1000

BEBR

FACULTY WORKING PAPER NO. 1000

# Tests For The Existence of a Dynamic Daily Market Model: Methods and Implications

*C. F. Lee*
*P. A. Cartwright*
*P. Newbold*

College of Commerce and Business Administration
Bureau of Economic and Business Research
University of Illinois, Urbana-Champaign

# BEBR

FACULTY WORKING PAPER NO. 1000

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

December 1983

Tests For The Existence of a Dynamic Daily Market Model: Methods and Implications

C. F. Lee, Professor
Department of Finance

P. A. Cartwright, Professor
University of Georgia

P. Newbold, Professor
Department of Economics

## Abstract

In this note we consider the extension of the simple market model to a form where leads and lags are allowed. Granger's test of causality is applied to daily data on both heavily and lightly traded stocks to check for possible leads and lags in the system. Also we attempt to assess the strength of these relationships.

## 1. Introduction

The market risk, or beta, of a security, or portfolio of securities, is generally estimated by fitting through least squares the market model

$$R_{jt} = \alpha_j + \beta_j R_{mt} + \varepsilon_{jt} \tag{1}$$

where $R_{jt}$ is the rate of return on the jth security and $R_{mt}$ the market rate of return. In equation (1) $\varepsilon_{jt}$ is a random error term. However, in relationships involving time series data, an essentially static formulation such as (1) will be inadequate if there are leads or lags in the system. Recently, following Dimson (1979), Reinganum (1982) has allowed for such a possibility by fitting to portfolios of stocks an elaboration of the model (1), including on the right-hand side terms in $R_{m,t-k}$, for both positive and negative k. This formulation, which resembles one employed by Sims (1972), suggests the possibility of testing for leads and lags between individual rates of return and market rates, using the methodology of the econometric causality literature. As Reinganum shows, the existence of strong leads or lags in the market model would lead to alternative definitions of beta. Moreover, if the rate of return on an individual security depends to some extent on previous market rates, this implies some predictability of future individual returns, and hence market inefficiency

In this paper, we apply to this problem the causality tests proposed by Granger (1969), using daily rate of return series on samples of both heavily traded and lightly traded common stocks. We have two objectives in implementing this test. First, we want to test for the existence of a dynamic daily market model, and second we wish to measure the strength, in terms of gains in forecasting rates of return, of any relationship. The Granger formulation is particularly well suited to this latter purpose.

Section 2 of the paper discusses the methodology used, while our empirical results are presented in the third section. In section 4 we explore the implications of these results, summarizing our findings in section 5.

## 2. Methodology

Let $R_{j,t}$ and $R_{m,t}$ denote two time series of returns, and consider the problem of forecasting $R_{j,n+1}$, based on the two sets of information

$$I_0 = [R_{j,n-k};\ k \geq 0]$$

and

$$I_1 = [R_{j,n-k},\ R_{m,n-k};\ k \geq 0]$$

In practical implementation, attention is restricted to predictors that are linear functions of members of the information sets. Then, if $R_{j,n+1}$ is better predicted, in the expected squared error sense, using information set $I_1$ than using $I_0$, $R_m$ is said to "cause" $R_j$, in the sense of Granger (1969). In an analogous fashion we can define the event "$R_j$ causes $R_m$." If, as can be the case, causality runs in both directions, the pair of time series is said to exhibit "feedback."

Granger proposes a test of causal direction based on the fitting of vector autoregressive models to a pair of time series. Assume that the vector process $\underset{\sim}{R}'_t = (R_{j,t}, R_{m,t})$ admits a stationary infinite order autoregressive representation

$$\underset{\sim}{R}_t = \underset{\sim}{\alpha} + \sum_{k=1}^{\infty} \Phi_k \underset{\sim}{R}_{t-k} + \underset{\sim}{\varepsilon}_t$$

where $\underset{\sim}{\alpha}$ is a vector, the $\Phi_k$ are 2x2 matrices of parameters, and $\underset{\sim}{\varepsilon}_t$ is zero-mean vector white noise[1], so that

$$E\,(\underset{\sim}{\varepsilon}_t\ \underset{\sim}{\varepsilon}'_t) = \Sigma$$

and

$$E(\underset{\sim}{\varepsilon}_t\ \underset{\sim}{\varepsilon}'_{t-k}) = \underline{0} \quad (k \neq 0)$$

Then, $R_m$ causes $R_j$ if and only if the (1,2) elements of the $\Phi_k$ are not all zero. Similarly, $R_j$ does not cause $R_m$ if and only if the (2,1) elements of these matrices are all zero. In practice, of course, it is not possible to estimate an infinite number of free parameters, so that the autoregression is truncated at some maximum lag K, chosen (often arbitrarily) to be sufficiently high to allow adequate description of any dynamic relationship.

