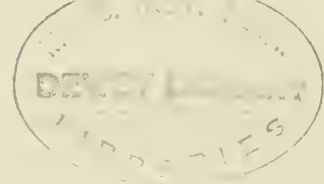




.M414

no. 1475-83

c. 2



WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

AGGREGATE DIVIDEND BEHAVIOR AND ITS
IMPLICATIONS FOR TESTS OF
STOCK MARKET RATIONALITY*

Terry A. Marsh
Robert C. Merton

WP #1475-83

September 1983

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139

AGGREGATE DIVIDEND BEHAVIOR AND ITS
IMPLICATIONS FOR TESTS OF
STOCK MARKET RATIONALITY*

Terry A. Marsh
Robert C. Merton

WP #1475-83

September 1983

*Dedicated to the scientific contributions and the
memory of John V. Lintner, Jr.

I. Introduction**

In a series of stimulating papers (1981a, 1981b, and 1982), Robert Shiller uses seemingly powerful variance bounds tests to show that variations in aggregate stock market prices are much too large to be justified by the variation in subsequent dividend payments. Under the assumption that the real expected return on the market remains essentially constant over time, Shiller concludes that the excess variation in stock prices identified in his tests provides strong evidence to reject the Efficient Market Hypothesis. Even if the real expected return on the market does change over time, Shiller further concludes that the amount of variation in that rate necessary to "save" the Efficient Market Hypothesis is so large that the measured excess variation in stock prices cannot be attributed to this source.

We need hardly mention the significance of such a conclusion. If Shiller's rejection of market efficiency is sustained, then serious doubt is cast on the validity of the most important cornerstone of modern financial economic theory. To be sure, of the hundreds of earlier tests of efficient markets, there have been a few which appear to reject market efficiency [cf. "Symposium on Some Anomalous Evidence on Capital Market Efficiency," Journal of Financial Economics (June-September 1978)]. For the most part, however, these studies are joint tests of both market efficiency and a particular equilibrium model of differential expected returns across stocks such as the Capital Asset Pricing Model, and therefore, rejection of the joint hypothesis may not imply a rejection of market efficiency. Even in their strongest interpretation, such studies have at most rejected market efficiency for select segments of the market. On the other hand, to reject the Efficient

Market Hypothesis for the whole stock market and at the level suggested by Shiller's analysis goes far beyond the narrow issue of whether or not some investors can beat the market. It implies broadly that production decisions based on stock prices will lead to inefficient capital allocations. More generally, if the application of rational expectations theory to the virtually "ideal" conditions provided by the stock market fails, then what confidence can economists have in its application to other areas of economics where there is not a large central market with continuously-quoted prices; where entry to its use is not free; and where short sales are not feasible transactions?

Rejection of market efficiency is essentially an empirical matter. Theory may suggest the correct null hypothesis--in this case, that stock market prices are efficient--but it cannot tell us whether or not real-world speculative prices as seen on Wall Street or LaSalle Street are indeed efficient. As Samuelson (1965) long ago noted in his seminal paper on efficient markets:

You never get something for nothing. From a nonempirical base of axioms, you never get empirical results. Deductive analysis cannot determine whether the empirical properties of the stochastic model I posit come close to resembling the empirical determinants of today's real-world markets. (p. 42)

Nevertheless, data cannot be analyzed without reference to a model, and in general, the interpretation of the results will not be invariant to the specification of the model. Shiller's analysis is no exception.

The Shiller model assumes that dividends are generated by an exogeneous, stationary process with a trend. If, as he further posits, the expected return on a share of stock is constant, then in an efficient market, the current price of a share can be represented as the present discounted value of expected future dividends. If the discount rate is constant, it follows that

in an efficient market, unanticipated price changes are simply reflections of revised expectations about future exogeneously-determined dividends. That is, the process for prices is endogeneous relative to the dividend process. If, moreover, dividends follow a stationary process with a trend, then the endogeneous process for price must also be a stationary process with a trend. These posited conditions are used to deduce certain relations between the time series of stock prices and aggregate dividends which must be satisfied if the market is efficient and if this model of the dividend process is correct. Shiller finds that the time series data violate these relations.

In this paper, we derive an alternative model for dividend and stock price behavior which is consistent with both market efficiency and the time series data. There are two fundamental differences between this model and the Shiller model: First, our model assumes that dividends follow a controlled, and therefore, essentially endogeneously-determined, process whereas the Shiller model assumes that the dividend process is exogeneous. Second, our model implies that asymptotically, both dividends and prices follow a geometric Brownian motion. Hence, in contrast to the Shiller model, neither dividends nor prices will have an asymptotic steady-state or stationary distribution. We will show that the empirical findings associated with the variance bound tests in the Shiller model are entirely consistent with the predictions of our model. This, together with the other theoretical and empirical evidence to be presented, leads us to conclude that the interpretation of Shiller's findings as implying that stock prices are "too volatile" is the result of a misspecification of the dividend process and not the result of market inefficiency.

In our development, the joint nature of the stock price and dividend process is emphasized because an empirical refutation of Shiller's claim of market inefficiency cannot be made on the basis of stock return data alone. Suppose, for example, that the stock return data support the empirical conclusion that stock prices follow a geometric random walk. While inconsistent with Shiller's model of stock prices in an efficient market, such a conclusion need not rule out market inefficiency. That is, if, as Shiller claims, stock prices are too volatile to be consistent with the stochastic process for dividends, and if this excess volatility can, for example, be attributed to irrational and random waves of optimism and pessimism among investors, then stock prices could appear to follow a random walk and yet stock prices would not reflect rational economic value. The random character of stock prices, which to some is stereotypical of market efficiency, can to others constitute prima facie evidence of market inefficiency.

In Section II of the paper, we present the theoretical foundation of our model of the stock price and dividend processes and develop its econometric specification. In Section III, we fit the model using stock market and aggregate dividend data for the period 1926-1981. In Section IV, we compare and contrast our findings with the regression analysis used to support the Shiller model of the dividend process. In the light of our alternative description of the dividend process, we reexamine the Shiller variance bound test methodology in Section V and show that his empirical results are entirely consistent with market efficiency.

II. Aggregate Dividend and Stock Price Model

In this section, we develop our model of the stock price and dividend processes. As we all know, the time series properties of aggregate stock market returns have been examined in numerous studies using everything from simple runs tests to sophisticated Kalman filtering techniques and spectral analysis. On the whole, these studies empirically support the view that stock returns are not forecastable which is, of course, consistent with the Efficient Market Hypothesis.¹ In contrast, relatively few studies have been devoted to the analysis of the time series behavior of aggregate dividends. As indicated in our introduction, Shiller's claim of market inefficiency depends critically on the validity of his model of the dividend process. Hence, the bulk of our analysis is focused on the dividend process. We tackle first the development of the stock price dynamics, because while essential, it is not likely to be controversial. Indeed, at the level of generality here, the stock price dynamics are almost definitional and are therefore, in no way inconsistent with the Shiller model.

As in the Shiller model, we too assume that the expected real rate of return on a share of the market portfolio is constant over time and denote that real rate by α . If $P(t)$ denotes the price of a share in the market portfolio at time t (again, in real terms), then by the well-known valuation equation, $P(t)$ can be written as:

$$P(t) = \epsilon_t^M \left\{ \int_t^{\infty} e^{-\alpha(s-t)} D(s) ds \right\} \quad (1)$$

where $D(s)$ denotes aggregate real dividends per share, and ϵ_t^M is the expectation operator over the probability distribution for future

dividends generated by the information set available to investors in the market as of time t .

From (1), we derive the dynamic equation for the evolution of stock price which can be separated into two components: the expected change and the unanticipated change. The expected change in the price between time t and $t + h$ can be written as:

$$\epsilon_t^M \{P(t+h) - P(t)\} = [k(h)P(t) - \bar{D}(t;h)]h \quad (2)$$

where $k(h) \equiv (e^{\alpha h} - 1)/h$ is the expected rate of return on the stock

over the interval h and $\bar{D}(t;h) \equiv e^{\alpha h} \epsilon_t^M \left\{ \int_t^{t+h} \exp[-\alpha(s-t)] D(s) ds \right\} / h$

is the expected end-of-period value of cumulative dividends paid during the period expressed as a flow per unit of time. From (1), the unanticipated component of the change in price can be written as:

$$P(t+h) - P(t) - \epsilon_t^M \{P(t+h) - P(t)\} \quad (3)$$

$$= \int_{t+h}^{\infty} \exp[-\alpha(s-t-h)] \{ \epsilon_{t+h}^M [D(s)] - \epsilon_t^M [D(s)] \} ds$$

which reflects the revision between t and $t+h$ of the market's expectation of future dividends. Note that from the perspective of any date equal to or earlier than t , the expected value of the right-hand side of (3) is always zero. Hence, it follows, by construction, that the sequence of unanticipated price changes described by (3) form a Martingale.

Any stock price process posited for (2) and (3) must satisfy the restriction that share owners enjoy limited liability (i.e., $P(t) \geq 0$). Moreover, to avoid arbitrage, $P(t) = 0$ if and only if $D(s) \equiv 0$ for all $s > t$. We can, therefore, rule out a priori the arithmetic random walk

model for stock price change. Because k is constant and $\bar{D}(t;h)$ is nonnegative, the unanticipated change in price in (4) cannot be described by a Gaussian distribution with a constant variance. Indeed, the variance of the unanticipated change in price must depend on the level of price in a systematic way. Therefore, any attempt to use standard regression analysis to fit the arithmetic change in price will not only produce inefficient estimates because of heteroscedasticity, but will also produce biased estimates of the coefficients for those explanatory variables which are correlated with the level of $P(t)$.

Given that prices must be positive,² it is more useful to express the price dynamics in terms of percentage change. From (2) and (3), we have that

$$\frac{P(t+h) - P(t)}{P(t)} = q(t;h)h + w(t;h) \quad (4)$$

where $q(t;h) \equiv [k(h) - \bar{D}(t;h)/P(t)]$ is the expected growth rate in price between t and $t+h$ and $w(t;h)$ is the unanticipated percentage change in price which is equal to the right-hand side of (3) divided by $P(t)$.

To locate (4) within the financial economics literature, we note that the standard continuous-time version of (4) is a stochastic differential equation of the $\hat{\text{Ito}}$ type which we write as:

$$\frac{dP}{P} = q(t)dt + \sigma dz \quad (5)$$

where $q(t) \equiv \alpha - D(t)/P(t)$; σ is the instantaneous standard deviation of the rate of return on the stock; and dz is a standard Wiener process.³ The prototype version of (5) assumes both α and σ are constants which implies that total returns on the stock (i.e., dividends plus price appreciation) follow a geometric Brownian motion.

Since in later econometric work, we will make the usual assumptions of i.i.d. Gaussian residuals, we choose to introduce those assumptions here and take advantage of the considerable analytical simplicity made possible by working with (5) rather than (4). That is, we assume that $w(t;h)$ in (4) can be expressed as:

$$w(t;h) = \frac{1}{P(t)} \int_t^{t+h} \sigma P(s) dz \quad (6)$$

where σ is assumed to be a constant.

While we assume (6) for analytical convenience, it does not create any material "model bias" with respect to the substantive issues of the paper. For example, the well-known Martingale property of a Brownian process does not imply that the stock market is efficient. As we have already noted, any posited process for stock price must satisfy the Martingale property for $w(t;h)$, and therefore, a nonanticipating process like a Brownian motion is almost a necessity for consistency. Although the assumption of (5) with a constant σ implies that cumulative total returns over a given interval are lognormally distributed with a variance rate which grows with the length of the interval, the stock price itself need not be lognormally distributed because of the "dividend-drag" term in (5). Indeed, depending upon the properties of the dividend process, (5) is entirely consistent with stock price having a steady-state or long-run stationary distribution around some fixed time trend. In short, nothing we have assumed about the stock price dynamics is inconsistent with the Shiller model of stock prices and dividends.

