

Division of Research
Graduate School of Business Administration
The University of Michigan

December 1984

THE ROLE OF DIVIDEND CHANGES IN THE PREDICTION
OF FUTURE CHANGES IN EARNINGS: AN EMPIRICAL
ANALYSIS OF QUARTERLY DATA

Working Paper No. 408

Antonio Zoratto Sanvicente
Visiting Assistant Professor
The University of Michigan
and
Professor Livre-Docente
Universidade de Sao Paulo, Brazil

FOR DISCUSSION PURPOSES ONLY

None of this material is to be quoted or
reproduced without the expressed permission
of the Division of Research.

1. Introduction

One of the first authors to suggest that current dividends depended not only on current and past earnings, but also on future earnings, was Lintner [1956]. Based on the results of interviews with the managers of twenty-eight U.S. industrial corporations in the early 1950's, Lintner posited the following model of changes in the firms' dividend policies:

$$\Delta D_{i,t} = a_i + c_i (D_{i,t}^* - D_{i,t-1}) + u_{i,t} \quad (1)$$

where $\Delta D_{i,t} \equiv D_{i,t} - D_{i,t-1}$ = observed change in dividends per share declared by firm i in period t ;

$D_{i,t}^*$ \equiv desired dividend level = $r_i P_{i,t}$, where r_i is firm i 's target payout ratio, and $P_{i,t}$ is its earnings per share in period t .

According to Lintner's own discussion of this model, it is important to note that:

- (1) t is measured in units of one year;
- (2) the constant a_i should generally be positive, reflecting the greater reluctance to decrease dividends; and
- (3) $u_{i,t}$ "...represents the discrepancy between the observed $\Delta D_{i,t}$ and that expected on the basis of other terms in the equation. It will absorb discrepancies due to each company's preference for dividend rates in rounded units per share, as well as the impact of all other considerations insofar as they are not systematically reflected in the values assigned to the two parameters c_i and r_i and the constant term, which is in the nature of a trend factor." (Lintner [1956, pp. 107-108] In other words, $u_{i,t}$ is an error term which is uncorrelated with the independent variables $D_{i,t}^*$ and $D_{i,t-1}$).

In a reply to a comment by Durand [1959], Modigliani and Miller [1959] also discussed the relationship between current dividends and expected earnings, and labeled it the 'information content of dividends.' In particular, Modigliani and Miller conjectured that changes in dividends should tend to reflect an alteration in management's views about the firm's future earnings prospects. This conclusion drew upon evidence, also gathered by Lintner, of the tendency by firms to follow stable dividend policies.

In this paper, past use of models of changes in dividend policy is reviewed. In the first place, it is seen that, if the Lintner model is assumed to be correct, obvious omitted-variables problems arise in the alternative models. This poses consequences for inferences about the information content hypothesis, which are pointed out in the next section.

In section 3, the discussion shows that even strict adherence to Lintner's own formulation also leads to econometric problems in empirical applications based on ordinary least squares regression. The hypothesis of interest is formalized in section 4, which also proposes an alternative approach - multiple time series analysis. This approach circumvents the specification and estimation problems inherent to the use of ordinary least squares in tests of the hypothesis of information content of dividends. Section 5 contains part of the preliminary data analysis, as well as the sample and the data utilized. It is indicated there that conditioning predictions of earnings changes on non-zero changes in dividends may be a more fruitful design in testing the hypothesis of information content. The final results are reported in section 6.

While proposing a new procedure (multiple time series analysis) for testing the "information content hypothesis," the present paper addresses a necessary condition for dividend changes conveying information about future changes in earnings.

As pointed out in Sanvicente [1982] and in Handjinicolaou and Kalay [1984], positive (negative) abnormal stock price changes in response to unexpectedly high (low) dividends were interpreted in the past as sufficient evidence about the existence of significant informational effects; however, such price responses are also consistent with a hypothesis of wealth redistribution from bondholders to stockholders, and vice-versa. Thus, examination of bond price reactions is a better procedure for ascertaining which of the possible effects is dominant.

However, for dividend changes to convey information about future changes in earnings, it is required, at a more basic level, that the formation of earnings expectations makes use of dividend change data. No price reaction will be observed (abnormal or otherwise) in security prices if dividends are not used by market participants in revising their estimates of future earnings.

By examining the joint time-series properties of quarterly earnings and dividends, on a per-share basis, the present paper is an attempt at answering that more basic question.

2. Lintner's Model and Econometric Problems in Some Empirical Tests of the Hypothesis of Information Content

As Lintner himself indicated, without affecting the error term equation (1) can be rewritten as

$$D_{i,t} = a_i + b_i P_{i,t} + d_i D_{i,t-1} + u_{i,t} \quad (2)$$

where $b_i \equiv c_i r_i$, and $d_i \equiv 1 - c_i$.

Hence, we can reformulate the model for changes in dividends per share by writing

$$\begin{aligned}\Delta D_{i,t} &\equiv D_{i,t} - D_{i,t-1} = a_i + b_i P_{i,t} + d_i D_{i,t-1} + u_{i,t} - \\ &\quad - a_i - b_i P_{i,t-1} - d_i D_{i,t-2} - u_{i,t-1} = \\ &= b_i \Delta P_{i,t} + d_i \Delta D_{i,t-1} + \varepsilon_{i,t}\end{aligned}\quad (3)$$

where $\varepsilon_{i,t} = u_{i,t} - u_{i,t-1}$.¹ It will also be useful to write equation (3) in the following way:

$$\Delta D_{i,t} = b_{1,i} P_{i,t} + b_{2,i} P_{i,t-1} + d_{1,i} D_{i,t-1} + d_{2,i} D_{i,t-2} + \varepsilon_{i,t} \quad (3')$$

where $b_{1,i} = b_i$, $-b_{2,i} = b_i$, $d_{1,i} = d_i$, and $-d_{2,i} = d_i$.

Fama and Babiak [1968] examined several 'forms' of Lintner's model,² and concluded that the best predictive form - in terms of R^2 and prediction errors relative to a holdout period of dividend changes - was

$$\Delta D_{i,t} = b'_{1,i} P_{i,t} + b'_{2,i} P_{i,t-1} + d'_{1,i} D_{i,t-1} + \varepsilon'_{i,t} \quad (4)$$

Watts [1973] used the Fama-Babiak form of Lintner's model to test the information-content-of-dividends hypothesis. He used annual earnings and dividends per share of 310 firms for the 1947-1966 period. His procedure involved two steps:

- (1) estimate equation (4) by ordinary least squares and compute

$$\begin{aligned}\hat{\varepsilon}'_{i,t} &= \Delta D_{i,t} - E(\Delta D_{i,t}) = \Delta D_{i,t} - \hat{b}'_{1,i} P_{i,t} - \hat{b}'_{2,i} P_{i,t-1} - \\ &\quad - \hat{d}'_{1,i} D_{i,t-1}\end{aligned}$$

where the $\hat{\varepsilon}'_{i,t}$ are interpreted as estimates of the unexpected change in dividends in year t .

- (2) estimate θ_1 in the model

$$P_{i,t+1} \equiv P_{i,t+1} - P_{i,t} = \gamma_i + \theta_i \hat{\varepsilon}'_{i,t} + \omega_{i,t} \quad (5)$$

With the null hypothesis being $\theta_i = 0$, and its alternative $\theta_i > 0$, information content would be said to present in dividend changes if this null hypothesis were rejected for its alternative. Watts found the average $\hat{\theta}_i$ to be equal to 0.508, with an average t-statistic of 0.406. His conclusion was:

"...the time-series regressions suggest that the relationship between the unexpected change in dividends and the change in future earnings is positive on the average. However, the relationship is hardly general or generally strong." (Watts [1973, p. 201])

If one takes the Lintner model as the correct expression of dividend policy, it is clear that the Fama-Babiak and Watts versions introduce an omitted-variables problem by ignoring the term in $d_{2,i}^D D_{i,t-2}$ - simply compare equations (3') and (4).

The consequences of the omission of relevant variables are discussed, for example, in Johnston [1972] and Pindyck and Rubinfeld [1981]: use of ordinary least squares in estimating (4) produces biased and inconsistent estimators. The direction of bias, in the omitted-variables problem, depends both on the sign of the true coefficient of the excluded variable, and on the variable's covariance with the included variables.

Denoting by $\hat{\varepsilon}_{i,t}$ the true value of the estimate of unanticipated dividend change, we can write

$$\hat{\varepsilon}_{i,t} = \hat{\varepsilon}'_{i,t} - v_{i,t} \quad (6)$$

where $v_{i,t}$ is the measurement error implied by using $\hat{\varepsilon}'_{i,t}$ instead of $\hat{\varepsilon}_{i,t}$, due to the omission of $D_{i,t-2}$.

Since the true regression model for the second stage of the Watts procedure in equation (5), but he in fact uses

$$\Delta P_{i,t+1} = \gamma_i + \theta_i \hat{\varepsilon}'_{i,t} + (\omega_{i,t} - \theta_i v_{i,t}) = \gamma_i + \theta_i \hat{\varepsilon}'_{i,t} + \omega'_{i,t} \quad (7)$$

if $v_{i,t}$ and $\omega_{i,t}$ are uncorrelated,

$$\text{cov}(\omega'_{i,t}, \hat{\varepsilon}'_{i,t}) = E[(\omega_{i,t} - \theta_i v_{i,t})(\hat{\varepsilon}'_{i,t} - v_{i,t})] = \theta_i \sigma_v^2 \neq 0,$$

so that the least squares estimate of θ_i is biased and inconsistent:

$$\text{plim } \hat{\theta}_i = \frac{\theta_i}{1 + \sigma_v^2 / \text{var}(\hat{\varepsilon}'_{i,t})}$$

implying that the coefficient θ_i will be underestimated when ordinary least squares regression is applied (as done by Watts, see p. 200 of his article [1973]). This will penalize the alternative hypothesis, of information content in unanticipated dividend changes.