To implement the test, then, consider the model

$$R_{j,t} = \alpha_1 + \sum_{k=1}^{K} \phi_{11,k} R_{j,t-k} + \sum_{k=1}^{K} \phi_{12,k} R_{m,t-k} + \varepsilon_{1,t} \qquad (2)$$

where $\varepsilon_{1,t}$ is a random error term. The null hypothesis that $R_m$ does not cause $R_j$ is then checked by testing $\phi_{12,k} = 0$ (k=1, 2, ..., K), using the usual F-test based on the ordinary least squares fitting of

(2). Similarly, to test the null hypothesis that $R_j$ does not cause $R_m$, we can fit

$$R_{m,t} = \alpha_2 + \sum_{k=1}^{K} \phi_{21,k} R_{j,t-k} + \sum_{k=1}^{K} \phi_{22,k} R_{m,t-k} + \varepsilon_{2,t} \tag{3}$$

where $\varepsilon_{2,t}$ is a random error term, and test $\phi_{21,k} = 0$ (k=1, 2, ..., K).

Now, the use of the term "causality" in connection with these tests is certainly controversial. However, for our present purposes, we do not need to get involved in this philosophical debate. Rather, we simply require tests of predictability in our search for possible leads and lags. For example, if the rate of return for an individual stock is better predicted when information on past market rates is added to the information on past individual rates, we will say that the market rate leads the individual rate.

The empirical tests we report in the next section are based on relatively long series of daily data on rates of return. Given such an abundance of data, the tests of our null hypotheses will be very powerful. We will, therefore, want to distinguish between statistical significance and practical significance. For instance, it may be that the null hypothesis $\phi_{12,k} = 0$ $(k = 1,2,\ldots,K)$, following from the fitting of (2), can be rejected at, say, the 1% significance level. In order to assess the magnitude of the effect found, we would want to measure the extent of the gains in predictability resulting from the addition of previous market rates of return to the information set. To do this we find the ratio of the, degrees of freedom corrected, mean squared error for the full model to that of the model where the $\phi_{12}$ are restricted to be zero. The restricted model, of course, corresponds to the case where the individual rate of return is predicted on the basis of its past history alone. In this way we obtain a direct measure of the strength of any leads or lags found.

## 3. Empirical Results

The data analyzed in this study are daily rates of return on 27 heavily traded and 22 lightly traded corporate stocks.[2] Table 1 provides a listing of the stocks used. For the heavily traded stocks our sample consisted of 1,515 observations from January 3, 1975 to December 31, 1980. Our series for lightly traded stocks contained 1,010 observations for the period beginning January 3, 1977 and ending December 31, 1980. Data were taken from the CRSP tape. Our market rate was the CRSP value weighted index of N.Y.S.E. and AMEX stocks.

---

Insert Table 1 about here

---

We estimated the equations (2) and (3), testing the respective null hypotheses $\phi_{12,k} = 0$ (k=1, 2, ..., k) and $\phi_{21,k} = 0$ (k=1, 2, ..., K). For the maximum permitted lag, K, we used values of both 5 and 10. We verified that this was adequate by examining the multivariate partial autocorrelations, as discussed, for example, in Ansley and Newbold (1979). There exists no procedure which can be shown to be optimal for choosing the best value of K in Granger causality tests. A high order autoregression is intended to provide an adequate, computationally convenient, representation of the underlying stochastic generating process. A value of K that is "too low" will invalidate the assumed significance levels of the tests, while a value that is "too high" will lower the power of the tests. Since our sample sizes are large, we felt that the second of these considerations was less important than the first. Time series analysts generally take one of two approaches to order determination in autoregressive models. One possibility is to employ one of the order estimation criteria, whose theoretical properties are discussed in the context of univariate time series models by Hannan (1980). Alternatively, model selection can be based on the sample partial autocorrelations. For our data these sample partial autocorrelations invariably indicated that an autoregressive order of at most five would be adequate. However, to be conservative with regard to the adequacy of the autoregressive approximation, we also employed K = 10. As already indicated, for samples of this size, the effect of this more elaborate formulation on the power of the tests should be minimal.