This completes our remarks concerning the stock price process. Before turning to the task of developing the dividend model, we pause here to briefly review some of the history of research on dividends. Although faithfully

discussed in every financial management textbook, corporate dividend policy has received relatively modest attention in the development of modern financial economic theory. A casual review of the empirical literature suggests that it has fared little better there. The number of empirical finance papers which use dividend series as their principal data file is small when compared with the number using either stock returns or earnings.

This seeming lack of research interest in dividends can undoubtedly be traced to the classical work of Miller and Modigliani (1961) which demonstrates the irrelevance of dividend policy for determining the firm's cost of capital. In brief, their proposition is as follows: The current value of the firm is equal to the expected future net cash flows generated by the firm's investments, discounted at the cost of capital. Management of the firm can choose virtually any time pattern for distribution of these flows to its stockholders (i.e., dividend policy) subject only to the overall constraint that the present value of expected future payouts cannot exceed the present value of the firm's net cash flows.⁴ Under the MM proposition, the expected return required by investors to induce them to hold the shares of the firm is invariant to the dividend policy selected by management, and hence, the cost of capital to be applied to the firm's investments is not affected by dividend policy. It therefore, follows that for a given investment policy, the current value of the firm will be the same for any choice of dividend policy. Thus, under the MM hypothesis, the current market value of the firm cannot be changed by a change in dividend policy, and in that sense, dividend policy "does not matter." If dividend policy does not matter, it is perhaps not surprising that financial researchers would turn their attention to other, more fruitful, topics for study.

Exceptions to the MM view are, of course, to be found in the literature. Gordon (1959; 1962) and Lintner (1962) claim that dividend policy does affect the firm's cost of capital and provide some evidence to support the view that a higher dividend payout reduces the cost of capital (i.e., investors prefer dividends). Others argue that personal and corporate taxes cause dividend policy to affect the firm's cost of capital, but in the direction that a higher payout raises the cost of capital (i.e., investors prefer capital gains). Black and Scholes (1974), Litzenberger and Ramaswamy (1979), Hess (1983), and Miller and Scholes (1982) represent just a part of an ongoing theoretical and empirical debate over this issue. Despite these exceptions, the empirical evidence to date is inconclusive for rejecting the Miller and Modigliani proposition.

As noted in their 1961 paper (p. 431), Miller and Modigliani did recognize the possibility that dividend changes are used to convey information not otherwise known to the market. Presuming a tax disadvantage from paying dividends, Ross (1977) and Bhattacharya (1979) use a signalling model approach to formalize this notion. In a study of 310 firms, Watts (1973) provides some empirical support for the informational content of dividends when he finds that on average, unexpected dividends and unexpected price changes are positively correlated around announcement dates. The evidence is, however, that the correlations are rather small.

In summary, some theories predict that dividend policy matters and others predict it does not, and the empirical studies provide little compelling evidence for one over the other. Moreover, except for certain bond indenture restrictions and accumulated earnings tax penalties, there do not appear to be any significant legal, accounting convention, or corporate tax factors to exert pressures on corporate managers to follow any particular dividend policy.

With normative theories having so little to say about the dividend process, it is perhaps not surprising that empirical researchers turned to positive theories of dividend policy to specify their models. The prototype for these models is the Lintner model (1956) based on stylized facts first established by him in a classic set of interviews of managers about their dividend policies. Briefly, these facts are: (i) Managers believe that their firms should have some long-term target payout ratio; (ii) In setting dividends, they focus on the change in existing payouts and not on the level; (iii) A major unanticipated and nontransitory change in earnings would be an important reason to change dividends; (iv) Most managers try to avoid making changes in dividends which stand a good chance of having to be reversed within the near future.

Lintner fit an econometric model based on these interviews that can be written as follows:

$$\Delta D_{jt} = a_j + c_j(\rho_j^* E_{jt} - D_{j,t-1}) + u_{jt} \quad (7)$$

where ΔD_{jt} is the change in firm j 's nominal dividends from period $t - 1$ to period t , E_{jt} is its nominal earnings in period t , and ρ_j^* is its target dividend payout rate. In his sample of companies over the period 1918-1941, Lintner found that (7) explained some 85% of annual dividend changes, and that the average payout ratio for all firms

$\bar{\rho}^*$ was about 50% with an average speed of adjustment \bar{c} equal to about 0.30. Using similar types of models, subsequent empirical work by Fama and Babiak (1968), Petit (1972), and Watts (1973) supports Lintner's original findings.

Significant by its absence from our brief review of the theory and empirical study of dividends is any reference to the analysis of aggregate

dividends for the whole stock market. The apparent lack of research interest in aggregate dividend behavior is perhaps not surprising since much of what is interesting about dividend policy is likely to be firm specific. For example, clientele effects and indenture restrictions which could in principle affect an individual firm's dividend policy are likely to "wash out" in any aggregate dividend analysis. Similarly, the informational content of dividends issue is likely to be considerably less important for the stock market as a whole than for an individual firm. It is, indeed, difficult to see how one could identify a meaningful announcement date for aggregate dividends to perform event studies along the lines of Watts (1973). There have, of course, been analyses of aggregate dividend time series,⁵ but none to our knowledge have used a Lintner-type behavioral model as we propose to do here.

As indicated by Lintner's stylized fact (i), managers apparently believe in target payout ratios, and so the specification of our model of aggregate dividend behavior reflects that fact. As most textbook discussions seem to agree, these target payout ratios are set with respect to some notion of long-run sustainable or "permanent" earnings rather than current earnings. That is, a change in current earnings which is viewed by management as essentially transitory would not be likely to give rise to a noticeable change in dividends. This view is also consistent with Lintner's stylized facts (iii) and (iv).

Unfortunately, except for the special case of a firm whose future earnings are certain and generated without further net new investment, the textbooks are not specific in defining permanent earnings. Our interpretation (which is consistent with this special case) defines the permanent earnings per share of a firm at time t as equal to the expectation as of time t of that level of

uniform payments which could be made by the firm to a single share in perpetuity. For an all-equity financed firm, permanent earnings are determined as follows: Let $\Pi(s)$ denote the real after-tax profit from the physical and financial assets of the firm at time s and $I(s)$ denote the real net new investment by the firm at time s . $I(s) =$ [gross new physical investment + purchases of financial assets - depreciation - sales of physical and financial assets]. If α denotes the firm's real cost of capital, then the discounted value of the expected cash flows available for distribution to each share outstanding at time t is given by:

$$V(t) \equiv \varepsilon_t \left\{ \int_t^{\infty} [\Pi(s) - I(s)] e^{-\alpha(s-t)} ds \right\} / N(t) \quad (8)$$

where ε_t denotes the expectation operator, conditional on information available as of time t and $N(t)$ denotes the number of shares outstanding. $V(t)$ is sometimes called the "intrinsic value" (per share) of the firm, and permanent earnings per share are determined by creating a perpetual annuity from this intrinsic value. That is, if $E(t)$ denotes permanent earnings per share of the firm at time t , then:⁶

$$E(t) = \alpha V(t) \quad (9)$$

In our behavioral equation for the dividend process, it is corporate managers' assessments of permanent earnings which are relevant for the evolution of aggregate dividends. For this purpose, we denote managers'

determination of permanent earnings by $E^m(t) = \alpha V^m(t)$ where $V^m(t)$ is given by (8) with the expectation operator $\epsilon_t = \epsilon_t^m$ based on the probability distribution for future $\Pi(s)$ and $I(s)$ generated by the managers' information sets as of time t .

If Δ_t denotes the forward difference operator between $t - 1$ and t so that $\Delta_{t+1} D \equiv D(t + 1) - D(t)$ where $D(t)$ equals aggregate dividends per share of the market portfolio at time t , then our equation for the evolution of aggregate dividends is given by:

$$\frac{\Delta_{t+1} D}{D(t)} = g(t) + \lambda \left[\frac{\Delta_t E^m}{E^m(t - 1)} - \epsilon_{t-1}^m \left(\frac{\Delta_t E^m}{E^m(t - 1)} \right) \right] + \gamma [b - \log(D(t)/E^m(t))] + u(t + 1) \quad (10)$$

The first term, $g(t) \equiv \alpha r(t)$, is the usual expression for the growth rate of dividends where $r(t) \equiv 1 - D(t)/E^m(t)$ is the retention rate (in terms of permanent earnings) at time t . It reflects the standard textbook proposition that if the current payout is high relative to permanent earnings and therefore the retention rate $r(t)$ is low, then dividends per share will be expected to grow more slowly than if the current payout were lower and the retention rate were correspondingly higher. The rest of the terms on the right-hand side of (10) describe the deviation of the growth in dividends from this normal rate.

The specification of the behavioral equation in terms of the change in dividends rather than the levels is motivated by Lintner's stylized fact (ii). We express the change in percentage terms because it is likely that the magnitude of the change in dividends is not independent of the level. Especially because $D(t)$ cannot be negative, we believe that few people would consider the chance of a \$0.20 change in the dividend per share to be just as

likely when $D(t) = \$1.00$ as when $D(t) = \$10.00$, and therefore, to express the change in absolute terms is to create problems of heteroscedasticity in the error terms.⁷

The second term in (10) which is multiplied by λ is our representation of Lintner's stylized fact (iii) that managers change dividends from their normal growth rate in response to an unanticipated change in permanent earnings. Note that, as specified, the reaction to a change in permanent earnings is lagged one period. That is, an unanticipated change in permanent earnings between $t - 1$ and t causes a change in dividends between t and $t + 1$. We choose this lag specification for a number of reasons: first, an unanticipated change in permanent earnings, by definition, cannot be known until it happens, so any reaction in dividends to such a change must occur at the same time or later. Unlike delays in the reaction of speculative prices to new information, there are no arbitrage opportunities created by managers if they delay changing dividends. Second, from Lintner's stylized fact (iv), it is plausible that managers would choose to delay changes in the dividend to avoid "quick" reversals. Third, a firm changes its dividend at most once a quarter and many firms make significant changes only at the end of their fiscal year. Fourth, even if individual firms' managers react instantaneously, the reaction in aggregate dividends will appear to be lagged because of different announcement dates and different speeds of reaction across firms. From stylized fact (iv), managers are likely to make only a partial rather than a full adjustment to an unanticipated change in permanent earnings in which case λ should be somewhere between zero and one.

The short-run extrapolative characteristics of the posited reaction to unanticipated permanent earnings changes do not ensure that dividends will

approach any particular target payout ratio. Therefore, to reflect Lintner's stylized fact (i), we include in the third term of (10) a long-run regressive component which causes the level of dividends to asymptotically tend toward a target payout ratio. In (10), b is related to the target or long-run expected value of $\log[D/E^m]$. If the dividend payout ratio is to converge to its long run target, then the speed of convergence to this target, γ , should be positive.

In the empirical studies of both Lintner's model and subsequent dividend models based on his original formulation, the equations corresponding to our (10) are treated as regression equations. We too assume that equation (10) is both a structural equation and a causal equation because our view of the economic process is that an unanticipated change in permanent earnings causes a predictable change in next period's dividends, and not the reverse.

As noted at the outset of the paper, the direction of causality between permanent earnings and dividends is the key to the proper interpretation of the empirical evidence on relative stock market volatility. Given the significance of this point to the substantive conclusions of the paper, we pause in our development of the structural equations of the model to elaborate further on the economic foundation for the causality assumption in equation (10).