3. Econometric Problems When the 'Exact' Lintner Model is Used

Note that equation (3), upon successive substitution of $\Delta D_{i,t-k}$, for $k = 1, 2, \dots, \ell$ can be rewritten as³

$$\begin{aligned} \Delta D_{i,t} &= b_i \Delta P_{i,t} + d_i (b_i \Delta P_{i,t-1} + d_i \Delta D_{i,t-2} + \varepsilon_{i,t-1}) + \varepsilon_{i,t} = \\ &= b_i \Delta P_{i,t} + d_i b_i \Delta P_{i,t-1} + d_i^2 (b_i \Delta P_{i,t-2} + d_i \Delta D_{i,t-3} + \\ &\quad + \varepsilon_{i,t-2}) + d_i \varepsilon_{i,t-1} + \varepsilon_{i,t} = \\ &= b_i \sum_{k=0}^{\ell} d_i^k \Delta P_{i,t-k} + d_i^\ell \Delta D_{i,t-\ell} + \sum_{k=0}^{\ell} d_i^k \varepsilon_{i,t-k} \end{aligned} \quad (8)$$

Hence, it is actually the case that Lintner's model - in its 'exact' version in the form of an 'adaptive regression' scheme - expresses the current change in dividend policy as a discrete linear combination of (1) current and past values of changes in earnings per share, (2) past dividend changes, and (3) current and past disturbances.

Therefore, estimation of (3) by linear regression in the first place reflects some arbitrary decision about the value of ℓ - the number lags. This may introduce misspecification bias in the form of omitted variables. It is also apparent that regression model poses two serious econometric problems: (1) lagged dependent variable, caused by the presence of $\Delta D_{i,t-2}$ on the right-hand side of (3), and (2) assumes away autocorrelated errors, which may well exist, as indicated by (8). Ordinary least squares estimation of (3) will produce biased and inconsistent estimators (see Johnston [1972]), who also discusses estimation techniques for models with such problems).

4. Multiple Time Series Analysis

Concomitant reference to Lintner's model - or any other model of changes in dividend policy, such as the Fama-Babiak and Watts versions - and proposing a model for the impact of dividend changes on changes in earnings in subsequent periods, may be represented by a simultaneous equation formulation, viz:

$$\Delta D_{i,t} = \pi_{i,1,0} + \sum_{j=1}^{\ell} \pi_{i,1,j} \Delta P_{i,t-j} + \sum_{j=1}^{\ell} \psi_{i,1,j} \Delta D_{i,t-j} + e_{i,1,t} \quad (9)$$

$$\Delta P_{i,t} = \pi_{i,2,0} + \sum_{j=1}^{\ell} \pi_{i,2,j} \Delta P_{i,t-j} + \sum_{j=1}^{\ell} \psi_{i,2,j} \Delta D_{i,t-j} + e_{i,2,t} \quad (10)$$

In such a formulation, testing for the significance of any $\psi_{i,2,j}$, for $j = 1, 2, \dots, \ell$ lags, would pose no special problems. Since all right-hand side variables are pre-determined, the system is identified; furthermore, it is already in its reduced form. Hence, ordinary least squares could be applied to each equation separately. This presupposes, however, that (1) the error terms in each equation are not serially correlated, and that (2) there are no omitted-variables problems, which is equivalent to knowing the value of ℓ for each term in both equations.

However, these potentially serious problems, which by themselves or in combination with each other eliminate the pre-determined character of the right-hand side variables, may be avoided. This will be the case if the multiple time series analysis approach is employed. For, as discussed by Tiao and Box [1981], use of this approach not only makes use of maximum likelihood estimation algorithms, but permits us to obtain models whose structure is determined by the data. In other words, it allows us to jointly determine the value of k and the structures of $e_{i,1,t}$ and $e_{i,2,t}$.

In this approach, which is summarized in the Appendix, which also contains an example formed by an equally-weighted portfolio of the firms in the sample described in the next section of the chapter, an iterative procedure is followed. This involves three stages:

- (1) identification, or preliminary specification of possible candidate models, embodied in the determination of the maximum orders of the autoregressive and the moving average components;
- (2) estimation, which is performed by maximum likelihood algorithms; and
- (3) diagnostic checking of estimated models, with particular interest in determining whether the resulting residuals can be considered as white noise.

In this approach, and because the variables of interest are changes in dividends and earnings, if we let

$$z_{1,t} \equiv \Delta D_{i,t}$$

$$z_{2,t} \equiv \Delta P_{i,t}$$

for any firm i , vector autoregressive-moving average (ARMA) models can be represented as

$$\begin{aligned}
 \begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix} &= \begin{bmatrix} \phi_{1,11} & \phi_{1,12} \\ \phi_{1,21} & \phi_{1,22} \end{bmatrix} \begin{bmatrix} z_{1,t-1} \\ z_{2,t-1} \end{bmatrix} + \begin{bmatrix} \phi_{2,11} & \phi_{2,12} \\ \phi_{2,21} & \phi_{2,22} \end{bmatrix} \begin{bmatrix} z_{1,t-2} \\ z_{2,t-2} \end{bmatrix} + \\
 &+ \dots + \begin{bmatrix} \phi_{p,11} & \phi_{p,12} \\ \phi_{p,21} & \phi_{p,22} \end{bmatrix} \begin{bmatrix} z_{1,t-p} \\ z_{2,t-p} \end{bmatrix} + \begin{bmatrix} a_{1,t} \\ a_{2,t} \end{bmatrix} - \\
 &- \begin{bmatrix} \theta_{1,11} & \theta_{1,12} \\ \theta_{1,21} & \theta_{1,22} \end{bmatrix} \begin{bmatrix} a_{1,t-1} \\ a_{2,t-1} \end{bmatrix} - \begin{bmatrix} \theta_{2,11} & \theta_{2,12} \\ \theta_{2,21} & \theta_{2,22} \end{bmatrix} \begin{bmatrix} a_{1,t-2} \\ a_{2,t-2} \end{bmatrix} - \\
 &- \dots - \begin{bmatrix} \theta_{q,11} & \theta_{q,12} \\ \theta_{q,21} & \theta_{q,22} \end{bmatrix} \begin{bmatrix} a_{1,t-q} \\ a_{2,t-q} \end{bmatrix}
 \end{aligned}$$

which is equivalent to

$$\begin{aligned}
 z_{1,t} &= \phi_{1,11}z_{1,t-1} + \phi_{1,12}z_{2,t-1} + \dots + \phi_{p,11}z_{1,t-p} + \phi_{p,12}z_{2,t-p} + \\
 &+ a_{1,t} - \theta_{1,11}a_{1,t-1} - \theta_{1,12}a_{2,t-1} - \dots - \theta_{q,11}a_{1,t-q} - \\
 &- \theta_{q,12}a_{2,t-q}
 \end{aligned} \tag{11}$$

$$\begin{aligned}
 z_{2,t} = & \phi_{1,21}z_{1,t-1} + \phi_{1,22}z_{2,t-1} + \dots + \phi_{p,21}z_{1,t-p} + \phi_{p,22}z_{2,t-p} + \\
 & + a_{2,t} - \theta_{1,21}a_{1,t-1} - \theta_{1,22}a_{2,t-1} - \dots - \theta_{q,12}a_{1,t-q} - \\
 & - \theta_{q,22}a_{2,t-q}
 \end{aligned} \tag{12}$$

where $\{\phi_p\}$ and $\{\theta_q\}$ are the coefficients of the autoregressive and moving-average components, respectively, p is the maximum order of the autoregressive part, and q the maximum order of the moving-average component. The $a_{1,t-k}$, $a_{2,t-k}$, for $k = 0, 1, 2, \dots, q$ are normally distributed, serially uncorrelated random variables, with zero mean and constant variances $\sigma_{a_1}^2$ and $\sigma_{a_2}^2$.

Since the interest is in finding out whether changes in dividends contribute significantly to predictions of changes in earnings in subsequent periods, a test of the information content hypothesis may be postulated quite simply as (see equation (12)):⁴

$$H_0: \phi_{1,21} = \phi_{2,21} = \dots = \phi_{p,21} = 0$$

$$H_A: \text{at least one } \phi_{k,21} > 0, \text{ for } k = 1, 2, \dots, p$$

5. Results of Applying Multiple Time Series Analysis

For a sample of twenty five industrial corporations, quarterly dividends and earnings, adjusted for stock splits and stock dividends, were collected for the 1951-1980 period from several issues of Value Line's Ratings and Opinions. The sample firms and the respective periods covered are listed in Table 1.