Table 2 shows the results of our F-tests for the 27 heavily traded stocks. As can be seen from this table, we frequently found strong statistical evidence of the market rate leading the individual rate. This pattern of findings is fairly consistent, whichever maximum lag length is used. On the other hand, there was much less strong evidence of the market rate lagging the individual rates. This is perhaps not surprising as, a priori, we would not expect a great deal of success when employing movements in a single stock rate of return to predict the overall market rate.

---

Insert Table 2 about here

---

Turning now to Table 3, we can put these results in perspective. We note there that, although the statistical evidence indicating lagged relationships is very strong, our estimates suggest that the relationships themselves are rather weak. The estimated prediction mean squared error for the individual rate, when market rates were included in the information set, was never reduced to less than 97.5% of the prediction mean squared error resulting from the use of past individual rates alone. For the majority of these series, the reduction in mean squared error from the use of the additional information was less than 1%. The relationships in the other direction were even weaker. The best that could be achieved by using past individual rates was a reduction of 1% in mean squared forecast error for the market rate.

---

Insert Table 3 about here

---

In Tables 4 and 5 we present the corresponding results for the 22 lightly traded stocks in our sample. It appears from Table 4 that the evidence for market lead is even stronger here than in the case of heavily traded stocks. Moreover, we notice from Table 5 that the relationships, though still of modest strength, are rather less weak for the lightly traded stocks. Mean squared error ratios were less than 0.97 for 7 stocks when the maximum permitted lag was 5 days, and for 6 stocks when the maximum permitted lag was 10 days. Nevertheless, the smallest value found for this ratio was still as much as 0.942 in the former case and 0.931 in the latter. It seems pretty clear that, while strong statistical evidence of lags can be found, only modest gains can be expected in trying to exploit these lags in the prediction of future individual rates of return on a daily basis, even for lighly traded stocks. As is to be expected, we see from the tables that little is to be gained in attempting to use the individual rates for lightly traded stocks to predict the market rate.

---

Insert Tables 4 and 5 about here

---

We were concerned that our findings could be influenced by the omission of a relevant variable related to both individual rates and market rate.[3] Accordingly, we repeated our analysis subtracting the risk-free interest rate from individual and market rates; that is, we tested for leads and lags between $R^*_{jt}$ and $R^*_{mt}$, where

$$R^*_{jt} = R_{jt} - R_{ft}; \qquad R^*_{mt} = R_{mt} - R_{ft}$$

and $R_{ft}$ is the risk-free rate, for which we employed the daily Federal Funds rate, obtained from the Wall Street Journal. Qualitatively our findings are unchanged. The null hypotheses of no leads or lags are very often rejected at the usual significance levels. However, as shown in Tables 6 and 7, the relationships found were not terribly strong. Once again, the strongest relations appear to involve the market rate leading the rate of return on lightly traded stocks.

---

Insert Tables 6 and 7 about here

---

## 4. Implications of the Empirical Results

The empirical results of the previous section provide strong statistical evidence supporting the existence of a dynamic daily market model, confirming the findings of Reinganum (1982). However, while Reinganum's major concern was with the precision of ordinary least squares estimates of beta, we have focused attention on assessing the strength of the dynamic aspects of the relationship--distinguishing between practical significance and statistical significance.[4] Using the criterion of prediction mean squared error, we have found that the practical importance of the dynamic specification is not terribly great, though market leads over lightly traded stock yields do appear to be somewhat stronger than those over heavily traded stock yields. Certainly we have found strong evidence for the existence of leads and lags, but it appears that, for many practical purposes, these will be only of minor importance. Hence, while we find strong statistical evidence of market inefficiency, the extent of that inefficiency appears not to be very severe. This finding is generally consistent with Hillmer and Yu's (1980) concerns about markets' adjustment speed with respect to information release.