As already noted, there are no important legal or accounting constraints on dividend policy, and hence, managers have almost complete discretion and control over the choice of dividend policy. If, however, managers are rational, then they will, at least, choose a dividend policy which is feasible in both the short and long runs. Such "rational" policies must satisfy the constraint that the discounted value of expected future dividends per share is

equal to the discounted value of expected future net cash flows as given by (8). Because managers set dividend policy, this constraint is properly specified in terms of their probability assessments. Hence, from (9), it follows that a rational dividend policy must satisfy:⁸

$$\epsilon_t^m \left\{ \int_t^\infty D(s) e^{-\alpha(s-t)} ds \right\} = E^m(t)/\alpha \quad (10')$$

This constraint on rational dividend choice is very much like the intertemporal budget constraint on rational consumption choice in the basic lifetime consumption decision problem for an individual. In this analogy, the capitalized permanent earnings of the firm, $E^m(t)/\alpha$, correspond to capitalized permanent income or wealth of the individual, and the dividend policy of the firm corresponds to the consumption policy of the individual. Just as there are an uncountable number of rational consumption plans which satisfy the consumer's budget constraint for a given amount of wealth, so there are an uncountable number of distinct dividend policies which satisfy (10') for a given value of permanent earnings. Hence, like rational consumers in selecting their plans, rational managers have a great deal of latitude in their choice of dividend policy.⁹ The fact that individual firms pursue dividend policies which are vastly different from one another is empirical evidence consistent with this view.

Pursuing the analogy further: if because of an exogeneous event (for example, a change in preferences), a consumer changes his planned pattern of consumption, then it surely does not follow from the budget constraint that this change in the future time path of his consumption will cause his current wealth to change. Just so, it does not follow from (10') that a change in

dividend policy by managers will cause a change in their current assessment of permanent earnings.¹⁰ For a fixed discount rate, α , it does however follow from (10') that an unanticipated change in permanent earnings must necessarily cause a change in expected future dividends, and this is so, for the same feasibility reason, that with a constant discount rate, an unanticipated change in a consumer's wealth must necessarily cause a change in his planned future consumption.

The direction of causation between unanticipated changes in permanent earnings and changes in subsequent dividends posited in our model is, thus, consistent with the direction of causation between changes in current wealth and changes in future consumption that is normally assumed for the life cycle model in a fixed-discount rate world.¹¹

If, as appears to be the case in the Shiller model, dividends are assumed to follow an exogeneous process, then the direction of causation between permanent earnings and future dividends is exactly the reverse. Under this exogeneity assumption, (10') can no longer be a constraint on the time path of future dividends and must instead become a constraint on current permanent earnings. As managers receive information about this exogeneous dividend process, they, in effect, must revise their assessment of permanent earnings according to (10') because they know that the dividend process is feasible. In such a formulation, unanticipated changes in permanent earnings are merely a reflection of managers' revised expectations about future exogeneously-determined dividends. If applied in an analogous fashion to the lifecycle model, the corresponding assumption would be that the individual's time path of consumption follows an exogeneous process which, from the budget constraint, would imply that an unanticipated change in his current wealth is

just a reflection of the consumer's revised expectations about his future consumption.

In making the case for causality in equation (10), we are not unaware of the possibility that there are other exogeneous variables which may cause managers to change dividends. If this is the case and if, further, these variables are correlated with unanticipated changes in current permanent earnings, then, of course, equation (10) is flawed as a causal equation. If, however, managers are rational predictors of permanent earnings, then an unanticipated change in permanent earnings this period will be uncorrelated with all variables which are observable prior to this period (including both past dividends and permanent earnings). It therefore must also be uncorrelated with all future unanticipated changes in permanent earnings. Thus, if there are other exogeneous variables which explain next period's change in dividends, it seems unlikely that they would be correlated with this period's unanticipated change in permanent earnings. Hence, the assumption that equation (10) is a proper regression equation is likely to be robust with respect to other "missing" explanatory variables.

This property of rationally-predicted permanent earnings together with the lagged structure of equation (10) may perhaps at first suggest that the causality issue can be resolved empirically by applying an appropriate version of the Granger-Sims test of causality. A careful review of this possibility will, however, lead to the well-known conclusion of the identification problem that statistical tests alone are not sufficient to establish causality and that ultimately, this issue can only be resolved by a priori economic reasoning (cf. Zellner (1979)).

In summary, our case for causality in equation (10) is as follows: First, we can identify the economic mechanism (namely, the managers) by which changes in unanticipated permanent earnings can cause subsequent changes in dividends. Second, there is substantial empirical evidence (independent of the regression results reported here) that managers do exercise their control in an attempt to smooth the time path of dividends and that managers view unexpected changes in permanent earnings as a significant reason to change dividends. Third, certainly by comparison with dividend policy, managers have little, if any, control over unanticipated changes in permanent earnings.

In specifying (10), our intent is to construct a simple model of the dividend process which nevertheless captures the basic stylized facts of management behavior. We have therefore assumed a simple one-period lagged adjustment. It is possible that the dividend process may involve higher order lags with different speeds of adjustment, and as already noted, there may be other "missing" variables which enter into the process. As we will show, the fundamental empirical results derived from this simple model are likely to be robust with respect to refinements which include such additional variables.

Equations (4), (6), and (10) describe the structural equations of our model for the stock price and dividend dynamics.

If managers behave rationally and the market is reasonably efficient, then the market's estimate of a firm's intrinsic value should on average be equal to the intrinsic value estimate made by that firm's management. We, therefore, assume that the discounted value of the expected future aggregate net cash flows of all firms as estimated from the market's information set is equal to the aggregate sum of the intrinsic values where the intrinsic value of each firm is estimated from the information set of that firm's

management.¹² To complete the structural specification of our model, we write this market efficiency condition as:

$$V^M(t) = V^m(t) \quad \text{for all } t \quad (11)$$

where $V^M(t)$ is given by (8) with $\epsilon_t = \epsilon_t^M$.

Under the hypothesis that our model is correct, it is straightforward to show from (10) that the dividend-to-stock price ratio has a long-run or asymptotic stationary distribution. The Shiller model also predicts this result. The Shiller model, however, further predicts that the level of the stock price and the level of aggregate dividends have an asymptotic stationary distribution around a deterministic exponential trend. In sharp contrast, our model predicts that both the change in price and the change in dividends asymptotically follow a geometric Brownian motion, and therefore, neither the level of stock price nor the level of dividends has a steady-state distribution (with or without a trend). This difference in prediction, together with the causality issue surrounding the exogeneity of the dividend process, are the important differences between the Shiller model and our model for the purpose of evaluating the evidence on whether or not the stock market is too volatile, and hence, whether or not the market is inefficient. Because the source of these important differences between the two models flows directly from their different posited dynamic behavior for dividends, we turn now to the empirical analysis of this aspect of the two models.

III. Empirical Estimation of the Model

In this section, we estimate our model of dividend behavior, and in Section IV, we compare its performance with the alternative specification of the Shiller model.

To estimate our structural equation (10) requires a time series of aggregate managerial assessments of permanent earnings which, as we all know, is not available. The structural equations of our model can, however, be combined to derive a reduced form equation which can be used to estimate the coefficients in (10).

From the accounting identity,¹³ We can rewrite equation (1) for the price as:

$$P(t) = \varepsilon_t^M \left\{ \int_t^{\infty} e^{-\alpha(s-t)} [\Pi(s) - I(s)] ds \right\} / N(t) \quad (12) .$$

Using the market efficiency condition (11) and the definition of permanent earnings in (9), we can rewrite (12) as:

$$P(t) = E^m(t) / \alpha \quad . \quad (13)$$

Because the expected return on the market, α , is assumed to be constant, it follows from (13) that the percentage change in stock market price is equal to the percentage change in managers' assessment of permanent earnings.

Substituting for $E^m(t)$ from (13), we rewrite (10) as:

$$\begin{aligned} \frac{\Delta_{t+1} D}{D(t)} = & \left[\alpha - \frac{D(t)}{P(t)} \right] + \lambda \left[\frac{\Delta_t P}{P(t-1)} - \left(\alpha - \frac{D(t-1)}{P(t-1)} \right) \right] \\ & + \gamma \left[b' - \log \frac{D(t)}{P(t)} \right] + u(t+1) \end{aligned} \quad (14)$$

where $b' \equiv b + \log \alpha$. By rearranging terms, we can rewrite (14) as

$$\frac{\Delta D_{t+1}}{D(t)} + \frac{D(t)}{P(t)} = a_0 + a_1 \left[\frac{\Delta_t P + D(t-1)}{P(t-1)} \right] + a_2 \log \frac{D(t)}{P(t)} + u(t+1) \quad (14')$$

where $a_1 = \lambda$; $a_2 = -\gamma$; and $a_0 = (1 - \lambda)\alpha + \gamma b'$.

Although (12) follows from the accounting identity, (13) is not an identity. It is a specification which is valid under the hypothesis that market prices are rational predictors of firms' future net cash flows. If market price is a very noisy estimator of managements' assessments of permanent earnings, then (14) should be a poor predictor of the evolution of dividends. This would be the case if on the one hand, managers are rational predictors of the future cash flows but on the other, the market price is subject to wide swings caused by waves of irrational investor optimism and pessimism. Alternatively, (14) could be a poor predictor because management dividend decisions in the aggregate are not well described by behavioral equation (10). Such is the nature of a reduced form equation.

If, however, market price provides a good estimate of managements' assessments of permanent earnings, then (14) should be a good predictor of the dividend process. Such would be the case if managers are rational predictors of future cash flows and the market is efficient. It is, of course, possible for the market to be inefficient and for (14) to be a good predictor of the dividend process. This would be the case if the market is inefficient but managers rely on market price as their estimate of permanent earnings. Alternatively, managements' assessments of permanent earnings and the market price could be irrational in the same way because managers' expectations are influenced by the same elements which cause investors to be irrationally optimistic or pessimistic. If, however, this were the case, then using

dividend data to test market efficiency is doomed because both the price and dividend processes would be irrational. In short, whether in our model or the Shiller model, the most that can be determined from such time series comparisons is some measure of "relative rationality."

For (14') to be a proper reduced form equation, its right-hand side variables must be predetermined relative to its left-hand side variable. As discussed at length in the development of the dividend behavioral equation (10) in Section II, an unanticipated change in permanent earnings this period (over which managers have no control) is an exogeneous variable relative to the change in next period's dividends which managers control almost completely. From (12) and structural equation (13), it therefore follows that an unanticipated change in this period's price is exogeneous relative to next period's dividend change, and hence, (14') is a proper reduced form equation. In this limited sense of a reduced form, an unanticipated change in this period's price "causes" a (predictable) change in next period's dividends.

To estimate the reduced form equation (14'), we use annual data¹⁴ constructed from monthly dividend and price series for the value-weighted NYSE index contained in the Center for Research in Security Prices data set over the period 1926 to 1981. To the extent possible with composite data, we have constructed series of net dividends and price changes which are scaled on a unit basis to adjust for growth in the index by new issues of stock over time. We note that our data set is different from those used by Shiller. However, using our data set, we were able to obtain essentially the same empirical findings as reported by Shiller for his data sets. We, therefore, expect that the results reported here for our model will also obtain if it were fit to his data sets.