In turn, Table 2 reports the individual cross-correlations computed for lags $\ell = -1, 0, 1, 2, 3$, and 4. That is, Table 2 reports the values of

Table 1

LIST OF 25 U.S. CORPORATIONS IN THE SAMPLE FOR THE TEST OF THE
HYPOTHESIS OF INFORMATION CONTENT OF DIVIDENDS

<u>Name of Corporation</u>	<u>Period (Year, quarter)</u>	<u>Number of observations</u>
Allegheny Ludlum	1951-I to 1980-IV	120
Amax, Inc.	1951-I to 1980-IV	120
Athlone Inds.	1951-I to 1980-IV	120
Braniff Internat.	1951-I to 1980-IV	120
Chrysler	1951-I to 1980-IV	120
Columbia Pictures	1951-I to 1980-IV	120
Crane	1951-I to 1980-IV	120
Dresser Inds.	1951-I to 1980-IV	120
Fairchild Inds.	1952-I to 1980-IV	116
Federal Paperboard	1952-I to 1980-IV	116
Ferro	1954-I to 1980-IV	108
Flintkote	1951-I to 1980-IV	120
Fruehauf	1951-I to 1980-IV	120
General Tire and Rub.	1951-I to 1980-IV	120
Internat. Harvester	1951-I to 1980-IV	120
Lone Star Inds.	1951-I to 1980-IV	120
M. Lowenstein & Sons	1951-I to 1980-IV	120
Rockwell Internat.	1955-I to 1980-IV	104
SCM	1951-I to 1980-IV	120
Signal Companies	1951-I to 1980-IV	120
Smith (A.O.)	1951-I to 1980-IV	120
Tenneco, Inc.	1951-I to 1980-IV	120
Twent. Cent. Fox	1952-I to 1980-IV	116
United Brands	1956-I to 1980-IV	100
U.S. Inds.	1951-I to 1980-IV	120

Table 2

CROSS-CORRELATION COEFFICIENTS: ALL DIVIDEND CHANGES INCLUDED

	Value of ℓ					
	-1	0	1	2	3	4
All. Ludlum	-0.1042	-0.0050	0.0217	0.0525	0.0228	0.0843
Amax	0.3687*	0.0544	0.0544	-0.1081	0.2413*	0.0183
Athlone	0.2115*	-0.0130	0.0111	-0.1641	0.1694	-0.0274
Braniff	0.0401	0.0250	0.0238	-0.0400	0.0600	-0.0419
Chrysler	0.1393	-0.0650	0.1056	-0.1170	0.3252*	-0.2663*
Columbia	0.0292	-0.0260	0.0963	-0.0444	-0.0003	0.0445
Crane	0.4556*	-0.0260	0.3338*	-0.3177*	-0.0211	-0.0318
Dresser	0.2215*	-0.0800	0.0117	0.0266	0.1737	-0.0913
Fairchild	-0.0821	0.0180	0.0444	-0.0247	0.0161	0.0442
Fed. Ppboard	-0.0031	-0.0640	0.0616	0.1115	-0.0229	0.0353
Flintkote	0.0109	0.0000	0.0626	-0.0323	-0.0160	0.0065
Ferro	0.1097	-0.0710	0.0688	-0.0041	0.1262	-0.1375
Fruehauf	-0.0546	0.0340	-0.0148	0.0331	0.1222	0.0142
Gen. Tire	0.0868	-0.1290	0.1461	-0.0409	0.1079	-0.1538
Int. Harv.	0.0281	0.0590	-0.3304*	0.0553	0.3519*	-0.0605
Lowenstein	0.0626	0.3410*	-0.0655	0.1357	0.0362	0.0419
Lone Star	-0.0128	-0.0070	0.0680	-0.0150	-0.0289	-0.0013
Rockwell	0.1755	-0.0590	0.0576	-0.0203	0.1309	-0.0909
SCM	0.0019	0.2510*	-0.1279	0.0826	-0.1058	0.2539*
Smith	0.0864	0.0440	0.1079	-0.0519	-0.1052	0.0012
Signal	-0.1270	0.1890*	-0.0060	-0.0136	-0.0178	0.0077
Tenneco	0.0530	-0.1000	0.1124	-0.0216	0.0387	-0.0873
T. C. Fox	0.1808	-0.1950*	0.2312*	-0.1213	0.0512	-0.1132
Unit. Brds.	-0.0291	-0.0420	0.1778	0.0792	-0.0196	-0.0228
U.S. Inds.	0.1127	0.0060	0.0100	0.0439	0.0396	-0.0923

Note: with n ranging from 120 to 100, $n^{-1/2}$ ranges from 0.091 to 0.100; correspondingly, the limit of two standard errors is given by 0.182 to 0.200. The coefficients greater than two standard errors in absolute value are indicated by *.

$$\hat{\rho}_{1,2}(\ell) = \frac{\{\sum_t (z_{1,t} - \bar{z}_1)(z_{2,t+\ell} - \bar{z}_2)\}}{\{\sum_t (z_{1,t} - \bar{z}_1)^2 \sum_t (z_{2,t} - \bar{z}_2)^2\}^{1/2}} \quad (13)$$

where \bar{z}_1 and \bar{z}_2 are the means of the series of changes in quarterly dividends and earnings, respectively.

As indicated in the Appendix, examination of patterns of cross-correlations is one phase of the identification stage. With an approximate standard error of $n^{-1/2}$, where n is the number of observations, we can see, in Table 2, that the null hypothesis is rejected for the alternative of positive content in very few instances. Therefore, these results, which are part of the stage of model identification in the multiple time series approach, already indicate that it is not likely that one will find a pattern of positive and significant values for $\phi_{k,21}$ (see equation (12)). For the twenty-five individual firms, as well as an equally-weighted portfolio of the sample firms, the estimation and diagnostic checking stages were performed for the sake of completeness. For the equally-weighted portfolio, more detailed results may be found in the Appendix. In the case of the twenty-five individual firms, only for four firms was the null hypothesis of $\phi_{k,21} = 0$ rejected for its alternative (see section 4).

6. Earnings Changes Following Non-zero Changes in Dividends

Given that the summary results in the Appendix indicate that the best predictor of future dividends is the immediately preceding dividend, and that firms change their quarterly dividends with low frequency, combining observations of zero dividend changes and non-zero dividend changes may be responsible for the lack of cross-correlation observed. In other words, a spuriously low correlation coefficient may be due to temporal aggregation of both types of changes. This has been pointed out by Laub [1972], in a study attempting to estimate the coefficients of Lintner's model, i.e., equation (1).

With this possibility in mind, and after excluding all observations of zero dividend changes for each firm, cross-correlation coefficients were again computed for lags $\ell = -1, 0, 1, 2, 3$, and 4. The results are reported in Table 3.

It is apparent that, although there is some improvement towards concluding for information content, relative to the results in Table 2, the more reasonable conclusion is for the null hypothesis of no information content in dividend changes. In other words, given the variability of changes in earnings, the conclusion is that dividend changes have not passed the test as necessary condition for conveying information about subsequent changes in earnings. Therefore, with more careful specification and estimation procedures, the basic results of Watts [1973] are confirmed, for quarterly data, and individual firms.

Table 4 reports the results obtained when the following models are estimated:

$$\Delta P_{t+\ell} = \alpha + \beta \Delta D_t + d_t \quad (14)$$

$$\Delta P_{t+\ell} = \alpha^\dagger + \beta^\dagger \Delta D_t^\dagger + d_t \quad (15)$$

where $\Delta P_{t+\ell}^\dagger$ and ΔD_t^\dagger are defined as $\Delta P_{t+\ell}$ and ΔD_t weighted by the reciprocal of the standard deviation of residuals of individual firms' ordinary least squares regressions.

Therefore, models (14) and (15) constitute an attempt at pooling all observations about the twenty-five different firms, and, particularly in the case of (15), taking into account the different variances found in residuals from individual regressions, as is typically done in generalized least squares regression.

It is clear, from the results, that the alternative hypothesis of positive information content in quarterly dividend announcements tends to be accepted. This indicates that, in tests involving data for individual firms, the

Table 3

CROSS-CORRELATION COEFFICIENTS: FOR NON-ZERO DIVIDEND CHANGES ONLY

	Value of ℓ					
	-1	0	1	2	3	4
All. Ludlum	0.0635	-0.0215	-0.2346	0.4559*	-0.3754	-0.2392
Amax	0.1103	-0.1123	0.7366*	-0.4799*	-0.2361	-0.2267
Athlone	0.0405	-0.0087	0.3131*	-0.5449*	-0.2181	-0.0428
Braniff	0.0775	0.0569	0.0734	-0.0640	-0.0457	0.0767
Chrysler	0.2727	-0.1447	0.2719	-0.3430	0.2945	-0.4297*
Columbia	0.7806*	-0.3315	0.3810	-0.1226	-0.0438	0.4634*
Crane	0.4932*	-0.0815	0.7213*	-0.4445*	0.2685	-0.3737*
Dresser	0.0231	-0.1478	0.4902*	0.0050	0.0074	-0.2945
Fairchild	0.2270	0.0939	-0.0449	0.0062	-0.0427	0.0810
Fed. Ppboard	-0.0057	0.4756	-0.0991	-0.0403	-0.1366	-0.0223
Flintkote	0.0260	0.1433	0.0144	-0.2000	0.1575	0.1239
Ferro	0.2832	0.6526*	0.2163	-0.0559	0.0722	-0.1947
Fruehauf	-0.0677	0.1380	-0.1480	0.0972	-0.0112	-0.0199
Gen. Tire	0.2640	-0.0313	0.0146	-0.1958	0.0540	-0.0565
Int. Harv.	-0.5616*	0.3574	0.1341	-0.0403	-0.1510	-0.0148
Lowenstein	-0.1599	0.5050*	0.1633	-0.0009	-0.5019*	-0.0860
Lone Star	0.2802	0.0068	0.0232	-0.1273	0.3451	0.0488
Rockwell	0.5003*	0.6713*	0.0939	0.1508	-0.1471	-0.0852
SCM	-0.2678	0.4420*	0.0024	0.0903	-0.2896	0.0690
Smith	0.1707	0.1066	0.2881	-0.3899	-0.3697	-0.1981
Signal	-0.0029	0.2306	-0.1692	-0.0712	0.1429	-0.1912
Tenneco	0.3485	-0.2147	0.0419	-0.2323	0.6600*	-0.2756
T. C. Fox	0.7780*	-0.6044*	0.5247*	-0.5144*	0.5227*	-0.3734*
Unit. Brds.	0.3911	-0.1849	-0.1175	-0.2594	0.2909	-0.0350
U.S. Inds.	0.0262	0.0335	0.2927	-0.3329	0.1077	-0.0124

Note: n ranges from 72 to 16; hence, the two-standard error limit ranges from 0.236 to 0.500. An asterisk denotes the cases in which the coefficients are greater than two standard errors.