Moreover, two additional factors suggest that any underlying relationships may be even weaker than we have reported. Newbold (1978) has shown that, if one or other of a pair of time series variables is measured with errors, spurious causal relationships can arise between the measured variables, and any true relationship can be magnified. Since the market rate of return index used in practice is a proxy for the unobserved "true" rate, measurement errors could constitute a partial explanation of some of our findings of dynamic relationships. In addition, Tiao and Wei (1976) have shown that, through time aggregation, inherently uni-directional relationships can take on the spurious appearance of feedback. This could account for our rather surprising finding that rates of return for some lightly traded stocks appear to lead, as well as lag behind, the market rate. Although it is impossible to quantify the effects of these two factors, they must be kept in mind when attempting to interpret empirical findings on dynamic specification.

## 5. Summary

In this paper, we have considered the possibility of departure from the simple static market model for individual rates of return, contemplating the existence of leads and lags in the system. Using long series of daily data, for both heavily traded and lightly traded stocks, we have found strong statistical evidence indicating that the simple market model is mis-specified, when testing against alternatives involving a market lead. This evidence is particularly strong for the lightly traded stocks. We conclude, however, that the practical extent of this mis-specification is not terribly severe, though it is more so for lightly traded than for heavily traded stocks. In the former case, certainly, it seems clear that modest gains in forecast quality can be achieved when past values of the market rate are used to aid in the prediction of individual rates of return.

## Footnotes

[1]Any instantaneous relationship between the two series is absorbed in the off-diagonal elements of the covariance matrix $\Sigma$.

[2]These were the most and least heavily traded stocks on which we were able to obtain complete daily records in the sampling period. Trading volume was used to provide this measure.

[3]We are grateful to an anonymous referee for this suggestion.

[4]In principle, our results support both the Dimson (1979) and Reinganum (1982) methods of estimating beta coefficients. However, our results indicate that the market lag variables in their specifications are likely to be less important than the market lead variables.

## References

Ansley, C. F. and Newbold, P. (1979). Multivariate partial autocorrelations, Proceedings of Business and Economic Statistics Section, American Statistical Association, 349-53.

Dimson, E. (1979). Risk measurement when shares are subject to infrequent trading, Journal of Financial Economics 7, 197-226.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods, Econometrica 37, 424-38.

Hannan, E. J. (1980). The estimation of the order of an ARMA process, Annals of Statistics 8, 1071-81.

Hillmer, S. C. and P. L. Yu (1980). Markets' adjustment speed with respect to information release, Journal of Financial Economics 7, 321-45.

Newbold, P. (1978). Feedback induced by measurement errors, International Economic Review 19, 787-91.

Reinganum, M. R. (1982). A direct test of Roll's conjecture on the firm size effect, Journal of Finance 37, 27-35.

Sims, C. A. (1972). Money, income and causality, American Economic Review 62, 540-52.

Tiao, G. C. and Wei, W. S. (1976). Effect of temporal aggregation on the dynamic relationship of two time series variables, Biometrika 63, 513-23.

TABLE 1

Corporate Stocks Used in the Study

(i) Heavily Traded Stocks

1. Allied Chemical
2. Aluminum Corporation of America
3. American Brands
4. American Can Corporation
5. A.T.T.
6. Bethlehem Steel
7. DuPont
8. Eastman Kodak
9. Exxon
10. General Electric
11. General Foods
12. General Motors
13. Goodyear
14. INCO, Ltd.
15. International Harvester
16. International Paper
17. Merk and Co.
18. 3M Corporation
19. Owens Illinois
20. Proctor and Gamble
21. Sears
22. Standard Oil of California
23. Texaco
24. Union Carbide
25. U. S. Steel
26. Westinghouse
27. Woolworth

(ii) Lightly Trades Stocks

1. Aro Corporation
2. Bethlehem Corporation
3. Breeze Corps. Inc.
4. Carriers and General Corp.
5. Community Public Service Company
6. Continental Metals Corp.
7. Equifax Inc.
8. Fidelity Union Bancorporation
9. First Connecticut Small Business Investment
10. First National State Corp.
11. GATX Corporation
12. Great Northern Iron Ore
13. Hastings Manufacturing Co.
14. Holiday Inns Inc., Spl. Stk. A
15. Indiana Gas Inc.
16. Jaclyn Inc.
17. O'Okeip Copper Company
18. Pacific Tin Consolidated
19. Shopwell Inc.
20. South Jersey Industries
21. Wies Markets Inc.
22. Winn-Dixie Stores Inc. B

