a once -off shock to the system, into contributions of a particular price's own and other prices' variance, in explaining the shock over a time path. Finally, a recent method provided by Toda and Yamamoto (1995) to conduct statistical causal inference in systems integrated and cointegrated of arbitrary orders was employed in a complementary manner to gain some understanding of the linkages of the system in the long-run.

The policy implications of our dynamic analysis are quite clear. This study is an attempt at placing the analysis of stock market linkages in a temporal Granger causal framework, by binding the relationship among six disaggregate prices of the Malaysian stock market, using daily data. Put briefly, our results hold import for policy / investment considerations in the following ways:

(i) The evidence of fractional cointegration rules out the possibility of the estimated relationship being 'spurious' and this intuitively implies that, although deviations from an equilibrium condition may exhibit substantial persistence in the short-run, it will still be mean-reverting over the long-run. This, in itself, is a very important and valuable finding for assisting financial analysts who may use information optimally in acknowledging that evidence does exist of mean-reversion in systems of stock prices.

(ii) Due to the presence of fractional cointegration among the six prices, while standard cointegration tests fail to reveal any cointegrating relationship, the fractional cointegration analysis applied in this study yields far more favourable results.

(iii) The Granger-causal chain implied by our dynamic analysis (based on both VFIECM and VDCs) tends to, more or less, broadly suggest that, of all the prices included in the system, palm-oil shares appear to be relatively more exogenous over both the short and the long-run. Our findings indicate consistently that the price discovery process was focused in the palm oil market in Malaysia in the sense that this market played, relatively, the leading role (both in the short- and long-term) being the most exogenous of all. In other words, it was the initial receptor of any exogenous shocks to the long run equilibrium relationship, resulting in disequilibrium in the short run and then all other markets (which are bound together through fractional cointegration) had to bear the burden of short-run adjustment endogenously in different proportions to bring about the long-run equilibrium. That implies that the palm oil market picks up the new information first and all other markets follow suit. The price discovery role played by the palm-oil prices in the long-run is further confirmed by the Toda-Yamamoto long-run causal analysis. Here, also, palm-oil prices are clearly the most exogenous (or least endogenous) in comparison, in the Granger-causal sense. Furthermore, significant short-run linkages appear to run from the industrial share prices to all other prices, apart from those of palm-oil shares, as well as bivariate causality between the industrial-tin and industrial-rubber prices. These are major empirical findings, which contain both intuitive appeal and strong policy implications, especially in the Malaysian context.

(iv) Another finding (based mostly on the variance decomposition analysis) that has policy implications is to bring to light the substantive contributions of the industrial prices in explaining the shocks to most other prices, except that of palm-oil which appeared to be most exogenous of all.

(v) Apart from these conclusions, this analysis has contributed in illustrating how alternative methodologies may be used in a complementary fashion to unearth previously unfounded relationships, both as short- and long-run phenomenon. While most studies to date use integration and cointegration as stringently strict pre-conditions to fulfil prior to analysis, we have suggested relevant and rigorous methodology to deal with situations with a much wider class of cases that a practitioner, particularly in the field of finance, may be confronted with. Moreover, these techniques could hold enormous potential in investigating possibilities of mean-reversion and multiple equilibria in multivariate systems where equilibrium is theoretically postulated or empirically elusive.

References

Booth, G.G., F.R. Kaen and P.E. Koveos (1982), R/S Analysis of Foreign Exchange Rates Under Two International Monetary Regimes, *Journal of Monetary Economics*, 10, 407-415.

Brocato, J. (1994), Evidence on Adjustments in Major National Stock Market Linkages Over the 1980s, *Journal of Business Finance and Accounting*, 643-667.

Cheung, Y-W (1993), Long Memory in Foreign-Exchange Rates, *Journal of Business and Economic Statistics*, 11, 93-101.

Chowdhury, A.R (1994), Stock Market Interdependencies: Evidence from the Asian NIEs, *Journal of Macroeconomics*, 16, 629-651.

Dickey, D. and W. Fuller (1979), Distribution of the Estimators for Autoregressive Time-Series with a Unit Root, *Journal of the American Statistical Association*, 74, 427-431.

Dickey, D. and W. Fuller (1981), Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root, *Econometrica*, 49, 1057-1072.

DeJong, D.N., J.C. Nankervis, N.E. Savin and C.H. Whiteman (1992), Integration Versus Trend Stationarity in Time Series, *Econometrica*, 60(2), 423-433.

Diebold, F.X. and G.D. Rudebusch (1989), Long Memory and Persistence in Aggregate Output, *Journal of Monetary Economics*, 24, 189-209.

Granger, C.W.J. (1986), Developments in the Study of Cointegrated Economic Variables, *Oxford Bulletin of Economics and Statistics*, 48, 213-228.

Granger, C.W.J. and R. Joyeux (1980), An Introduction to Long Memory Time Series Models and Fractional Differencing, *Journal of Time Series Analysis*, 1, 15-39.