Table 4

SUMMARY OF REGRESSION OF ΔP_{t+l} ON ΔD_t : NON-ZERO DIVIDEND CHANGES ONLY, WITH (A) EQUAL WEIGHTS, (B) WEIGHTS EQUAL TO THE RECIPROCAL OF INDIVIDUAL REGRESSION EQUATION RESIDUALS

	Value of l					
	-1	0	1	2	3	4
Equal weights:						
$\hat{\alpha}$	0.0091	-0.0084	-0.0024	-0.0052	0.0035	-0.1009
S.E. ($\hat{\alpha}$)	0.0048	0.0050	0.0050	0.0051	0.0041	0.0058
$\hat{\beta}$	0.8866	-0.3494	1.0313	-0.8052	0.5702	-0.9008
S.E. ($\hat{\beta}$)	0.1214	0.1256	0.1276	0.1296	0.1041	0.1482
$\overline{R^2}$	0.0647	0.0088	0.0788	0.0478	0.0374	0.0462
F	53.33	7.73	65.37	38.63	29.99	36.96
D.W.	2.1279	2.0694	2.5610	2.2035	1.8258	1.8554
$\hat{\rho}$	0.2569*	-0.1005*	0.2828*	0.2186*	0.1966*	-0.2178*
Weighted by the standard deviation of residuals:						
$\hat{\alpha}^\dagger$	0.0090	-0.0053	-0.0045	-0.0061	0.0049	-0.1102
S.E. ($\hat{\alpha}^\dagger$)	0.0037	0.0037	0.0037	0.0068	0.0042	0.0070
$\hat{\beta}^\dagger$	0.4606	0.3626	0.6519	0.6455	0.7337	0.1608
S.E. ($\hat{\beta}^\dagger$)	0.0065	0.0082	0.0082	0.0073	0.0042	0.1218
$\overline{R^2}$	0.0617	0.0237	0.0769	0.0944	0.2945	0.0010
F	50.67	19.45	63.77	79.06	312.02	1.74
D.W.	2.0801	1.9506	2.2335	2.0817	2.0258	2.0492
$\hat{\rho}^\dagger$	0.2508*	0.1582*	0.2796*	0.3092*	0.5436*	0.0485
Number of obs.:	757	760	754	750	746	741

Note: for prob. = 0.05, $F_{1,\infty} = 3.84$

information potential is obscured not only by the variability of earnings, but by inefficient estimates resulting from the small sample sizes.

7. Summary and Conclusions

The results of this chapter indicate that some evidence exists for concluding for the existence of information content in dividend changes. Once more, the test of the information content hypothesis was posed as the test of dividend changes as necessary conditions for predictions of earnings changes. While market agents may indeed use dividend changes in their models when forecasting earnings changes, the actual dividend changes may have already been anticipated, at the time of their occurrence. In any event, however, this paper's formulation of forecasting models for both earnings and dividend changes, taking into account possible interaction between the two series, can be useful for actually measuring unanticipated dividend changes.

In future research, unanticipated dividend changes measured as the difference between actual changes and changes forecast with the aid of multiple time series analysis may be employed in association tests with observed security price changes. In the particular case of corporate bonds, a bond valuation model should be used in order to account for the expected impact of concomitant changes in other determining factors - such as interest rate, time to maturity, variance of rate of return on the firm's assets - in addition to the change in dividend policy. For this purpose, an analytical framework distinguishing between wealth redistribution and informational effects should prove useful.

Footnotes

1. Under the assumption that

$$E(u_{i,t}) = 0, \text{ var}(u_{i,t}) = \sigma_u^2 \text{ for all } t, \text{ and } E(u_{i,t}u_{i,t-k}) = 0 \text{ for all } k \neq 0,$$

$$E(\varepsilon_{i,t}) - E(u_{i,t} - u_{i,t-1}) = 0,$$

$$\begin{aligned} \text{var}(\varepsilon_{i,t}) &= \text{var}(u_{i,t}) + \text{var}(u_{i,t-1}) - 2 \text{cov}(u_{i,t}, u_{i,t-1}) = \\ &= 2\sigma_u^2 = \sigma_\varepsilon^2 \text{ for all } t, \text{ and} \end{aligned}$$

$$\begin{aligned} E(\varepsilon_{i,t}\varepsilon_{i,t-k}) &= E[(u_{i,t} - u_{i,t-1})(u_{i,t-k} - u_{i,t-k-1})] = \\ &= E[(u_{i,t}u_{i,t-k}) - (u_{i,t-1}u_{i,t-k}) - (u_{i,t}u_{i,t-k-1}) + \\ &\quad + (u_{i,t-1}u_{i,t-k-1})] = 0 \text{ for all } k \neq 0. \end{aligned}$$

Hence, the new error term, $\varepsilon_{i,t}$, retains the properties of the original error term, $u_{i,t}$.

2. It seems apparent that, in modeling dividend policy changes, Lintner did not mean only that dividend changes are a function of current and past earnings and past dividends, as implied by equation (4). Rather, by postulating (1) as a model - with (2), (3) and (3') as consequences - Lintner meant that dividend changes are a function of current and past earnings and past changes in dividends.

Because of this issue, I refer to (1) and its algebraic equivalents - equations (2), (3), and (3') - as 'exact' forms of Lintner's model; equations such as (4) are called 'predictive' versions of Lintner's model.

Note that Fama and Babiak [1968] did not have as their objective testing Lintner's model; they simply compared several alternatives of equation (4): with and without an intercept, with an without lagged earnings, with

and without depreciation expense, the latter with the support of the results obtained by Brittain [1966]. In all their models, the failure to address the determination of the appropriate number of lags is a fixture; hence, in all their specifications one may argue that there is an omitted-variables problem, relative to what is seen in equation (8).

3. l denotes a finite number of lags, after which changes in earnings and changes in dividends do not contribute to the firm's policy relative to changes in dividends. If one proceeds with the substitution, i.e., $l \rightarrow \infty$, the term in $\Delta D_{i,t-k}$ should vanish, as d_i is less than unity.
4. Note that Watts [1973], in his test of the hypothesis, did not use the changes in dividends, but an estimate of unanticipated changes. He argued that changes in dividends would not possess information content if they were fully anticipated. While this may be correct when the impact of announcements is measured in terms of stock or bond price changes, it seems sufficient to verify whether expectations of changes in earnings are revised in the same direction of current changes in dividends. This is what will be the case if terms in $z_{1,t-k}$ are present in the estimated versions of equation (12).

In summary, for dividend changes to have information content about future earnings changes, it is at least necessary that expectations of earnings changes depend, on a positive way, on current and past changes in dividends, in addition to the current and past changes in earnings.

Appendix

Implementation of Multiple Time Series Analysis: Identification,
Estimation, and Diagnostic Checking of Equations (11) and (12)

Equations (11) and (12) constitute the representation of a vector autoregressive-moving average model for a two-series problem. For this, we can use the procedures and techniques developed by Tiao and Box [1981]. They are extensions of the univariate, linear discrete models of Box and Jenkins [1976].

In this appendix, a description of the Tiao and Box approach to the modeling of multiple time series is presented.¹ It is accompanied by the discussion of an example. The approach is applied to the quarterly dividends and earnings per share of an equally-weighted portfolio of the twenty-five corporations listed in Table 1. The period covered begins with the first quarter of 1956 and ends in the fourth quarter of 1980.

In this approach to modeling multiple time series, the objective is to use statistics which are readily computed from the data and permit the choice of subclasses of models for consideration. The approach proceeds iteratively, in three stages;

- (1) identification, or preliminary model specification;
- (2) estimation of model parameters; and
- (3) diagnostic checking of the model's appropriateness.

A.1. Identification

After appropriate differencing and/or transformation, the series $z_{1,t}$ and $z_{2,t}$ are examined with the help of two types of statistics:

(1) sample cross-correlation matrices $\hat{\underline{\rho}}(\ell) = \{\hat{\rho}_{mn}(\ell)\}$,

$$\text{where } \{\hat{\rho}_{mn}(\ell)\} = \frac{\sum_t (z_{m,t} - \bar{z}_m)(z_{n,t+\ell} - \bar{z}_n)}{\left\{ \sum_t (z_{m,t} - \bar{z}_m)^2 \sum_t (z_{n,t} - \bar{z}_n)^2 \right\}^{1/2}} \quad (\text{A.1})$$

for $m, n = 1, 2, \dots, k$, k being the number of series, so that $k = 2$ in this case. \bar{z}_m and \bar{z}_n are the means of the component series and ℓ denotes the number of lags.

As a device for summarizing the structure of the cross-correlation matrices, a plus sign is used to indicate a value greater than $2T^{-1/2}$, a minus sign a value less than $-2T^{-1/2}$, and a dot to indicate a value between $2T^{-1/2}$ and $-2T^{-1/2}$, where T is the number of observations in each series. This is done under the assumption that, if the series were white noise, then, for large T , the coefficients $\hat{\rho}_{mn}(\ell)$ would be normally distributed with mean zero and variance T^{-1} .

For the example discussed in this appendix, Illustrations 1 and 2 present plots of the transformed and differenced series of quarterly dividends and earnings for the portfolio. More specifically, we have:

(1) in Illustration 1, the values of

$$(1 - B)(1 - B^4)D_t \equiv z_{D,t}$$

(2) in Illustration 2, the values of

$$(1 - B)(1 - B^4)\ln(P_t + 1.0) \equiv z_{P,t}$$

The differences and transformations indicated above were required in order to make the two series exhibit stationary behavior, and this may be observed in the illustrations.