TABLE 2

Tests for Market Leads and Lags:
Heavily Traded Stocks; F Statistics

| Stocks | Maximum Lag 5 Days Dependent Variable $R_{jt}$ | Maximum Lag 5 Days Dependent Variable $R_{mt}$ | Maximum Lag 10 Days Dependent Variable $R_{jt}$ | Maximum Lag 10 Days Dependent Variable $R_{mt}$ |
|---|---|---|---|---|
| 1 | 6.458*** | 3.620*** | 3.994*** | 2.034** |
| 2 | 2.768** | 0.838 | 1.673* | 0.671 |
| 3 | 2.666** | 0.432 | 2.730*** | 0.463 |
| 4 | 8.239*** | 1.255 | 4.525*** | 1.475 |
| 5 | 1.491 | 2.396** | 3.257*** | 1.740* |
| 6 | 2.194* | 0.584 | 1.735* | 1.077 |
| 7 | 2.322** | 2.493** | 3.053*** | 1.436 |
| 8 | 1.591 | 0.828 | 1.739* | 1.483 |
| 9 | 3.411*** | 2.847** | 2.652*** | 1.973** |
| 10 | 1.975* | 1.358 | 1.816* | 0.670 |
| 11 | 1.413 | 1.492 | 0.993 | 0.977 |
| 12 | 2.845** | 0.668 | 2.386*** | 0.876 |
| 13 | 6.822*** | 1.848 | 4.428*** | 1.690 |
| 14 | 2.709** | 0.714 | 1.784* | 0.981 |
| 15 | 3.645*** | 0.189 | 2.693*** | 0.315 |
| 16 | 2.136* | 1.165 | 2.805*** | 1.016 |
| 17 | 1.365 | 1.509 | 1.643* | 1.000 |
| 18 | 3.889*** | 1.365 | 4.395*** | 0.954 |
| 19 | 7.989*** | 1.399 | 4.880*** | 1.057 |
| 20 | 2.262** | 2.110* | 3.717*** | 1.578 |
| 21 | 1.216 | 2.000* | 1.865** | 1.212 |
| 22 | 1.194 | 2.706** | 1.554 | 2.441*** |
| 23 | 3.061*** | 1.138 | 2.769*** | 0.883 |
| 24 | 2.998** | 1.509 | 2.788*** | 1.725** |
| 25 | 0.952 | 0.551 | 1.084 | 0.663 |
| 26 | 1.619 | 0.744 | 1.724* | 0.864 |
| 27 | 3.018** | 1.355 | 1.968** | 0.822 |

* Significant at 10% level
** Significant at 5% level
*** Significant at 1% level

TABLE 3

Tests for Market Leads and Lags:
Heavily Traded Stocks; Mean Squared Error Ratios

| | Maximum Lag 5 Days Dependent Variable | | Maximum Lag 10 Days Dependent Variable | |
|---|---|---|---|---|
| | $R_{jt}$ | $R_{mt}$ | $R_{jt}$ | $R_{mt}$ |
| Smallest | 0.977 | 0.991 | 0.975 | 0.990 |
| 0.97 to 0.99 | 2 | 0 | 4 | 0 |
| 0.98 to 0.99 | 2 | 0 | 10 | 0 |
| 0.99 to 1.00 | 21 | 17 | 12 | 12 |
| More than 1.00 | 2 | 10 | 1 | 15 |