Because annual data represents a fairly long discrete-time interval between observations, it is statistically more appropriate to represent the percentage changes in the variables in (14') by the change in the logarithm of these variables. The only effect of this substitution is to change slightly the interpretation of the constant term, a_0 , in (14').¹⁵ For annual data over the period 1927 to 1980, this discrete time version of (14') estimated by ordinary least squares (OLS) is given by

$$\log\left[\frac{D(t+1)}{D(t)}\right] + \frac{D(t)}{P(t)} = -0.101 + 0.437 \log\left[\frac{P(t) + D(t-1)}{P(t-1)}\right] + \frac{-0.042}{(0.050)} \log\left[\frac{D(t)}{P(t)}\right] + u(t+1) \quad (15)$$

$$\bar{R}^2 = 0.47 \quad DW = 1.53$$

In (15), $D(t)$ refers to aggregate NYSE dividends totalled over year t , and $P(t)$ refers to price at the end of year t . The numbers in brackets under the coefficients are standard errors, and not t -statistics. For example, the coefficient of the lagged logarithmic change in price has a standard error of 0.064, and a t -statistic of 6.83. The coefficient point estimates in (15) indicate that the deviations of real dividend changes from their normal growth rate covary positively and strongly with the previous year's unexpected price changes and negatively with the previous year's dividend yield. The Durbin Watson statistic suggests that there is positive autocorrelation in the residuals of the OLS fit of (15). Disturbance correlation can arise in various ways. As already noted, our simple model assumes one-period adjustment in dividends by management, whereas longer lags are entirely possible. Another source could be omitted variables which are serially correlated.¹⁶

In light of the autocorrelation in the residuals of (15) we reestimated (14') using generalized least squares, and the results are:

$$\log\left[\frac{D(t+1)}{D(t)}\right] + \frac{D(t)}{P(t)} = -0.234 + 0.444 \log\left[\frac{P(t) + D(t-1)}{P(t-1)}\right] \quad (16)$$

(0.198) (0.061)

$$-0.085 \log\left[\frac{D(t)}{P(t)}\right] + u'(t+1)$$

(0.082)

$$\bar{R}^2 = 0.53 \quad DW = 1.83$$

Although the GLS estimate which takes into account the positive autocorrelation appears to have slightly more explanatory power, the results from either the OLS or GLS fits are essentially the same in that they explain about 50 percent of aggregate NYSE real dividend changes. As will be shown in Section IV, the explanatory power of our single-equation aggregate time series model¹⁷ is considerably higher than that provided by the corresponding regression fit used to motivate the Shiller model.

The point estimate of 0.44 for the coefficient on the lagged percentage price change is positive, substantial in magnitude, and highly significant. This finding is consistent with the hypothesis that the market price is a good indicator of permanent earnings and that managers systematically change dividends in response to an unanticipated change in permanent earnings. Because the coefficient is both significantly greater than zero and significantly less than one, this finding is also consistent with the Lintner stylized fact that managers smooth dividends by responding in a partial adjustment fashion to an unanticipated change in permanent earnings. The well-established empirical fact that the variation in the percentage change in dividends is significantly less than the variation in the percentage change in

prices, might suggest to some that prices are "too volatile." However, the empirical verification in (15) and (16) of the partial adjustment mechanism posited in our model provides an explanation of this well-established fact that is entirely consistent with market price being a rational predictor of future dividends.

The estimated coefficient of the dividend-to-price ratio is negative in both (15) and (16) which is consistent with the hypothesis that this ratio converges to a long-run stationary distribution. The point estimates for the speed of adjustment are rather small which suggests that a substantial period of time is required for the dividend-to-price ratio to converge to its steady-state distribution. Thus, using the -0.085 estimate from (16), a conventional "half-life" calculation shows that it takes more than eight years for the expected value of this ratio to move halfway from its initial value to its expected steady-state value.¹⁸

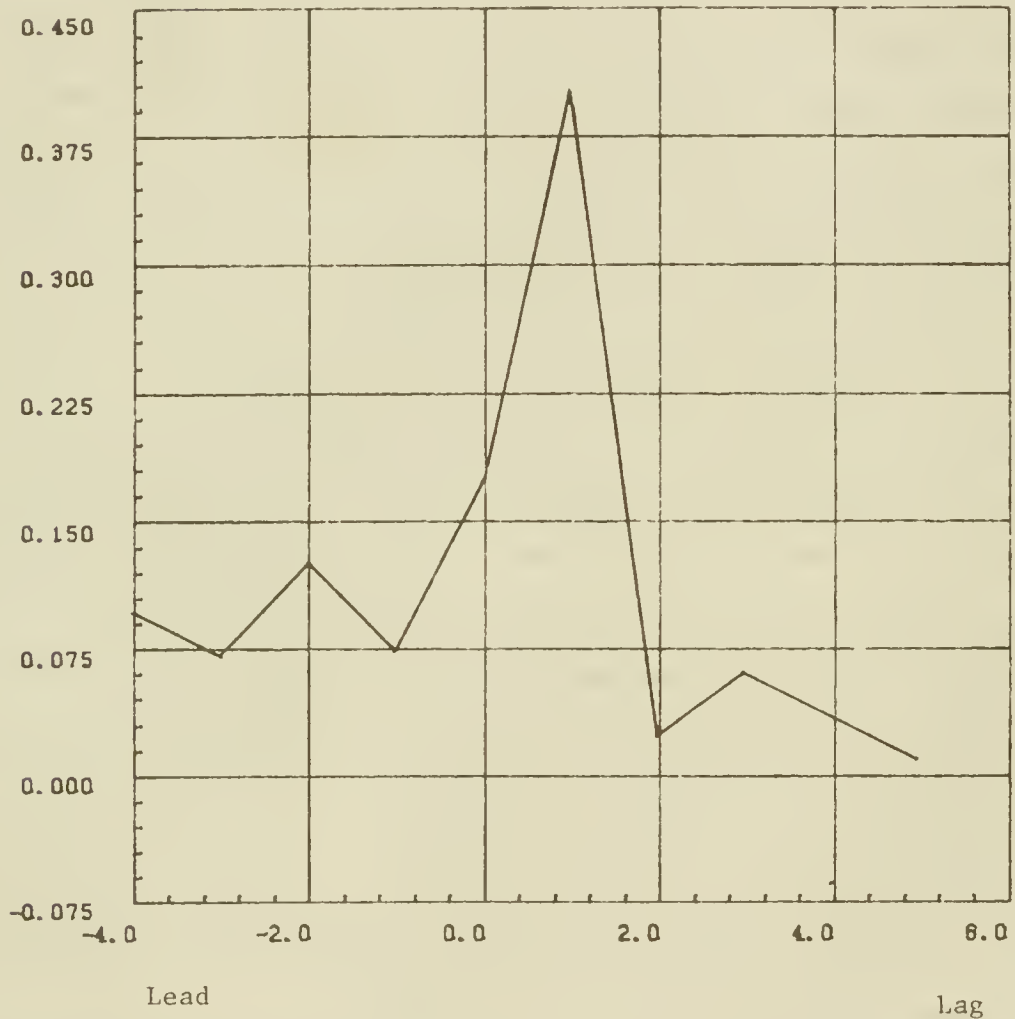
As the analysis in Sections IV and V will also confirm, these empirical findings are entirely consistent with the proposition that market price is a rational predictor of future dividends and that managers use rational assessments of permanent earnings to implement dividend policies along the lines of the Lintner stylized facts. The critical empirical results for supporting this conclusion are that (i) past (or possibly contemporaneous) changes in stock price explain a significant part of the change in dividends and (ii) the (partial) elasticity of dividends in response to a change in price is less than one. Ours is a simple model, and it is, of course, possible that more sophisticated models could produce empirical results contrary to our findings. If, however, the findings of such refined models were to support conditions (i) and (ii), then they would also support the

conclusion of our simple model that the observed volatility of stock prices and dividends in both the short and long runs is consistent with market efficiency. Because this conclusion depends so importantly on these conditions, we conclude this section with a further investigation of the empirical robustness of (i) and (ii).

Figure 1 plots the leads and lags of percentage changes in dividends regressed on the percentage changes in price estimated by the Hannan-efficient procedure. By inspection, the cross correlation at the lag in price change of one year specified in our model dominates that at all other leads and lags. In even a reasonably efficient market, one would not expect lagged variables of any sort to have meaningful predictive power for future price changes. It is therefore not surprising that changes in dividends are not significantly correlated with subsequent changes in price. The modest positive correlation between contemporaneous dividend and price changes is, of course, consistent with an efficient market and is perhaps suggestive of a mild information or announcement effect for dividend and price changes at the aggregate level. Indeed, we do find about an 8 percent correlation between contemporaneous (i.e., period $t + 1$) unanticipated price changes and the residuals from our regression (16). As noted in the discussion of the informational content of dividends in Section II, it is difficult to identify in a meaningful way an announcement date for aggregate dividends. Moreover, what is perceived to be contemporaneous correlation between dividend and price changes over the coarse grid of annual data may simply turn out to be lagged price changes explaining subsequent dividend changes when examined under the finer grids of quarterly or monthly data. Thus, an 8 percent correlation is likely to be a significant overstatement of the announcement effect of aggregate dividends.

Figure 1: Lead and lag structure of deviations in annual percentage real dividend changes, around their expected growth rate,¹ on percentage unexpected real price changes,¹ for the NYSE value-weighted companies over the period 1927-1980, computed using the Hannan-efficient procedure.

Point Estimate
of \hat{a}_1 in (*)



$$\text{Equation: } \log\left[\frac{D(t+j)}{D(t+j-1)}\right] + \frac{D(t+j-1)}{P(t+j-1)} = a_0 + a_1 \log\left[\frac{P(t)}{P(t-1)} + \frac{D(t-1)}{P(t-1)}\right] \quad (*)$$

$$j = -4, \dots, 0, \dots, +6$$

¹Dividends, $D(t)$, are defined as year t cash dividend payments for all NYSE companies, and the index "price" $P(t)$ is the beginning of year t value-weighted index. The dividend-price ratio, $D(t+j-1)/P(t+j-1)$ is, up to a constant, the expected rate of growth of dividends and prices in year $t+j$.

Unlike for speculative price changes, the Efficient Market Hypothesis does not rule out the change in dividends this period being predicted by variables which are observable prior to this period. Nevertheless, if the posited economic process underlying the specification of (14') is a reasonably accurate description of reality, then lagged price changes much beyond the one-year lag specified in (14') would be expected to exhibit relatively little explanatory power in forecasting this period's dividend change. If an unanticipated change in permanent earnings causes managers to change dividends, then strict rational behavior would seem to dictate that their decision be based on their most recent assessment of permanent earnings, and hence, earlier revisions in those assessments should have little impact on the change in dividends. In attempting to smooth the time path of dividends, it is possible that managers would choose to change dividends in response to changes in a moving average or distributed lag of unanticipated permanent earnings changes over an extended past history. Such behavior would create a dependency between the current change in dividends and lagged price changes of all orders. Because these averaging techniques embody so much "stale" information, it would appear that this approach to dividend smoothing leads to an inefficient use of the available information. If instead, managers change the dividend in a partial adjustment response to the most recent unanticipated change in permanent earnings, then they can achieve their dividend-smoothing objective and use the most up-to-date information.

Even if managers behave this rationally, there will still be some lag between a change in permanent earnings and the change it induces in subsequent dividends. As indicated in the discussion surrounding the specification of behavioral equation (10), at the level of aggregate dividends, it is unlikely

that this "minimal" lag is much shorter than a year.¹⁹ The correlations between the change in dividends and lagged price changes displayed in Figure 1 are therefore consistent with this view of rational behavior by managers. As a further, more quantitative test, we ran (14) with an additional five years of lagged unexpected price changes. None of these additional lagged variables had a coefficient point estimate more than one standard error from zero, and the F statistic for their inclusion is 0.361, which has a p value of 0.872.

If these empirical results had turned out differently, they would neither imply an arbitrage opportunity in the stock market nor an inefficiency in the allocation of capital. We need hardly mention again that managers have great latitude in their selection of dividend policies including the option to choose ones which are not based on the most up-to-date information. It is, however, reassuring for the overall validity of our model that the data tend to support such "superrational" behavior by managers even in the relatively unimportant area of dividend policy.