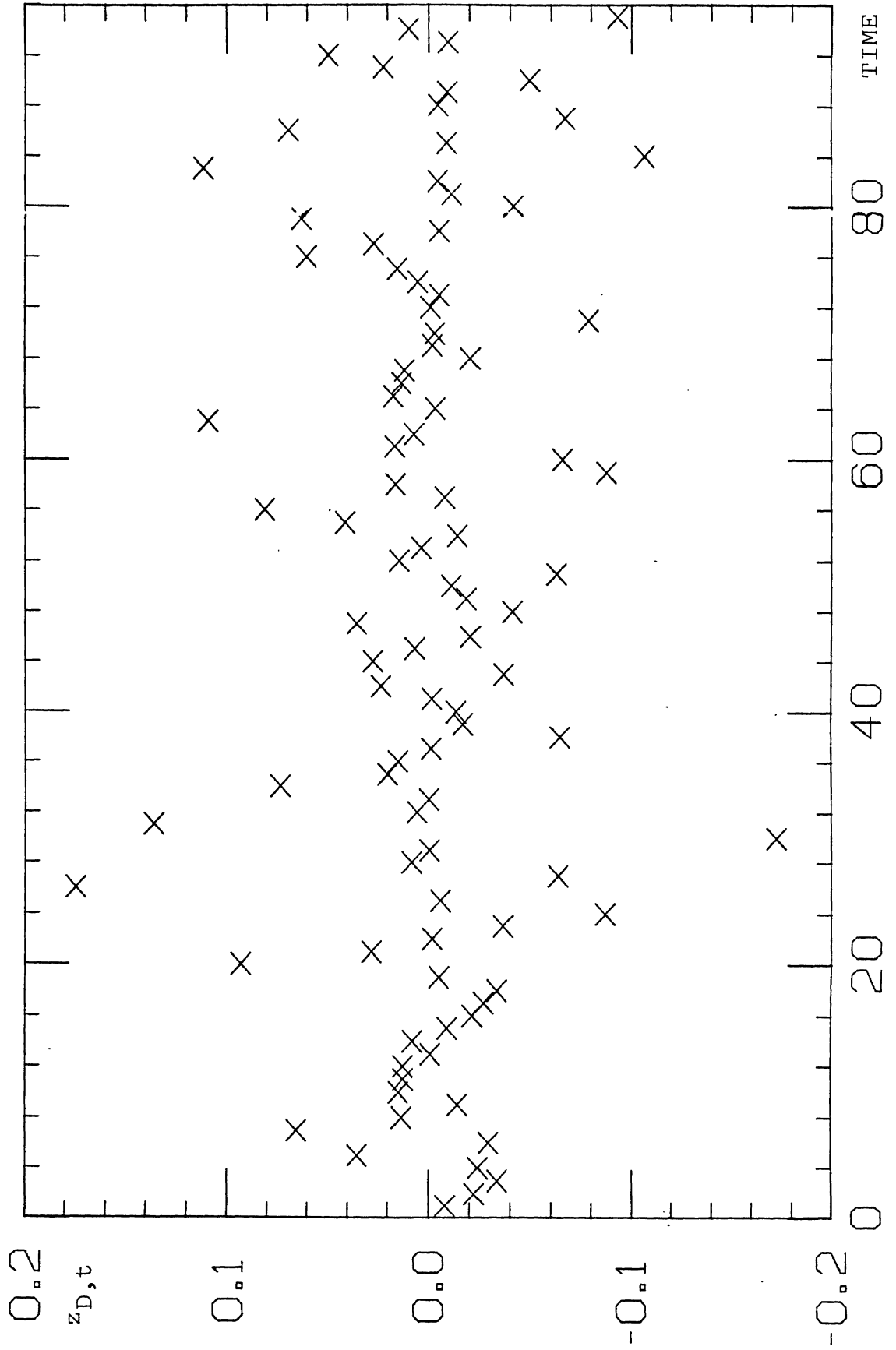


Illustration 1 TRANSFORMED SERIES OF D_t

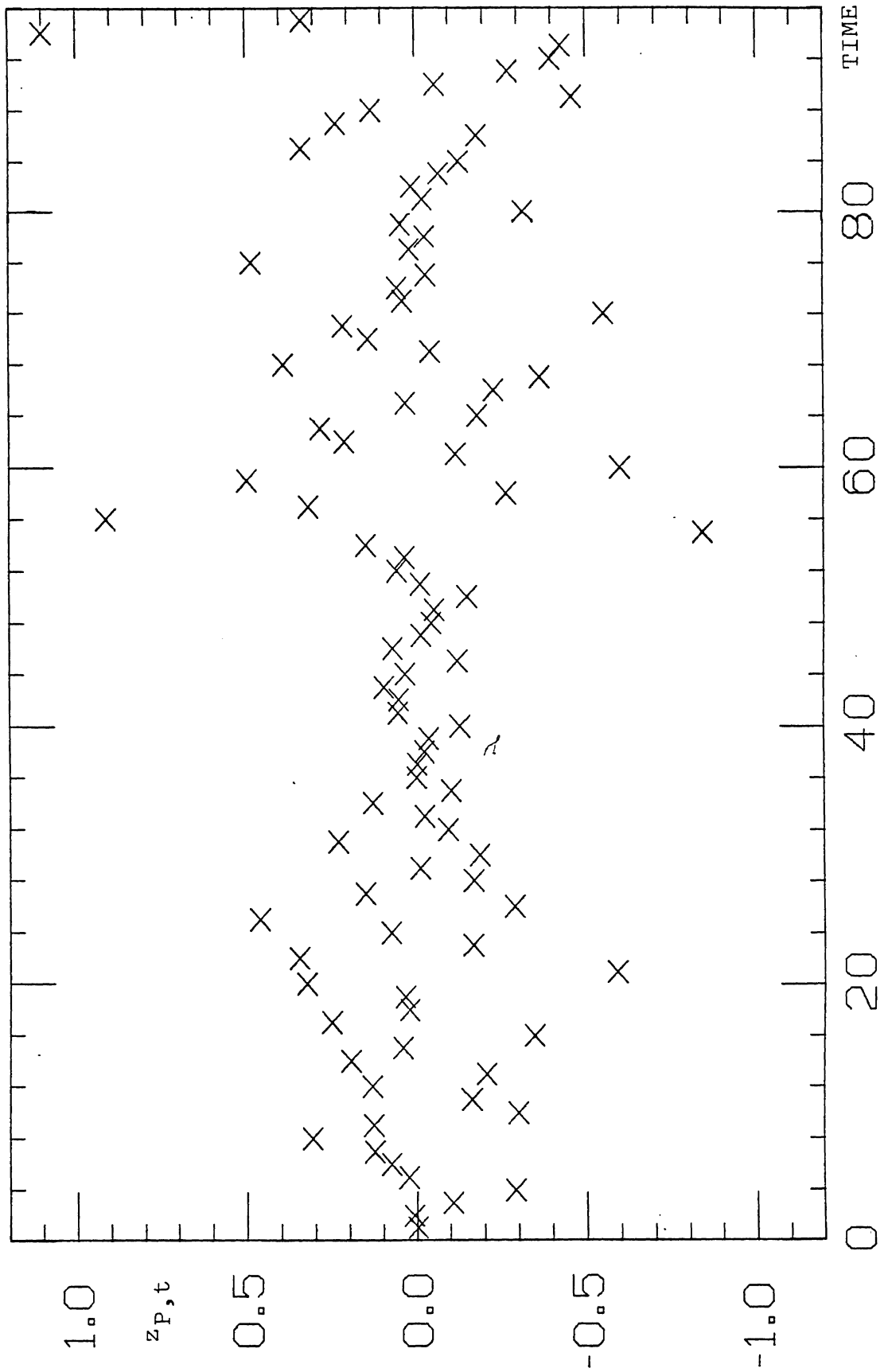


Illustration 2 TRANSFORMED SERIES OF P_t

In turn, Illustrations 3 through 5 present plots of the 24-lag sample auto-correlations and cross-correlations of $z_{D,t}$ and $z_{P,t}$. Table 5 contains the values of the correlations plotted in Illustrations 3 through 5. A summary of the correlation patterns using the indicator symbols is found in Table 6.

It is apparent, in the case of the earnings series, that the spikes at lags $\ell = 1$ and $\ell = 4$ suggests a moving-average model of the first order for both the regular and the seasonal components. In turn, a purely seasonal moving-average model, of order one, seems plausible for the dividends series. Note that no cross-correlation coefficient is indicated as being significantly different from zero, in confirmation of the results in section 5. Therefore, a vector $(0,1,1) \times (0,1,1)$ model, with seasonal period equal to four quarters, seems to be a good candidate.

(2) sample generalized partial correlations and related summary statistics:

Because the evidence in Illustrations 3-5 and Tables 5-6 seems to rule out a model with autoregressive components, use of generalized partial cross-correlation matrices becomes unnecessary, for this is an instrument suitable for the determination of the order of the autoregressive component of the model. However, for the sake of completeness, a brief description is given here.