Haubrich, G. (1993), Fractional Integration and the Permanent Income Hypothesis, *Review of Economics and Statistics*, 76, 255-260.

Hosking, J.R.M. (1981), Fractional Differencing, *Biometrika*, 68, 165-176.

Johansen, S. (1991), Estimation and Hypothesis Testing of Cointegrating Vectors in Gaussian Autoregression Models, *Econometrica*, 59, 1551-1580.

Johansen, S. and K. Juselius (1990), Maximum Likelihood Estimation and Inference on Cointegration with Applications to Money Demand, *Oxford Bulletin of Economics and Statistics*, 52, 169-210.

King, R.G., C.I. Plosser, J.H. Stock and M.W. Watson (1991), Stochastic Trends and Economic Fluctuations, *American Economic Review*, 81(Sept.) 819-840.

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt and Y. Shin (1992), Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?, *Journal of Econometrics*, 54(Oct / Dec), 159-178.

Lo, A.W. (1991), Long-Term Memory in Stock Market Prices, *Econometrica,* 59, 1279-1314.

Mathur, I. and V. Subrahmanyam (1990), Interdependencies Among the Nordic and U.S. Stock Markets, *Scandinavian Journal of Economics*, 92(4), 587-597.

Masih, R. and A.M.M. Masih (1995), A Fractional Cointegration Approach to Empirical Tests of PPP: New Evidence and Methodological Implications From an Application to the Taiwan / US Dollar Relationship, *Weltwirtschaftliches Archiv*, 131(4), 673-694.

Masih A.M.M. and R. Masih (1998), A Fractional Cointegration Approach to Testing Mean Reversion between Spot and Forward Exchange Rates: A Case of High Frequency Data With Low Frequency Dynamics, *Journal of Business Finance and Accounting (JBFA)*. 25(7)&(8), Sept / Oct, 987-1003.

Masih A.M.M. and R. Masih (2004), Fractional Cointegration, Low Frequency Dynamics and Long-Run Purchasing Power Parity : An Analysis of the Australian Dollar, *Applied Economics,* 36 (6), 593 -605.

Mehra Y.P. (1994), An Error-Correction Model of the Long-Term Bond Rate, *Federal Reserve Bank of Richmond Economic Quarterly*, 80(4), 49-68.

Mosconi, R. and C. Giannini (1992), Non-Causality in Cointegrated Systems: Representation, Estimation and Testing, *Oxford Bulletin of Economics and Statistics,* 54, 399-417.

Phillips, P.C.B. (1987), Time Series Regressions with a Unit Root, *Econometrica*, 55, 277-301.

Phillips, P.C.B. and P. Perron (1988), Testing for a Unit Root in Time Series Regression, *Biometrika*, 75, 335-346.

Rousseau, P.L. and D. Vuthipadadorn (2005), Finance, Investment, and Growth: Time Series Evidence from 10 Asian Economies, *Journal of Macroeconomics*, 27, 87 - 106

Schwert, G.W. (1987), Effects of Model Specification on Tests for Unit Roots in Macroeconomic Data, *Journal of Monetary Economics*, 20, 73-103.

Shea, G.S. (1991), Uncertainty and Implied Variance Bounds in Long-Memory Models of the Interest rate Term Structure, *Empirical Economics*, 16, 287-312.

Sims, C.A. (1982), Policy Analysis with Econometric Models, *Brookings Papers on Economic Activity,* 1, 107-152.

Sims, C.A., J.H. Stock and M.W. Watson (1990), Inference in Linear Time Series Models with a Unit Root, *Econometrica*, 58, 113-144.

Sowell, F.B. (1992a), Modeling Long Memory Behaviour with the Fractional ARMA Model, *Journal of Monetary Economics,* 29, 277-302.

Sowell, F.B., (1992b), Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models, *Journal of Econometrics*, 53, 165-188.

Stock, J.H. (1991), Confidence Intervals for the Largest Autoregressive Root, *Journal of Monetary Economics*, 28, 435-459.

Tegene, A. and F. Kuchler (1994), Evaluating Forecasting Models of Farmland Prices, *International Journal of Forecasting*, 10, 65-80.

Toda, H.Y. (1995), Finite Sample Performance of Likelihood Ratio Tests for Cointegrating Ranks in Vector Autoregressions, *Econometric Theory*, 11.

Toda, H.Y. and P.C.B. Phillips (1993), Vector Autoregressions and Causality, *Econometrica*, 61(6), 1367-93.

Toda, H.Y. and T. Yamamoto (1995), Statistical Inference in Vector Autoregressions with Possibly Integrated Processes, *Journal of Econometrics*, 66, 225-250.

Wolde-Rufael, Y. (2005), Energy Demand and Economic Growth, *Journal of Policy Modelling*, 27, 891 – 903.