With the exception of contemporaneous changes in other speculative prices, it is a well-established empirical fact that there are few, if any, observable variables which exhibit high contemporaneous correlation with changes in aggregate stock prices. It is, therefore, rather unlikely that the change in stock price is merely serving in (14') as a proxy for some other observable variables which if included, would cause the significance of the coefficient on the price change to disappear. Hence, our important condition (1) that the lagged change in stock price explains a significant part of the next period's dividend change, is likely to be empirically robust.

As a check on this belief, we ran (14') with the addition of various lagged dividend and accounting earnings changes. The results were confirming

in that there were at most modest improvements in corrected- R^2 , and in all cases, the coefficient on the lagged logarithmic change in stock price remained significantly greater than zero and significantly less than one. The latter result also serves to confirm the robustness of the posited partial adjustment of dividends in response to an unanticipated change in permanent earnings.

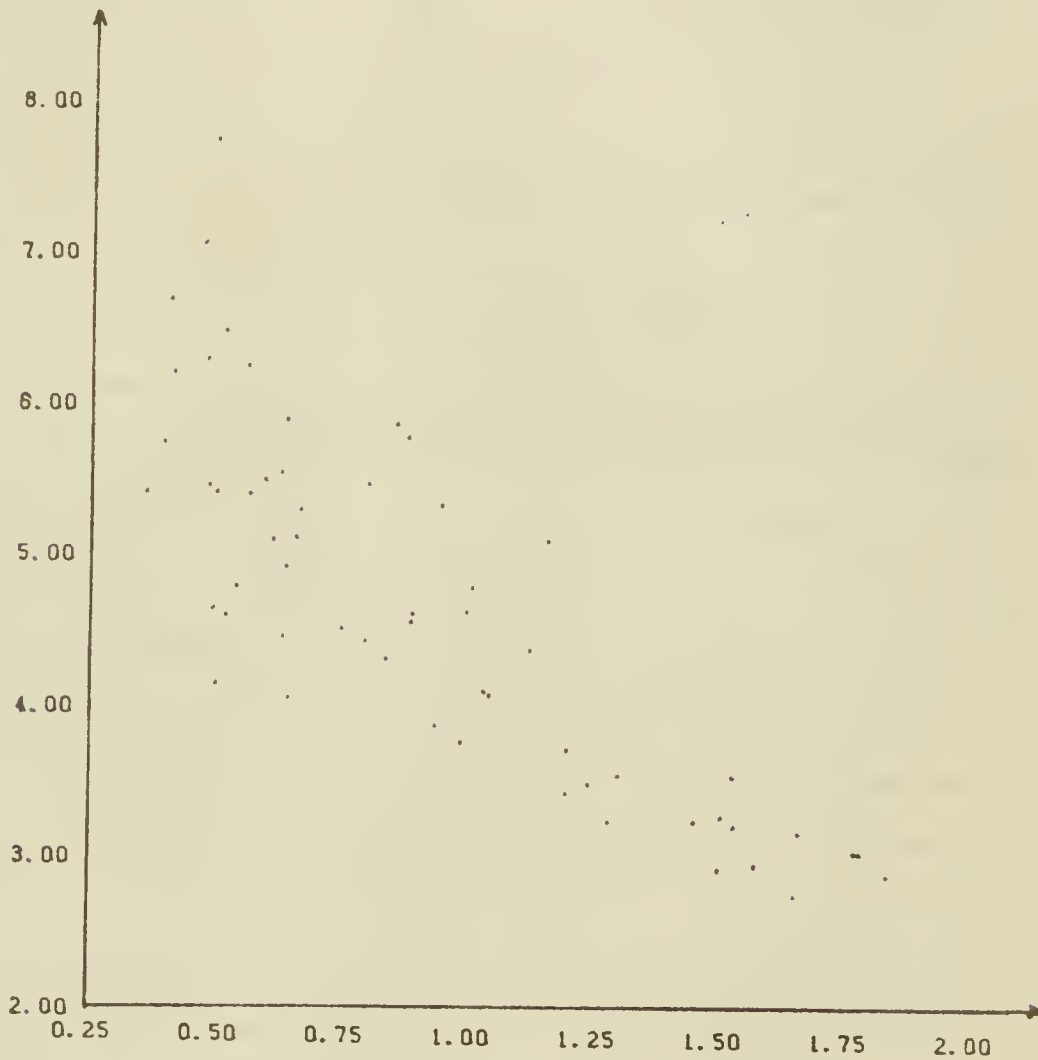
This regression result also supports our important condition (ii). That is, if the coefficient estimates are interpreted in a partial derivative sense, then the estimated partial elasticity of dividends in response to a change in price is less than one. It follows as a corollary to (ii) that the dividend-to-price ratio is not a systematically increasing function of the price. As further (admittedly, crude) evidence for this corollary, we present in Figure 2 a plot of the dividend-to-price ratio versus the price. A simple linear fit of the scatter produces a coefficient on the price which is negative and statistically significant. While we would hardly treat this result as convincing evidence for the stronger proposition that the dividend-to-price ratio is a systematically decreasing function of price, it surely supports our weaker proposition that the dividend-to-price ratio is not a systematically increasing function of price.

In summary, the empirical evidence, together with the theoretical foundation presented for its interpretation, leads us to conclude that conditions (i) and (ii) are very likely to remain robust even in the light of more-refined versions of the dividend model. We, therefore, treat these conditions as empirical facts for the balance of the analysis in this paper.

Armed with conditions (i) and (ii), we now prove that neither rationally-determined prices nor rationally-determined dividends have a

Figure 2: Plot of the dividend-price ratio against the level of the real NYSE value-weighted index price over the period 1927-1980.

Dividend-to-price ratio
(scaled by 10^{-2})



Level of real
NYSE prices
(scaled by 10^6)

nondegenerate steady-state distribution (with or without a trend). For analytic convenience only, we prove these results in the continuous-time version of the model. From equation (5), we have that the price dynamics can be written as:

$$\frac{dP}{P} = (\alpha - D/P)dt + \sigma dz \quad . \quad (17)$$

By Ito's lemma, it follows from (17) that:

$$d(\log P) = (\alpha - \frac{1}{2} \sigma^2 - D/P)dt + \sigma dz \quad . \quad (18)$$

As one of us has shown elsewhere,²⁰ if $\log P$ has a nondegenerate steady-state distribution, then there must exist numbers $\bar{P} < \infty$ and $\underline{P} > 0$ such that for (almost) all $P \geq \bar{P}$, $(\alpha - \frac{1}{2} \sigma^2 - D/P) < 0$ and for (almost) all $P < \underline{P}$, $(\alpha - \frac{1}{2} \sigma^2 - D/P) > 0$. By hypothesis, α and σ are constants. Hence, the full burden of these requirements falls upon the dividend-to-price ratio viewed as a function of price. As a practical matter, these requirements on D/P are likely to be met only if this ratio is a systematically increasing function of the price. From our corollary to empirical condition (ii), the dividend-to-price ratio is not such a function, and therefore, the price will not have a steady-state distribution. Moreover, the same conclusion applies as well to "detrended" price. That is, let $p(t) \equiv P(t)\exp[-gt]$ and $d(t) \equiv D(t)\exp[-gt]$ where g is some appropriately chosen deterministic trend rate. From (18), it follows that:

$$d(\log[p(t)]) = (\alpha - \frac{1}{2} \sigma^2 - g - d/p)dt + \sigma dz \quad (19) \quad .$$

Because $d(t)/p(t) = D(t)/P(t)$, the conditions on D/P for detrended price to have a steady-state distribution are the same as for the nondetrended price

except the constant " $\alpha - \sigma^2/2$ " is replaced by another constant " $\alpha - \sigma^2/2 - g$."

If, for analytical convenience, we neglect the empirically-established lag in the dividend response to an unanticipated change in price, then the continuous-time version of our reduced-form equation (14) can be written as:

$$\frac{dD}{D} = \left(\alpha - \frac{D}{P} + \gamma \left[b' - \log\left(\frac{D}{P} \right) \right] \right) dt + \lambda \sigma dz \quad (20)$$

and by Ito's lemma, we have from (20) that:

$$d(\log D) = \left[\alpha - \lambda^2 \sigma^2 / 2 - \frac{D}{P} + \gamma \left[b' - \log\left(\frac{D}{P} \right) \right] \right] dt + \lambda \sigma dz \quad (21)$$

By subtracting (18) from (21) and rearranging terms, we can write the dynamics for the dividend-to-price-ratio as:

$$d[\log(D/P)] = \gamma [\mu - \log(D/P)] dt - (1 - \lambda) \sigma dz \quad (22)$$

where $\mu \equiv b' + (1 - \lambda^2) \sigma^2 / 2\gamma$. By inspection of (22),

$\log[D/P]$ follows a classical Ornstein-Uhlenbeck process which is known to have a steady-state normal distribution with mean μ and variance $(1 - \lambda)^2 \sigma^2 / 2\gamma$. It follows from (22) that the dividend-to-price ratio has an asymptotic stationary lognormal distribution.

Although the dividend-to-price ratio has a steady-state distribution, the level of stock prices does not. It therefore follows that the level of dividends cannot have an asymptotic stationary distribution. Similarly, from the analysis of detrended price in (19), the level of detrended dividends, $d(t)$, does not have a stationary distribution. The conclusions drawn from our model are therefore in direct conflict with one of the fundamental postulates of the Shiller model: Namely, that dividends have a stationary distribution around some deterministic exponential trend.

Because the dividend-to-price ratio has a stationary distribution with a finite variance, it follows from (18) that for T large, $\text{Var}(\log[P(T)/P(0)])$ will be proportional to T . Hence, for large T , the dynamics for P in (17) can be well-approximated by a geometric Brownian motion with an instantaneous expected rate of growth equal to $(\alpha - \rho)$ where ρ is the expected "long-run" dividend-to-price ratio computed from the steady-state distribution for D/P . In this same sense, the asymptotic process for dividends will also be a geometric Brownian motion with $\text{Var}(\log[D(T)/D(0)])$ proportional to T .

In his reply to a comment on his work by Basil Copeland, Shiller (1983, p. 237) notes that "Even if we assumed log dividends were a random walk with trend with independent increments, stock prices still would show too much volatility." As he correctly points out, if $D(t+1)/D(t)$ is independent of $D(t'+1)/D(t')$ for $t \neq t'$, then the current dividend will be proportional to the current price (i.e., in our notation, $D(t) = \rho P(t)$). Hence, in such a model, the variance of logarithmic dividend changes will equal the variance of logarithmic price changes. Shiller goes on to report that for his Standard and Poor data set from 1871-1979, the sample standard deviation for log dividend changes is 0.127, whereas the sample standard deviation for log prices changes is 0.176. Because the ratio of sample variances of 1.93 is significant at the 1 percent level, he concludes that prices are too volatile to be consistent with this model. In our much-shorter 1926-1981 sample period, the standard deviation of dividend changes is virtually the same as in his sample (0.124), but the standard deviation of price changes is higher (0.203) which leads to a larger sample variance ratio of 2.64. We therefore agree with Shiller's conclusion, although our description would be that "the sample variations in dividend changes are too small to be consistent with this model."

" $D(t) = \rho P(t)$ " is the extreme polar case of our model where managers do not attempt to smooth dividends at all, and fully and immediately adjust dividends to reflect unanticipated changes in permanent earnings. From (18), the "short-run" instantaneous variance rate for logarithmic price changes is σ^2 whereas from (21), the corresponding variance rate for dividends is $\lambda^2 \sigma^2$. If managers fully adjust dividends in response to unanticipated changes in permanent earnings, then $\lambda = 1$, and the variance of dividend changes and price changes will be the same for all time intervals. If, as empirically appears to be the case, managers smooth the dividend time path by making partial adjustment responses, then $\lambda < 1$, and the variance of dividend changes will be strictly smaller than the variance of price changes. If, indeed, our model completely explained the process for dividend changes, then the coefficient estimate in (16) of 0.44 for λ , would imply that the ratio of the variances of annual log price changes to log dividend changes would exceed 5. Because the model explains only about 50 percent of the variation in dividends, the actual ratio is reduced to 2.64.