The generalized partial cross-correlation matrix is defined, in a purely autoregressive model of order p , or $AR(p)$, as

$$P(\ell) = \begin{cases} \underline{\Gamma}^{-1}(0)\underline{\Gamma}(1) & \text{for } \ell = 1 \\ [\underline{\Gamma}(0) - \underline{b}'(\ell,0)\underline{A}^{-1}(\ell,0)\underline{b}(\ell,0)]^{-1}[\underline{\Gamma}(\ell) - \\ - \underline{b}'(\ell,0)\underline{A}^{-1}(\ell,0)\underline{c}(\ell,0)] & \text{for } \ell > 1 \end{cases}$$

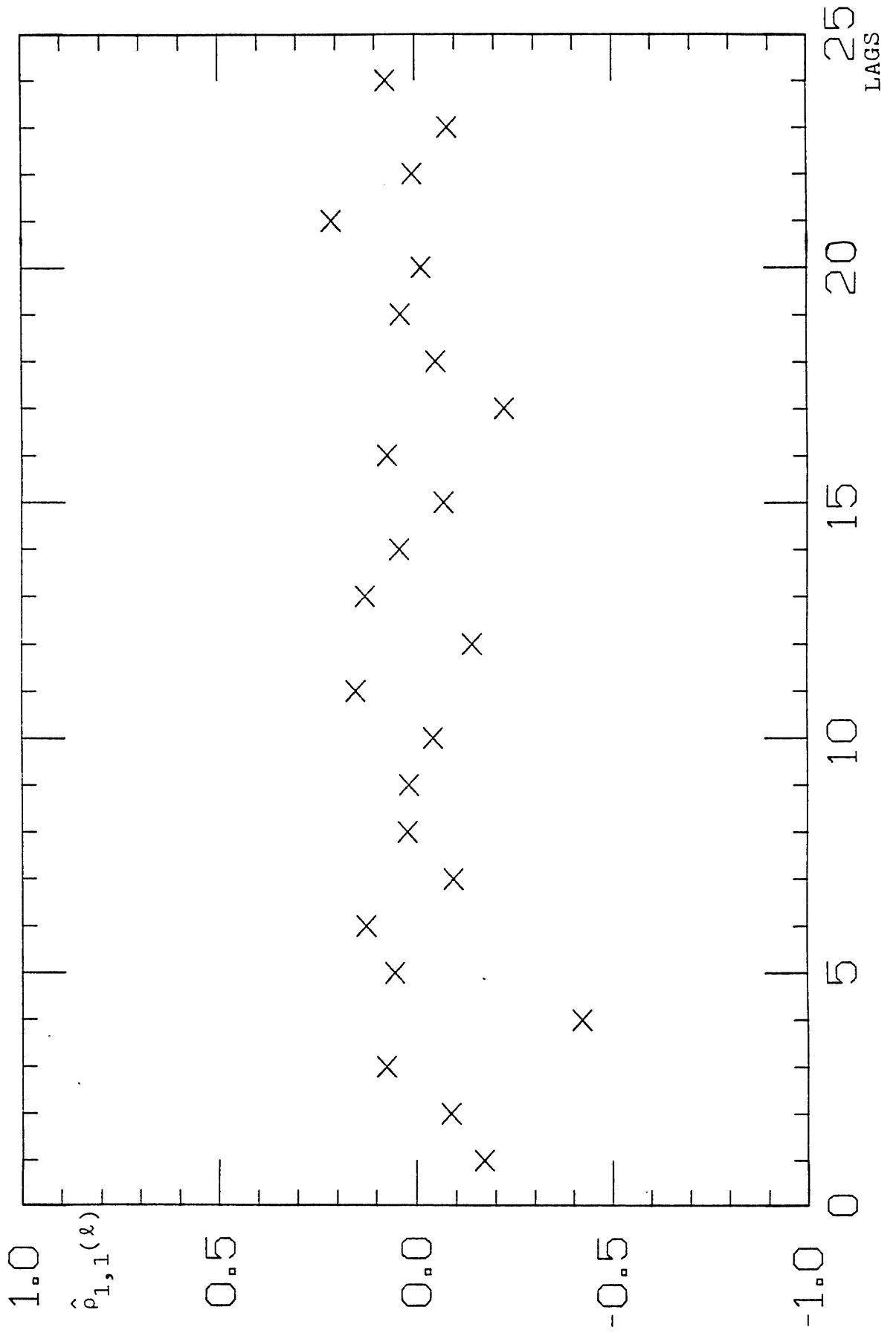


Illustration 3 AUTOCORRELATION FUNCTION FOR $z_{D,t}$

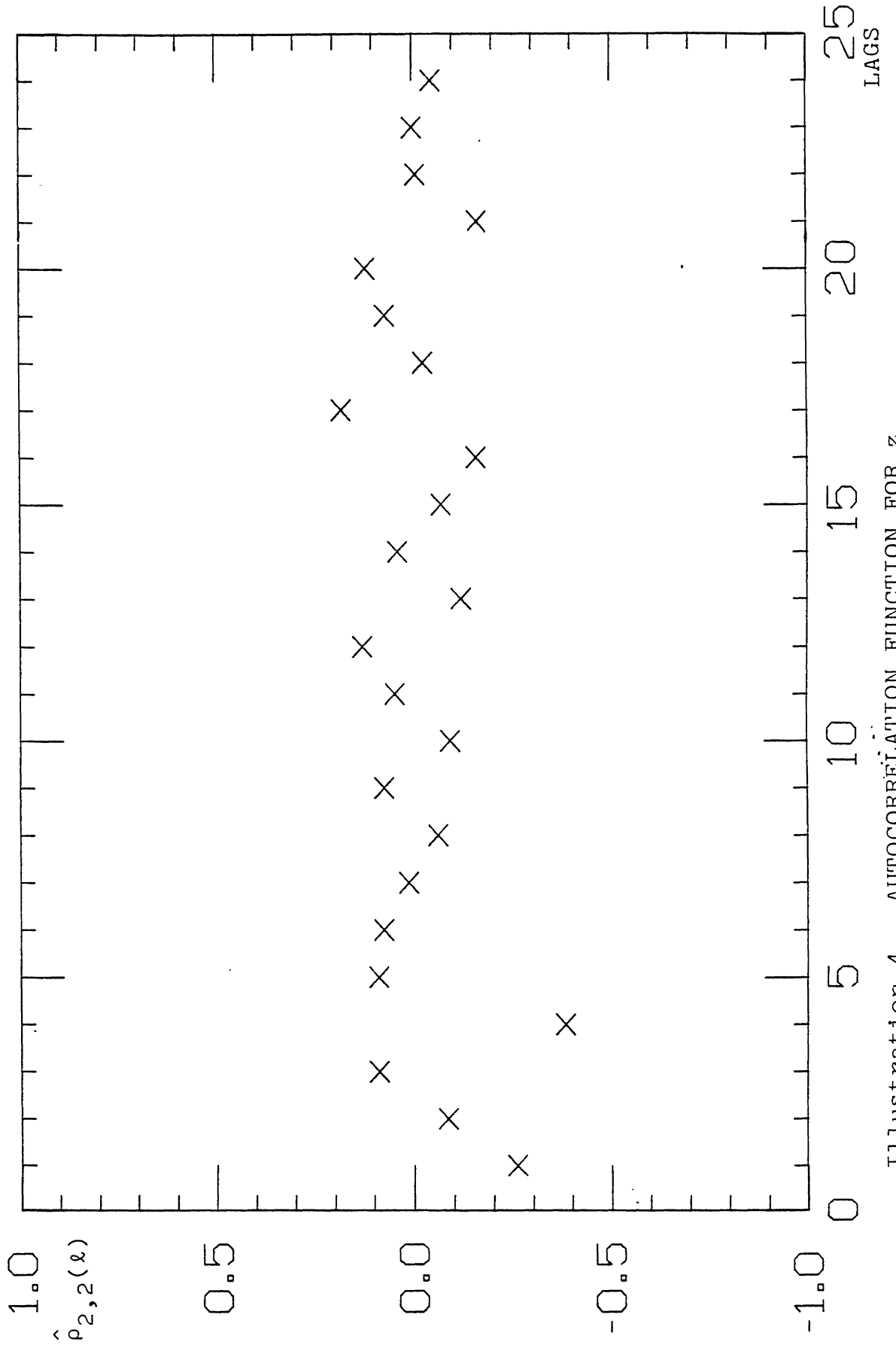


Illustration 4 AUTOCORRELATION FUNCTION FOR $z_{P,t}$

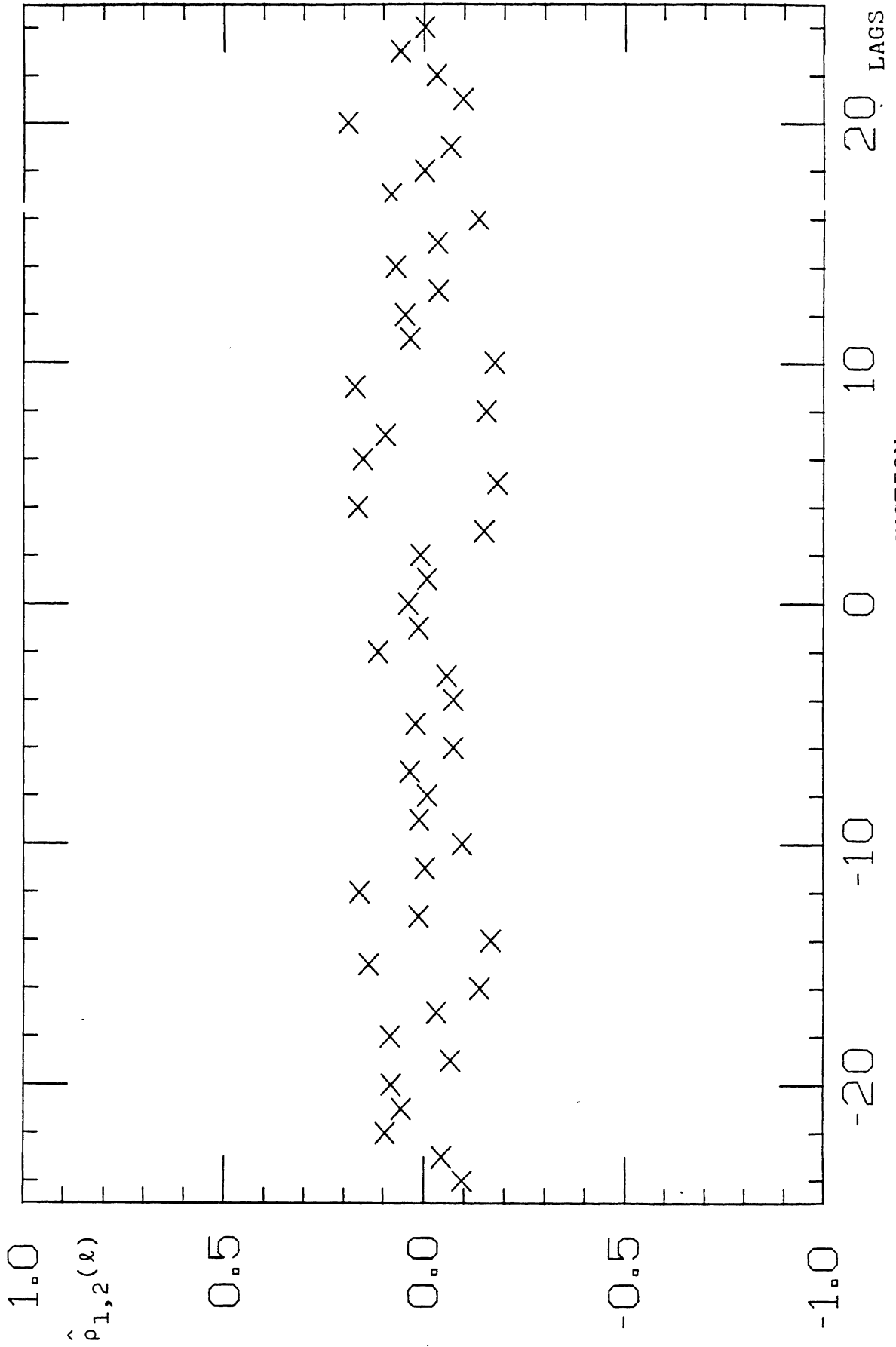


Illustration 5 CROSS-CORRELATION FUNCTION: $z_{D,t}$ vs. $z_{P,t}$

Table 5

SAMPLE AUTO- AND CROSS-CORRELATIONS OF $z_{D,t}$ AND $z_{P,t}$ (24 LAGS)

Sample correlation matrices
$$\begin{bmatrix} \hat{\rho}_{D,D}(\ell) & \hat{\rho}_{D,P}(\ell) \\ \hat{\rho}_{P,D}(-\ell) & \hat{\rho}_{P,P}(\ell) \end{bmatrix}$$

Lags 1-4:

-0.1718	0.0135	-0.0890	0.1149	0.0762	-0.0576	-0.4223	-0.0742
-0.0083	-0.2604	0.0084	-0.0870	-0.1493	0.0879	0.1672	-0.3840

Lags 5-8:

0.0544	0.0210	0.1268	-0.0740	-0.0947	0.0346	0.0212	-0.0078
-0.1823	0.0877	0.1543	0.0761	0.0975	0.0127	-0.1549	-0.0620

Lags 9-12:

0.0164	0.0120	-0.0445	-0.0950	0.1515	-0.0028	-0.1424	0.1623
0.1739	0.0729	-0.1761	-0.0942	0.0364	0.0464	0.0492	0.1284

Lags 13-16:

0.1282	0.0126	0.0400	-0.1667	-0.0723	0.1388	0.0719	-0.1393
-0.0349	-0.1215	0.0711	0.0396	-0.0339	-0.0719	-0.1362	-0.1587

Lags 17-20:

-0.2239	-0.0316	-0.0521	0.0836	0.0367	-0.0664	-0.0148	0.0828
0.0832	-0.1816	0.0001	-0.0270	-0.0675	0.0703	0.1911	0.1196

Lags 21-24:

0.2120	0.0571	0.0070	0.0983	-0.0835	-0.0450	0.0728	-0.0954
-0.0979	-0.1631	-0.0308	-0.0107	0.0585	0.0000	-0.0029	-0.0473

Table 6

SIGNIFICANCE INDICATORS FOR SAMPLE CROSS-CORRELATIONS IN TABLE 5

Summary of correlation matrices:

+ denotes value greater than $2(N)^{-1/2}$

- denotes value smaller than $-2(N)^{-1/2}$

. denotes value smaller than $2(N)^{-1/2}$ and greater than $-2(N)^{-1/2}$

$\hat{\rho}_{D,D}^{(\ell)}$	$\hat{\rho}_{D,P}^{(\ell)}$
...-.....
....-....+...
$\hat{\rho}_{P,D}^{(-\ell)}$	$\hat{\rho}_{P,P}^{(\ell)}$
.....	-...-.....
.....

Note: N denotes the effective number of observations after the transformations and/or differences applied to the original data. In this case, N = 95.

where the $\underline{\Gamma}(\ell)$ are the cross-covariance matrices, or

$$\underline{\Gamma}(\ell) = E(\underline{z}_{t-\ell} \underline{z}_t'), \text{ and}$$

$$\underline{A}(\ell, 0) = \begin{bmatrix} \underline{\Gamma}(0) & \dots & \dots & \underline{\Gamma}(\ell-2) \\ \cdot & \cdot & \cdot & \cdot \\ \underline{\Gamma}(\ell-2) & \dots & \dots & \underline{\Gamma}(0) \end{bmatrix}, \quad \underline{b}(\ell, 0) = \begin{bmatrix} \underline{\Gamma}'(-1) \\ \cdot \\ \underline{\Gamma}'(1) \end{bmatrix},$$

$$\text{and } \underline{c}(\ell, 0) = \begin{bmatrix} \underline{\Gamma}(1) \\ \cdot \\ \underline{\Gamma}(\ell-1) \end{bmatrix}.$$

from which it results that if \underline{z}_t follows an autoregressive process of order p , then

$$\underline{P}(\ell) = \begin{cases} \underline{\phi}_\ell & \text{for } \ell = p \\ \underline{0} & \text{for } \ell > p \end{cases}$$

where $\underline{\phi}_\ell$ and $\underline{0}$ are matrices of order k .

In this step of the identification stage, the sample values of $\underline{P}(\ell)$, or $\hat{\underline{P}}(\ell)$, are used for determining the possible value of p , the maximum order of the autoregressive component. In the Tiao-Box approach, this is done by fitting autoregressive models of order $\ell = 1, 2, \dots$ successively.

A vector AR(p) model can be written as