TABLE 4

Tests for Market Leads and Lags:
Lightly Traded Stocks; F Statistics

| Stocks | Maximum Lag 5 Days Dependent Variable $R_{jt}$ | Maximum Lag 5 Days Dependent Variable $R_{mt}$ | Maximum Lag 10 Days Dependent Variable $R_{jt}$ | Maximum Lag 10 Days Dependent Variable $R_{mt}$ |
|---|---|---|---|---|
| 1 | 4.915*** | 0.897 | 3.612*** | 0.928 |
| 2 | 7.352*** | 0.566 | 3.460*** | 0.853 |
| 3 | 2.571** | 0.508 | 1.794* | 0.945 |
| 4 | 13.405*** | 2.097* | 8.288*** | 2.604*** |
| 5 | 4.930*** | 1.775 | 3.336*** | 1.686* |
| 6 | 2.615** | 1.250 | 1.759* | 1.059 |
| 7 | 7.463*** | 0.525 | 5.327*** | 0.568 |
| 8 | 8.598*** | 1.393 | 5.093*** | 0.783 |
| 9 | 2.734** | 0.708 | 2.121** | 1.026 |
| 10 | 8.195*** | 0.279 | 4.854*** | 0.628 |
| 11 | 6.378*** | 1.373 | 3.958*** | 1.763* |
| 12 | 0.923 | 0.118 | 1.022 | 0.524 |
| 13 | 4.474*** | 0.276 | 2.312** | 0.629 |
| 14 | 2.699** | 1.240 | 1.758* | 2.160** |
| 15 | 8.328*** | 1.461 | 5.050*** | 1.162 |
| 16 | 3.750*** | 1.352 | 2.573*** | 1.023 |
| 17 | 2.988** | 0.880 | 2.125** | 1.009 |
| 18 | 1.337 | 0.954 | 1.611* | 0.856 |
| 19 | 2.636** | 2.046* | 1.628* | 1.122 |
| 20 | 3.184*** | 1.728 | 1.794* | 1.935** |
| 21 | 11.333*** | 1.461 | 7.211*** | 2.051** |
| 22 | 2.372** | 3.035*** | 1.681* | 2.299** |

* Significant at 10% level
** Significant at 5% level
*** Significant at 1% level

TABLE 5

Tests for Market Leads and Lags:
Lightly Traded Stocks; Mean Squared Error Ratios

| | Maximum Lag 5 Days Dependent Variable | | Maximum Lag 10 Days Dependent Variable | |
|---|---|---|---|---|
| | $R_{jt}$ | $R_{mt}$ | $R_{jt}$ | $R_{mt}$ |
| Smallest | 0.942 | 0.990 | 0.931 | 0.984 |
| 0.93 to 0.94 | 0 | 0 | 1 | 0 |
| 0.94 to 0.95 | 1 | 0 | 1 | 0 |
| 0.95 to 0.96 | 1 | 0 | 1 | 0 |
| 0.96 to 0.97 | 5 | 0 | 3 | 0 |
| 0.97 to 0.98 | 1 | 0 | 4 | 0 |
| 0.98 to 0.99 | 5 | 0 | 4 | 3 |
| 0.99 to 1.00 | 8 | 10 | 7 | 7 |
| More than 1.00 | 1 | 12 | 1 | 12 |

TABLE 6

Tests for Market Leads and Lags with Risk-free Rate Adjusted Data: Heavily Traded Stocks; Mean Squared Error Ratios

| | Maximum Lag 5 Days Dependent Variable | | Maximum Lag 10 Days Dependent Variable | |
|---|---|---|---|---|
| | $R^*_{jt}$ | $R^*_{mt}$ | $R^*_{jt}$ | $R^*_{mt}$ |
| Smallest | 0.990 | 0.955 | 0.978 | 0.990 |
| 0.94 to 0.96 | 0 | 1 | 0 | 0 |
| 0.96 to 0.98 | 0 | 2 | 1 | 0 |
| 0.98 to 1.00 | 24 | 17 | 23 | 26 |
| More than 1.00 | 3 | 7 | 3 | 1 |

TABLE 7

Tests for Market Leads and Lags with Risk-free Rate Adjusted Data: Lightly Traded Stocks; Mean Squared Error Ratios

| | Maximum Lag 5 Days Dependent Variable | | Maximum Lag 10 Days Dependent Variable | |
|---|---|---|---|---|
| | $R^*_{jt}$ | $R^*_{mt}$ | $R^*_{jt}$ | $R^*_{mt}$ |
| Smallest | 0.965 | 0.985 | 0.912 | 0.979 |
| 0.90 to 0.92 | 0 | 0 | 1 | 0 |
| 0.92 to 0.94 | 0 | 0 | 0 | 0 |
| 0.94 to 0.96 | 0 | 0 | 2 | 0 |
| 0.96 to 0.98 | 6 | 0 | 5 | 1 |
| 0.98 to 1.00 | 15 | 16 | 14 | 18 |
| More than 1.00 | 1 | 6 | 0 | 3 |