As is generally true of "smoothed" processes which are constrained to converge to a more-variable process, the variance rate of the percentage change in dividends increases as the interval over which it is computed is increased. It is, however, also the case that for $\lambda < 1$:

$$\frac{\text{Var}(\log[D(T)/D(0)])}{T} < \frac{\text{Var}(\log[P(T)/P(0)])}{T}$$

for any interval T and equality holds only in the limit as $T \rightarrow \infty$. Our model of rationally-determined prices and rationally-determined dividends, therefore, predicts that the variance rate of logarithmic dividend changes will always be smaller (and, at least for annual or shorter intervals, considerably

smaller) than the variance rate of logarithmic price changes. It is, hence, reassuring to find this prediction confirmed by Shiller's statistics which are based on a considerably longer sample period than our own.

In his 1981a paper (p. 428), Shiller discusses the higher-order moment properties of the stock price and dividend processes with a focus on the relation between infrequent arrivals of important information and the often-observed high-kurtosis (or "fat tail") sample characteristic of stock price changes. He demonstrates this relation by an illustrative example where dividends are taken to be independently and identically distributed. To capture the effect on stock price of infrequent arrivals of important information, he assumes that at each time t , with probability $1/n$, the market is told the current dividend level and with probability $(n - 1)/n$, the market has no information about current or future dividends. In this example, the kurtosis of the stock price change is shown to be n times greater than the kurtosis of the normal distribution posited for dividends. Our model, however, predicts the exact opposite result: Namely, dividend changes should exhibit relatively higher kurtosis than stock price changes. That is, although the variance of dividend changes is smaller than for stock price changes, the time series of dividends should contain more relatively small changes and more relatively large changes than the corresponding time series of stock prices.

The "lumpy" arrival of information (which may, in fact, cause the sample distribution of stock price changes to have fatter tails than a normal distribution) is not the source of this prediction about dividend changes. Instead, it comes as a result of managers smoothing the time path of dividends. To illustrate this point, we use an example which is very much

like Shiller's information example.

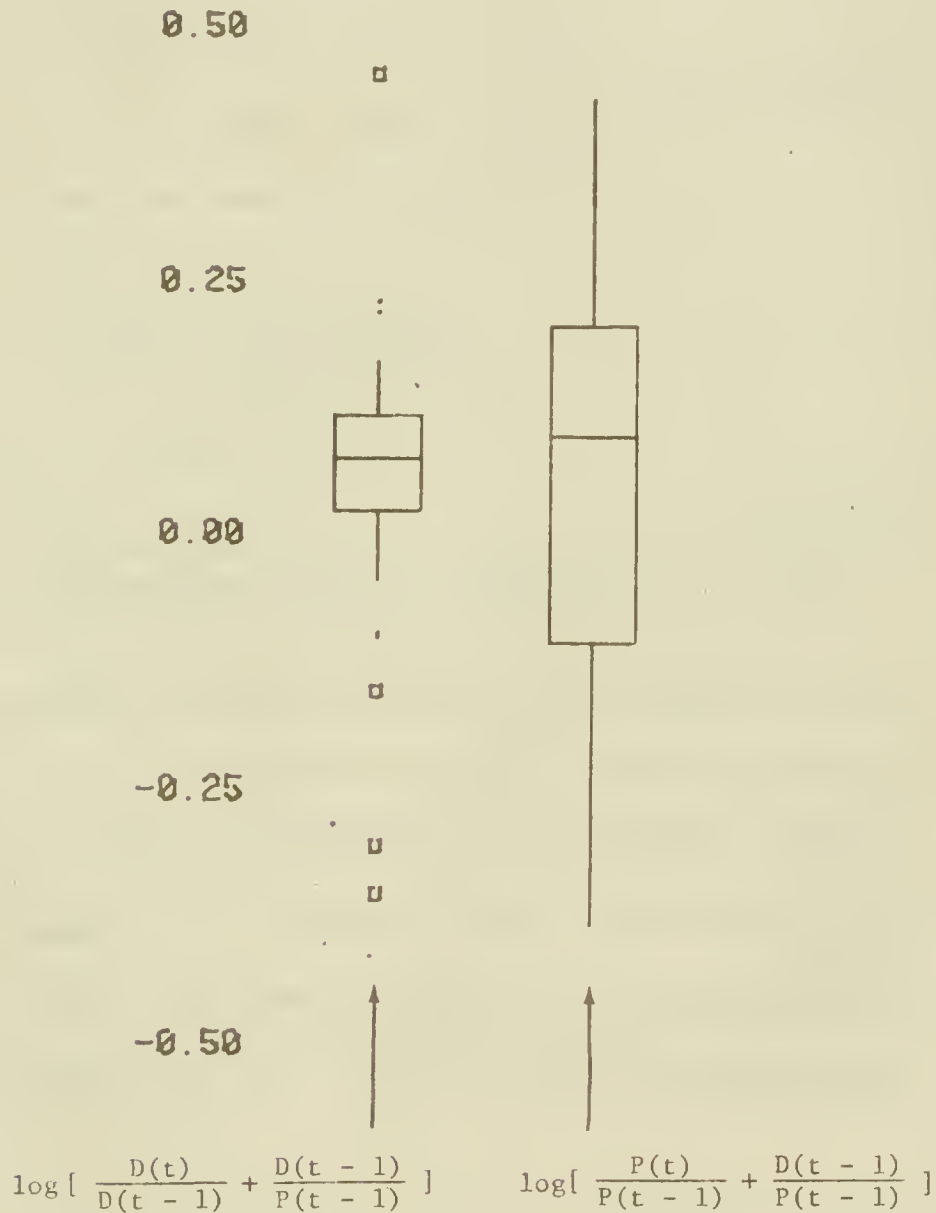
Suppose that unanticipated logarithmic changes in stock price are serially independent and identically distributed. Suppose further, that managers smooth the time path of dividends according to the following rule: At each time t , with probability $1/n$, they change the dividend to fully reflect the unanticipated change in stock price (i.e., $\log[D(t)/D(t-1)] = \log[P(t)/P(t-1)]$) and with probability $(n-1)/n$, they change the dividend to equal its expected normal growth rate (i.e., $\log[D(t)/D(t-1)] = g$). As expected for smoothed processes generally, in this example, the variance of dividend changes, σ^2/n , is smaller than the variance of stock price changes, σ^2 . If m_4 denotes the fourth central moment of the stock price change distribution, then the fourth moment of the dividend change distribution is m_4/n . Hence, the kurtosis of the dividend change process, $(m_4/n)/(\sigma^2/n)^2$, is simply n times the kurtosis of the stock price process, m_4/σ^4 . Thus, unless managers do not attempt to smooth dividends at all (i.e., $n = 1$), the kurtosis of the controlled dividend process will always exceed the kurtosis of the (uncontrolled) stock price process. Indeed, the more strongly that managers attempt to smooth dividends (i.e., the larger is n), the greater is the relative kurtosis of the dividend process.

In the light of this result, we estimated the kurtosis of each of the time series as a further empirical check on our model. As predicted, the estimated kurtosis for the annual logarithmic dividend changes, 7.377, is 2.79 times the estimate of 2.648 for the kurtosis of the logarithmic changes in stock prices. The sample kurtosis for stock prices is not much different than the kurtosis of 3 for a normal distribution whereas the sample kurtosis for

dividend changes is more than two times larger. As further evidence on the relative "fat tails" of the dividend process, the sample distributions of dividend and stock price changes are plotted in Figure 3, using Tukey's (1970) robust statistics. By inspection, these plots of the data are also consistent with the kurtosis prediction of our model.

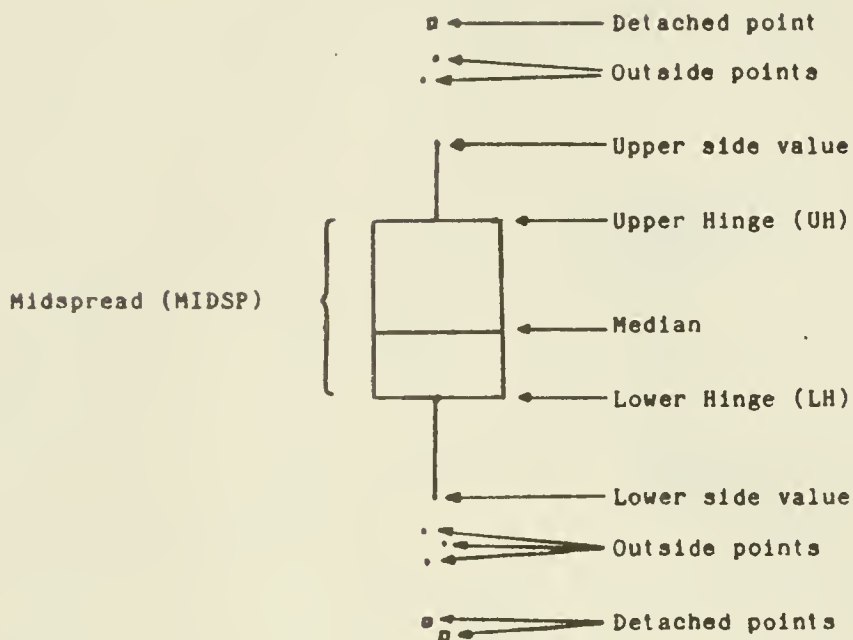
This completes the presentation of theoretical and empirical evidence on our model of rationally-determined prices and dividends. Because it is readily apparent that the predictions of the model directly conflict with those of the Shiller model, we proceed in the next sections to examine the evidence on that model's alternative specification of the dividend and rational stock price process.

Figure 3. Schematic or box-plot of the distribution of deviations in annual percentage real dividend changes around their expected growth rate,¹ and percentage unexpected real price changes,¹ for the value-weighted NYSE companies over the period 1927-1980.



(Figure continued.)

Figure 3. The following is provided to aid in interpreting the (Cont'd.) above graph.



Definitions:

The upper hinge (UH). If the number of observations (NOB) divided by 4 equals an integer i , then $UH = (X_{(4i)} + X_{(4i+1)})/2$. But if $NOB/4 = i + 1/4$, a non-integer, let $j = [4i + 1]$; then $UH = X_{(j)}$. Note that $X_{(i)}$ and $X_{(j)}$ are order statistics.

The lower hinge (LH). If $NOB/4 = i$, an integer; then $LH = (X_{(4i)} + X_{(4i+1)})/2$. But if $NOB/4 = i + 1/4$, where i is a noninteger, let $j = [4i + 1]$; then $LH = X_{(j)}$.

The midspread (MIDSP) is the distance $UH - LH$.

The upper side value is the largest data value less than $UH + MIDSP$. But if that value is less than UH , the upper side value equals UH .

The lower side value is defined similarly to the upper side value.

Outside points are data values between $UH + (3/2) * MIDSP$ and $UH + MIDSP$ or between $LH - (3/2) * MIDSP$ and $LH - MIDSP$.

Detached points are data values greater than $UH + (3/2) * MIDSP$ or less than $LH - (3/2) * MIDSP$.

IV. The Shiller Model: Empirical Evidence and Theory

The two important postulates of the Shiller Model of dividends and rational prices are as follows:

S.A.1 The dynamics of aggregate dividends are such that there exists a statistical equilibrium or steady-state probability distribution for the level of ("real" or inflation adjusted) dividends and this distribution has finite variance. A deterministic, exponential trend or growth rate for dividends is permitted, in which case detrended dividends, $d(t)(\equiv D(t)e^{-gt})$, have a steady-state probability distribution.