$$\underline{z}_t' = \underline{z}_{t-1}' \underline{\phi}_1' + \dots + \underline{z}_{t-p}' \underline{\phi}_p' + \underline{a}_t' \tag{A.2}$$

This can be expressed as a multivariate linear model:

$$\underline{Y} = \underline{X}_1 \underline{\phi}_1' + \dots + \underline{X}_p \underline{\phi}_p' + \underline{\varepsilon} \tag{A.3}$$

where

$$\underline{Y} = \begin{bmatrix} z'_{-p+1} \\ \cdot \\ \cdot \\ z'_T \end{bmatrix}, \underline{X}_1 = \begin{bmatrix} z'_{-p} \\ \cdot \\ \cdot \\ z'_{T-1} \end{bmatrix}, \dots, \underline{X}_p = \begin{bmatrix} z'_1 \\ \cdot \\ \cdot \\ z'_{T-p} \end{bmatrix}, \underline{\varepsilon} = \begin{bmatrix} a'_{-p+1} \\ \cdot \\ \cdot \\ a'_T \end{bmatrix}$$

and from which the least squares estimates $\hat{\phi}'_1, \dots, \hat{\phi}'_p$ can be obtained. When p is the largest lag of the sample partial cross-correlation matrices desired, an estimate of $\underline{P}(\ell)$, when an $AR(\ell)$ model is fitted, with $\ell = 1, 2, \dots, p$, is

$$\hat{\underline{P}}(\ell) = \hat{\underline{\phi}}_\ell$$

From Anderson [1971], it is known that, for a stationary $AR(\ell)$ model, the estimates $\hat{\underline{\phi}}_1, \dots, \hat{\underline{\phi}}_\ell$ asymptotically have the same distribution properties as in the classical multivariate linear model, in which the \underline{X} 's are assumed to be pre-determined.

Thus, we can compute the standard errors of elements of $\hat{\underline{P}}(\ell)$ and test their significance in a way similar to the t-tests of regression coefficients. This is based on 'standardized partial autoregression coefficients', defined as the ratios of the elements of $\hat{\underline{\phi}}_\ell$ to their respective standard errors.

In addition, a likelihood ratio statistic for the null hypothesis that $\underline{\phi}_\ell = \underline{0}$, against the alternative $\underline{\phi}_\ell \neq \underline{0}$, may be used. Let

$$\underline{S}(\ell) = (\underline{Y} - \underline{X}_1 \hat{\underline{\phi}}_1' - \dots - \underline{X}_\ell \hat{\underline{\phi}}_\ell')' (\underline{Y} - \underline{X}_1 \hat{\underline{\phi}}_1' - \dots - \underline{X}_\ell \hat{\underline{\phi}}_\ell')$$

be the matrix of residual sums of squares and cross-products after fitting an $AR(\ell)$ model. Then, with $\underline{S}(0) = \underline{Y}'\underline{Y}$, for $\ell = 1, 2, \dots, p$, the likelihood ratio statistic U is given by

$$U = \frac{|\underline{S}(\ell)|}{|\underline{S}(\ell-1)|}$$

and, using Bartlett's [1938] approximation, the statistic

$$M(\ell) = - (N - \frac{1}{2} - \ell k) \ln U$$

is, on the null hypothesis, distributed as χ^2 with k^2 degrees of freedom, and $N = T - p - 1$ is the effective number of observations, provided that a constant term is included in the model.

A.2 Estimation

Having tentatively selected the order of the model, efficient estimates of $\underline{\phi}$ and $\underline{\theta}$ are then computed by maximizing the likelihood function. From the autoregressive-moving average model, we can write

$$\underline{a}_t = \underline{z}_t = \phi_1 \underline{z}_{t-1} - \dots - \phi_p \underline{z}_{t-p} + \theta_1 \underline{a}_{t-1} + \dots + \theta_q \underline{a}_{t-q} \quad (\text{A.4})$$

In the Tiao-Box approach, both 'conditional' and 'exact' likelihood functions are considered. In the former case, with the series regarded as consisting of the $T - p$ vector observations $\underline{z}_{p+1}, \dots, \underline{z}_T$, the likelihood function is determined from $\underline{a}_{p+1}, \dots, \underline{a}_T$ using the preliminary values $\underline{z}_1, \dots, \underline{z}_p$ and setting $\underline{a}_p = \dots = \underline{a}_{p-q-1} = \underline{0}$, where q is the order of the moving average component.