S.A.2 The required expected real rate of return on the stock market (or "the discount rate") is constant (or at least fairly stable).

It follows from assumptions S.A.1 and S.A.2 that the dynamics of aggregate stock market prices in an efficient market will be such that the level of real, detrended stock prices, $p(t)(\equiv P(t)e^{-gt})$, will have a steady-state probability distribution with a finite variance.

Although there is little supporting empirical evidence, it is fairly standard in the financial economics literature to assume that the expected real return on the market is constant. Quite aside from its empirical validity, this assumption is theoretically consistent with the assumption of risk-averse investors if the total returns on the market portfolio are described by a geometric Brownian motion. As already noted, this is the standard or prototype process for stock returns used in the literature, and, the one we have assumed in our model. Unless investors are risk-neutral and the riskless real rate of interest is constant, the assumption of a constant expected return on the market is, however, inconsistent with a stationary

process for dividends and stock prices (with or without a trend).²¹ If investors were risk-neutral, then the expected return on the market would be equal to the riskless rate of interest. While there are precious few hypotheses about stock market returns that can be empirically rejected, one that can is the hypothesis that the expected excess return on the market is zero.²² It appears, therefore, that assumption S.A.2 of the Shiller model is wholly inconsistent with assumption S.A.1 of that model. While this inconsistency issue is pursued no further here, it remains an empirical puzzle as to why stock returns on average are so much higher than bond returns, if dividends can be empirically described by a stationary process.

As the empirical basis for the stationarity assumption S.A.1 in his various papers on the subject, Shiller relies almost exclusively on regression analyses to show that dividends follow an autoregressive process with a time trend.

In his most recent published evidence, Shiller (1983) reports that "If $\log D(t)$ is regressed on $\log D(t-1)$, a constant and a linear time trend for 1872 to 1978, the coefficient of $\log D(t-1)$ is 0.807, with an estimated standard error of 0.058," implying that "log dividends would always be expected to return half way to the trend line in three years," (p. 237). We repeated essentially the same OLS regression on our data set, and the results are:

$$\Delta \log D(t) = 2.492 - 0.249 \log D(t) + 0.004 t + u(t) \quad (23)$$

(0.890) (0.089) (0.002)

$$R^2 = 0.130 \quad DW = 1.495$$

Because the left side of (23) is the change in $\log D(t)$, the comparable autoregressive coefficient is $(1 - 0.249) = 0.751$ which is rather close to Shiller's 0.807 estimate. By the standards of a conventional t-test, the coefficients in both samples are significantly less than 1 with a -2.80 t-statistic in our sample and a -3.33 t-statistic in his. The t-statistic for the trend coefficient in (23) is 2.00. This finding serves to confirm our belief that the important empirical results derived from our 1926-1981 data set will not be significantly altered if fit to the longer 1871-1979 data set used by Shiller in his analysis.

Because the Durbin Watson statistic suggests positively autocorrelated residuals, the lagged endogeneous variable in (23) may cause the OLS coefficients to be biased. We therefore reestimated (23) using a GLS iterative technique,²³ and the results are:

$$\Delta \log D(t) = \begin{matrix} 5.225 \\ (1.211) \end{matrix} - \begin{matrix} 0.524 \\ (0.121) \end{matrix} \log D(t) + \begin{matrix} 0.009 \\ (0.003) \end{matrix} t + u(t) \quad (24)$$

$$\bar{R}^2 = 0.243 \quad \text{DW} = 1.85$$

Hence, when the OLS specification of Shiller's autoregressive model for dividends is correctly adjusted for autocorrelation, its measured explanatory power almost doubles. Only about half of the total 24 percent explanation of the variation in dividend changes can be attributed to the lagged dividend and time trend variables. The other half is attributable to the time series model of the disturbances or "unknown variables" in the regression.²⁴

Given the apparent statistical significance of the coefficients, it might appear from (23) and (24) that dividends follow an autoregressive process

which approaches a steady-state distribution (possibly around a positive trend). If (23) fully described the "true" dividend process (i.e., $\bar{R}^2 = 1$), then this result, along with a casual inspection of the stock return time series, would surely imply that stock prices are "too volatile." If, as implied by (24), dividends were to regress over 90 percent of the way to their deterministic trend line within one Presidential term, then uncertainty about the future path of dividends would be rather unimportant, and therefore, rational stock prices should exhibit trivial fluctuations. Because equities are the residual claims of the private sector, variations in their returns are "blown up" reflections of the uncertainties about the whole economy. If stock returns should have small variations, then the fluctuations in the economy should be even smaller. It would therefore seem that in such an environment we economists could safely neglect such uncertainties in the specification of our macroeconomic models. While perhaps an appealing hypothesis, the real world is not this way as further analysis of (23) and (24) will clearly indicate.

The fit of the autoregressive model (23) and (24) is rather poor, with half of the explanatory power of (24) represented by unspecified variables. With respect to a different regression but similar data, Shiller (1981a, p. 433) gives one possible explanation for low R^2 : Namely, "...regression tests are not insensitive to data misalignment. Such low R^2 might be the result of dividend or commodity price index data errors." Although we agree that such data errors can be a source for lower R^2 , our explanation is simply that the autoregressive process posited in (23) and (24) is not an accurate specification of the dividend process.

Motivated by the analysis of our model of the dividend process, we add the one-year lagged unanticipated change in the log of stock price to the specification of (24). By the same iterative GLS procedure used in (24), the results are:

$$\begin{aligned} \Delta \log D(t) = & 2.107 + 0.347 \log \left[\frac{P(t)}{P(t-1)} + \frac{D(t-1)}{P(t-1)} \right] \\ & (0.918) \quad (0.064) \\ & - 0.213 \log D(t) + 0.004 t + u(t) \\ & (0.092) \quad (0.002) \end{aligned} \tag{25}$$

$$\bar{R}^2 = 0.473 \qquad DW = 1.755$$

By inspection, the addition of the previous year's unexpected price change in (25) doubles the explanatory power of (24). This measured increase in \bar{R}^2 greatly understates the impact of this added variable because, in addition to increasing \bar{R}^2 by 100 percent, it also virtually eliminates the explanatory power of the remaining unspecified variables whose effects are captured by the GLS procedure.²⁵ We would also note that by adding the log price change variable, the absolute magnitudes of both the $\log D(t)$ and time trend coefficients are cut in half.

To further explore the relative importance of the specified variables, equation (25) was reestimated: first, with the time trend deleted, and second, with both the time trend and $\log D(t)$ removed. The results are:

$$\begin{aligned} \Delta \log D(t) = & 0.566 + 0.388 \log \left[\frac{P(t)}{P(t-1)} + \frac{D(t-1)}{P(t-1)} \right] \\ & (0.564) \quad (0.064) \\ & - 0.055 \log D(t) + u(t) \\ & (0.053) \end{aligned} \tag{26.a}$$

$$\bar{R}^2 = 0.437 \qquad DW = 1.814$$

and

$$\Delta \log D(t) = \frac{-0.014}{(0.016)} + \frac{0.402}{(0.063)} \log \left[\frac{P(t)}{P(t-1)} + \frac{D(t-1)}{P(t-1)} \right] + u(t) \quad (26.b)$$

$$\bar{R}^2 = 0.435$$

$$DW = 1.82$$

Comparing (25) with (26.a), the elimination of the time trend variable causes only a modest reduction in \bar{R}^2 , and it has little effect on the estimated coefficient of the log price change. In contrast, the magnitude of the regressive coefficient on $\log D(t)$ falls by 75 percent, and with a p value equal to 0.309, it is not statistically significant. It would appear that there is a strong interaction between $\log D(t)$ and the time trend which together with the GLS iterative procedure is responsible for the significant coefficients in (24). If either variable is removed, then the magnitude and the statistical significance of the coefficient of the remaining variable are nil. In this light, it is not surprising that the elimination of $\log D(t)$ as a variable in (26.b) has no effect on either the \bar{R}^2 of that equation or the coefficient of log price change.

Unless the log price change can be "explained" by some distributed lag of past dividends (which, as an empirical matter, it cannot), then it surely belongs in the specification of the dividend process. Because it alone accounts for over 90 percent of the explanatory power of (25), its omission from (23) and (24) is rather important. In sharp contrast, the elimination of either the time trend or the $\log D(t)$ variables has no significant effect on the fit. Hence, unless there are strong a priori economic reasons to believe that these variables belong in the specification of the dividend process, there appears to be no valid empirical reason for their inclusion.

In our model of the dividend process, there is no role for a time trend and its inclusion in equation (15) or (16), not only produces a highly insignificant coefficient, but actually causes the corrected OLS and GLS R^2 's to fall. Our model would, however, predict that changes in log dividends are related to $\log D(t)$ through the dividend-to-price ratio, $\log[D(t)/P(t)]$. Although not explicitly specified in our simple model, it is also possible that lagged changes in log dividends may also explain part of the adjustment process used by managers to decide upon subsequent dividend changes. If this were the case, then $\log D(t)$ may be a proxy for these lagged changes. The inclusion of lagged dividend changes would in no substantive way change the conclusions derived for the dividend and rational stock price processes in Section III. To investigate this possibility, we reestimate equation (16) with the addition of $\log D(t)$ and the fitted results are given by:

$$\begin{aligned} \log\left[\frac{D(t+1)}{D(t)}\right] + \frac{D(t)}{P(t)} &= \underset{(1.923)}{1.550} + \underset{(0.061)}{0.441} \log\left[\frac{P(t)}{P(t-1)} + \frac{D(t-1)}{P(t-1)}\right] \\ &- \underset{(0.267)}{0.247} \log\left[\frac{D(t)}{P(t)}\right] - \underset{(0.256)}{0.220} \log D(t) + u(t+1) \end{aligned} \tag{27}$$

$$\bar{R}^2 = 0.532 \qquad \text{DW} = 1.88$$

The F statistic for $\log D(t)$ in (27) is 3.23 which is insignificant at the 5 percent level. If $\log D(t - 1)$ is added to (27), its estimated coefficient is -0.135 while that of $\log D(t)$ becomes 0.064. While this result suggests that $\log D(t)$ in (25) may be a proxy for $\log[D(t)/D(t - 1)]$, the coefficient of $\log D(t - 1)$ is also statistically insignificant. As expected, the addition of these lagged dividend variables

in (27) had no effect on the point estimate of the coefficient of lagged log price change.²⁶

In summary, adding the trend and lagged dividend variables of the Shiller model to our specification does little to improve the explanatory power of our model in terms of \bar{R}^2 . Moreover, the estimated coefficients of these "added" variables are statistically insignificant. The other side of this result is that the addition of the variables from our model to the autoregressive specification (23) substantially increases the explanatory power of that model. As these variables are added, however, the statistical significance of the autoregressive variables is reduced. This result is perhaps surprising because the dividend time series is often thought to be quite smooth, while the price series is volatile, and therefore, a distributed lag of past dividends together with a time trend might be expected to do a better job than stock price changes in explaining subsequent dividend changes, almost independently of the "true" economic specification.²⁷

Although there appears to be no significant empirical evidence for regressivity in the time series of dividends, the lack of such evidence does not disprove the hypothesis that dividends have a stationary distribution around a deterministic trend. As discussed at length in Section II, the resolution of such an issue must ultimately come from economic reasoning. As Shiller (1983, p. 236) has noted on the specification issue, "Of course, we do not literally believe with certainty all the assumptions in the model which are the basis of testing. I did not intend to assert in the paper that I knew dividends were indeed stationary around the historical trend." We surely echo this view with respect to the theoretical assumptions underlying our own empirical model. Nevertheless, unlike our model's assumptions, the assumption

that dividends follow a stationary process with a trend does not appear to have any nonstatistical foundation. In our review of the variance bounds literature, we found no discussion of the theoretical structure which supports this assumption. This is especially important because there is neither an oral nor a written tradition in the financial economics literature that assumes dividends and rational stock prices have stationary distributions.²⁸

One could, of course, revive Malthus, and use his "limits-to-growth" view to justify the assumption of a steady-state distribution for the levels of dividends and prices. This theory, however, also rules out an exponential trend in these levels. Refitting equation (23) without the trend, we have that the OLS estimate is given by:

$$\Delta \log D(t) = \begin{matrix} 0.802 & - & 0.076 & \log D(t-1) & + & u(t) \\ (0.576) & & (0.055) & & & \end{matrix} \quad (28)$$

$$\bar{R}^2 = 0.017 \quad DW = 1.576 \quad .$$

By inspection of (28), it would appear that at least in the dividend series, there is no evidence at all to support this "zero-growth" model.