With

$$\underline{S}(\underline{\phi}, \underline{\theta}) = \sum_{t=p+1}^T \underline{a}_t \underline{a}_t'$$

the conditional likelihood function $\ell_c(\underline{\phi}, \underline{\theta}, \underline{\Sigma} | \underline{z})$, is distributed as

$$|\underline{\Sigma}|^{-(T-p)/2} \exp\{-1/2 \text{tr} \underline{\Sigma}^{-1} \underline{S}(\underline{\phi}, \underline{\theta})\}$$

where $\underline{\Sigma}$ is the residual covariance matrix.

Hillmer and Tiao [1979] have shown that this approximation can be inadequate if T is 'not sufficiently large' and one or more zeros of $|\underline{\theta}_{=q}(B)|$ lie on or close to the unit circle. The result is producing biased estimates of the moving average parameters.

In the case of the exact likelihood function, whose form is developed in Hillmer and Tiao [1979] as

$\ell(\underline{\phi}, \underline{\theta}, \underline{\lambda} | \underline{z})$ which is distributed as $\ell_c(\underline{\phi}, \underline{\theta}, \underline{\lambda} | \underline{z}) \ell_1(\underline{\phi}, \underline{\theta}, \underline{\lambda} | \underline{z})$

and where ℓ_1 depends only on $\underline{z}_1, \dots, \underline{z}_p$, if $q = 0$, and on all the vectors, $\underline{z}_1, \dots, \underline{z}_T$, if $q \neq 0$, the function is approximated by the transformation

$$\underline{w}_t = (\underline{I} - \underline{\phi}_1 B - \dots - \underline{\phi}_p B^p) \underline{z}_t = \underline{\theta}_q(B) \underline{a}_t$$

In the example discussed in this section, a tentative model was estimated by the conditional likelihood function approach. In full, the tentative model estimated is

$$\underline{z}_t = (\underline{I} - \underline{\theta}_1 B)(\underline{I} - \underline{\theta}_4 B^4) \underline{a}_t$$

where $\underline{z}_t = (z_{D,t}, z_{P,t})'$. The model is represented by the multiplicative expression $(0,1,1) \times (0,1,1)^4$. The first triplet $(0,1,1)$ denotes an ARIMA(0,1,1) model for the regular part of the vector of series of D_t and P_t : no autoregressive component ($p = 0$), $d = 1$, indicating the use of a stationarity-inducing first difference in the original series, and a first order moving average process, represented by $q = 1$. An identical process is indicated for the seasonal part of the series in the second triplet. The exponent over the second triplet indicates that the seasonal period is equal to $s = 4$ quarters.³

A.3 Diagnostic Checking

The third stage of the iterative model building approach is based on an analysis of the residual series $\{\hat{a}_t\}$, given by

$$\hat{a}_t = \underline{z}_t - \hat{\underline{\phi}}_1 \underline{z}_{t-1} - \dots - \hat{\underline{\phi}}_p \underline{z}_{t-p} + \hat{\underline{\theta}}_1 \hat{a}_{t-1} + \dots + \hat{\underline{\theta}}_q \hat{a}_{t-q} \quad (\text{V.1.5})$$

The analysis is performed by examining the cross-correlation matrices of the residuals, plots of residuals of both series against time, and the X^2 portmanteau test proposed by Hosking [1980].⁴ The results for the

$(0,1,1) \times (0,1,1)^4$ model proposed for the equally-weighted portfolio of the twenty five firms in the sample are presented in Table, 7. Based on the value of the X^2 test statistic, indicating that the residuals of both series may be regarded as white noise, the estimated model is considered appropriate.

Table 7

RESULTS OF MODEL ESTIMATED FOR PORTFOLIO OF TWENTY FIVE FIRMS

	$\hat{\theta}_{=1}$		$\hat{\theta}_{=4}$	
	0.2899*	0.0080	0.7477*	0.0076
	(0.1018)	(0.0195)	(0.0801)	(0.0139)
	0.0501	0.4525*	0.0919	0.9972*
	(0.4777)	(0.1002)	(0.2494)	(0.0369)
Constant terms: for $z_{D,t}$			0.0006	$\chi^2 = 68.9$
			(0.0008)	
		for $z_{P,t}$	-0.0008	
			(0.0010)	

Note: Standard errors are in parentheses; an asterisk denotes coefficients greater than twice their standard errors.

Footnotes to Appendix

1. Among the several alternative approaches referred to in Tiao and Box [1981], those of Granger and Newbold [1977], Wallis [1977] and Chan and Wallis [1978] are also discussed by Tiao and Box. In this application of a multiple time series analysis approach, no attempt is made at evaluating alternative approaches or comparing their results, if and when applied to the data at hand.
2. Although not a formal test of significance, this procedure involving indicators of standard errors of autocorrelations and cross-correlations is an extension of what is proposed by Box and Jenkins [1976, p. 290] for diagnostic checking of univariate models, a procedure in turn based on a result of Anderson [1942].
3. For a sample of ninety four firms covering the period from the first quarter of 1958 to the fourth quarter of 1971, Griffin [1977] found that for quarterly earnings a satisfactory model was 'a multiplicative first-order moving average process in first differences of the four-period differences of quarterly earnings' (p. 81 of his article). The present results confirm his finding.
4. The test statistic is

$$Q' = T^2 \sum_{\ell=1}^r \frac{1}{2} (T - \ell)^{-1} \text{tr}(\hat{\underline{C}}_{\ell} \hat{\underline{C}}_{\ell}^{-1} \hat{\underline{C}}_{\ell} \hat{\underline{C}}_{\ell}^{-1})$$

where the summation involves the r lags for which covariance matrices are computed, $\hat{\underline{C}}_{\ell}$ denotes the covariance matrices of the model's residuals, with $\ell = 1, 2, \dots, r$. The statistic Q' is distributed as $X^2_{k^2(r-p-q)}$. In this example, $k = 2$, $T = 100$, and r is chosen to be equal to twenty four. For the estimated model, at a probability level of 0.05 and with 88 degrees of freedom, X^2 is equal to 110.7

References

- Anderson, R. L., "Distribution of the Serial Correlation Coefficient,"
Annals of Mathematical Statistics, Vol. 13 (1942), pp. 1-13.
- Anderson, T. W., The Statistical Analysis of Time Series, New York, N.Y.,
Wiley, 1971.
- Bartlett, M. S., "Further Aspects of the Theory of Multiple Regression,"
Proceedings of the Cambridge Philosophical Society, Vol. 34 (1938),
pp. 33-40.
- Box, G. E. P., and Jenkins, G. M., Time Series Analysis, Forecasting and
Control, San Francisco, Holden-Day, 1976.
- Chan, W. Y. T., and Wallis, K. F., "Multiple Time Series Modelling: Another
Look at the Mink-Muskkrat Interaction," Applied Statistics, Vol. 27 (1978),
pp. 168-175.
- Durand, D., "The Cost of Capital, Corporation Finance, and the Theory of
Investment: Comment," American Economic Review, Vol. 49 (September,
1959), pp. 639-655.
- Fama, E. F., and Babiak, H., "Dividend Policy: An Empirical Analysis,"
Journal of the American Statistical Association, Vol. 63 (December, 1968),
pp. 1132-1161.
- Granger, C. W. J., and Newbold, P., Forecasting Economic Time Series, New York,
N.Y., Academic Press, 1977.

- Griffin, P. A., "The Time-Series Behavior of Quarterly Earnings: Preliminary Evidence," Journal of Accounting Research (Spring, 1977), pp. 71-83.
- Handjinicolaou, G., and Kalay, A., "Wealth Redistribution or Changes in Firm Value: An Analysis of Return to Bondholders and Stockholders Around Dividend Announcements," Journal of Financial Economics, Vol. 13 (1984), pp. 35-63.
- Hillmer, S. C., and Tiao, G. C., "Likelihood Function of Stationary Multiple Autoregressive-Moving Average Models," Journal of the American Statistical Association, Vol. 74 (September, 1979), pp. 652-660.
- Hosking, J. R. M., "The Multivariate Portmanteau Statistic," Journal of the American Statistical Association, Vol. 75 (September, 1980), pp. 602-608.
- Johnston, J., Econometric Methods, New York, N.Y., McGraw-Hill, 1972.
- Laub, P. M., "Some Aspects of the Aggregation Problem in the Dividends-Earnings Relationship," Journal of the American Statistical Association, Vol. 68 (September, 1972), pp. 552-559.
- Lintner, J., "Distribution of Incomes of Corporations among Dividends, Retained Earnings and Taxes," American Economic Review, Vol. 46 (May, 1956), pp. 97-113.
- Modigliani, F., and Miller, M. H., "The Cost of Capital, Corporation Finance, and the Theory of Investment: A Reply," American Economic Review, Vol. 19 (September, 1959), pp. 655-669.
- Pindyck, R. S., and Rubinfeld, D. C., Econometric Models and Economic Forecasts, New York, N.Y., McGraw-Hill, 1981.

Sanvicente, A. Z., Corporate Bond Prices in an Option Pricing Framework,
Ph.D. dissertation, Stanford University, 1982.

Tiao, G. C., and Box, G. E. P., "Modeling Multiple Time Series with
Applications," Journal of the American Statistical Association, Vol. 76
(December, 1981), pp. 802-816.

Wallis, K. F., "Multiple Time Series Analysis and the Final Form of Econo-
metric Models," Econometrica, Vol. 45 (1977), pp. 1481-1497.

Watts, R., "The Information Content of Dividends," Journal of Business, Vol. 46
(April, 1973), pp. 191-211.