Notable perhaps by its absence from this section is any discussion of tests of stationarity of stock price levels. We have not directly tested the time series of price changes for evidence of regressivity because the literature is almost uniform in failing to find any lagged variables which forecast future stock price changes. Moreover, using autocorrelation and Dickey-Fuller (1979) tests, Kleidon (1983) finds that neither the arithmetic nor the geometric Brownian motion models can be rejected against the trend model for the S&P 500 annual composite index over the period 1926-1979. We would expect that these same results would obtain for our data set.

Shiller (1981a, p. 432-433) does report that the dividend-to-price ratio appears to forecast next period's holding period returns. We replicated this result on our data set, and found, as did he, that the R^2 is about 0.06. As Shiller himself stresses, such regression tests are sensitive to data problems, and such problems could easily explain the positive relation between stock returns and lagged dividend yield. Study of our sample using blunt interocular analysis suggests that there are a few influential "outliers" in the annual data which cause the correlation, and the correlation disappears all together with monthly data. Even if the findings were "truly there," the modest R^2 is consistent with market efficiency if the expected return on the market changes and such changes are positively correlated with the current $D(t)/P(t)$. Finally, if the expected return on the market is constant and therefore, D/P does predict "abnormal" returns, then this finding would provide, at best, modest evidence of some market inefficiency. If, however, this evidence is seriously accepted, then it is also evidence against a stationary distribution for stock prices. As was shown in Section III, a condition for stock prices to have a stationary distribution is that the dividend-to-price ratio must be an increasing function of price. The claimed positive correlation between the current dividend-to-price ratio and subsequent price changes directly conflicts with this condition.

In closing this section, we emphasize that throughout his published studies of stock market volatility, Shiller has not relied upon his regression models as the principal evidence for his claim of excess stock price volatility. For that, he has used "long-run" variance bounds test. Nevertheless, as our analysis in the next section will show, his interpretation of those tests' results depends crucially on the assumption

that (detrended) dividends are generated by a stationary process--a hypothesis which the analysis in this section shows to be empirically unfounded.

V. Variance Bounds Tests and the Shiller Model

In his June 1981 American Economic Review article, Shiller concluded that:

measures of stock price volatility over the past century appear to be far too high--five to thirteen times too high--to be attributed to new information about future real dividends if uncertainty about future dividends is measured by the sample standard deviations of real dividends around their long-run exponential path (p. 434).

Although this assertion of implied market efficiency is not consistent with the mainstream of thought in modern financial economics, there is a long history of academic economists and financial practitioners making similar claims.²⁹ These earlier claims could be largely dismissed, however, because their empirical foundations were, at best, anecdotal. In contrast, the estimates of large excess stock price volatility in the Shiller study arise from seemingly robust statistical tests applied to data from the very long sample period 1871-1979. As such, the study presents a serious empirical challenge to the validity of efficient market (and more generally, rational expectations) models as "reasonable" abstractions of real-world economic behavior.

As noted in the previous section, Shiller does not draw his cited conclusions from regression tests, but instead chooses the alternative methodology of variance bounds tests which he believes is better-suited for the investigation of long-term stock market volatility.³⁰ The strict validity of his tests requires that the time path of detrended dividends follow a stationary process--an assumption which our analysis has already shown to be without empirical or theoretical foundation. We nevertheless, explore the properties of the statistics computed in these tests to see whether or not their estimated values are consistent with our alternative model of rationally-determined dividends and prices.

Shiller (1981a) performs two variance bound tests of rational stock prices within the context of his model. The first of these tests, the "p* test," compares the variance of the level of detrended real stock prices with the variance of an "ex post rational" detrended real price series, $\{p^*(t)\}$. The second, called the "innovations test," compares the variance of the unanticipated change in detrended real stock prices to the variance of detrended real dividends. "Detrending" as used here is the same as in the development at the end of Section III where, for example, the detrended dividend at time t is given by $d(t) \equiv D(t)e^{-gt}$ where g is the long-run deterministic growth rate of real dividends and stock prices. Hence, $g = \alpha - \rho$ where ρ is the long-run dividend component of the total expected real rate of return on stocks, α . Our analysis of the tests begins with a brief description of Shiller's (1981a) presentation, but using our notation.

Define the ex post rational detrended price per share in the market portfolio at time t to be:

$$p^*(t) \equiv \sum_{k=0}^{\infty} \eta^k d(t+k) \quad (29)$$

where $\eta \equiv 1/(1+\rho)$.³¹ $p^*(t)$ is called an ex post rational price because it is the present value of actual subsequent (to time t) dividends. If the actual detrended stock price at time t is an ex ante rational price, then it follows from (1) and (29) that:

$$p(t) = \varepsilon_t[p^*(t)] \quad (30)$$

where, as defined before, ε_t is the expectation operator conditional on all information available as of time t . Therefore, $p(t)$ is forecast of $p^*(t)$. As Shiller (1981a, p. 422) points out, if $p(t)$ is an ex ante

rational price, then it is an optimal forecast of $p^*(t)$. If $u(t) \equiv p^*(t) - p(t)$ is defined to be the forecast error, then $u(t)$ and $p(t)$ should be uncorrelated if $p(t)$ is such an optimal forecast. Under this hypothesis, it therefore follows that $\text{Var}[p^*(t)] = \text{Var}[p(t)] + \text{Var}[u(t)] \geq \text{Var}[p(t)]$. That is, in a set of repeated experiments where a forecast $p(t)$ and a sequence of subsequent dividends, $d(t+k)$, $k = 0, 1, \dots$, are "drawn," it should turn out that the sample variance of the $p^*(t)$ exceeds the sample variance of the forecasts $p(t)$.

If, as is posited in the Shiller model, detrended dividends follow a stationary process, then rational detrended prices will also be stationary. From this assumption, it follows by the Ergodic Theorem that the time series ensembles of $\{p(t)\}$ and $\{p^*(t)\}$ can be used to test the "cross-sectional" proposition, $\text{Var}[p^*(t)] \geq \text{Var}[p(t)]$.

Because the sample time period from which the data are drawn is finite, it is, of course, necessary to truncate the summation in (29) to compute an estimate of $p^*(t)$. If, as Shiller notes (1981a, p. 425), the time series sample is "long enough," then the truncated estimate of $p^*(t)$ should provide a reasonable estimate of the true value.³² To correct for this truncation, Shiller sets a particular value for the "terminal" $p^*(T)$ which he chooses to be the average of the detrended stock prices over the sample period. That is,

$$p^*(T) = \bar{p} \quad (31)$$

where $\bar{p} \equiv \left[\sum_{k=0}^{T-1} p(t) \right] / T$, where T is the number of years in the sample.

From (29) and (31), the constructed p^* series is given by:

$$p^*(t) = \sum_{k=0}^{T-t-1} \eta^k d(t+k) + \eta^{T-t} p^*(T), \quad t = 0, \dots, T-1 \quad (32)$$

Using (32), the sample variance of p^* is computed by:

$$\text{Var}(p^*) = \left(\sum_{k=0}^{T-1} [p^*(t) - \bar{p}^*]^2 \right) / T \quad (33)$$

where $\bar{p}^* \equiv \left[\sum_{k=0}^{T-1} p^*(t) \right] / T$. With a similar computation for the sample variance of p , the null hypothesis of the p^* test for rational stock prices can be written as:

$$\text{Var}[p^*] \geq \text{Var}[p] \quad . \quad (34)$$

As reported in Shiller's Table 2 (1981a, p. 431), the variance bound in (34) is grossly violated by both his Standard and Poor's 1871-1979 data set and his modified Dow Industrial 1928-1979 data set.

In his second test of rational stock prices, Shiller (1981a, pp. 425-427) uses the time series of "price innovations" (i.e., the unanticipated part of the annual change in detrended stock price) which he denotes by $\delta p(t) \equiv p(t) - p(t-1) + d(t-1) - \rho p(t-1)$. Under the maintained assumption that detrended dividends have a stationary distribution, he derives as a condition for rational stock prices that:

$$\text{Var}(d) \geq \text{Var}(\delta p) [(1 + \rho)^2 - 1] \quad (35)$$

where $\text{Var}(d)$ and $\text{Var}(\delta p)$ denote the sample variances of the level of detrended dividends and the innovations of price changes, respectively. As reported in the cited Table 2, the null hypothesis of rational stock prices is, once again, grossly violated by both the data sets.

Although Shiller reports none of the finite sampling properties for the estimated statistics in (34) and (35), it would appear unnecessary since the sample period is so long and the point estimates are so large. The validity of these seemingly-powerful variance bounds tests depends crucially on the

assumption that detrended real dividends have a stationary distribution. Because the theoretical and empirical evidence presented in the preceding sections is wholly inconsistent with this assumption, there would appear to be little point in further investigation of these tests. We, however, undertake such further analysis to show that Shiller's estimated test statistics are, in fact, consistent with our model where the dividend and stock price processes asymptotically follow geometric Brownian motions.

If $p(t)$ and $d(t)$ follow such processes, the variances of the price and dividend are, of course, not well-defined in the time series sense that they were used in Shiller's variance bound tests. However, as defined in (33), $\text{Var}(p^*)$, $\text{Var}(p)$, and $\text{Var}(d)$ can be simply treated as statistics constructed from the random variables $\{p(t)\}$ and $\{d(t)\}$, and for any finite T , the moments of their distributions will exist. As will be shown, moreover, the conditional expectation of the ratios, $\text{Var}[p^*]/\text{Var}[p]$ and $\text{Var}[\delta p]/\text{Var}[d]$, will exist even in the limit as $T \rightarrow \infty$.

We now show that the properties of the p^* -test statistics within the context of our model are such that inequality given in (34) is exactly reversed. To prove this result, consider the extreme polar case of our model where managers fully and immediately adjust dividends to maintain the target dividend payout ratio ρ [I.e., $D(t) = \rho P(t)$].

Proposition: If $D(t) = \rho P(t)$ for all t , then, for each and every sample path of stock price realizations, $\text{Var}(p^*) > \text{Var}(p)$, with equality holding if and only if all realized prices are identical in the sample $t = 0, \dots, T - 1$.

Proof: By hypothesis, $d(t) = \rho p(t)$ for $t = 0, \dots, T - 1$. From (32), $p^*(t)$ can therefore be rewritten as:

$$p^*(t) = \rho \sum_{k=0}^{T-t-1} \eta^k p(t+k) + h^{T-t} p^*(T), \quad j = 1, \dots, T \quad (36)$$

From (31) and (36), $p^*(t)$ can be expressed as a weighted sum of the realized prices in the sample period which we write as:

$$p^*(T-j) = \sum_{k=1}^T w_{jk} p(T-k), \quad j = 1, \dots, T \quad (37)$$

where $j \equiv T - t$ and for each j , $j = 1, \dots, T$, the weights $\{w_{jk}\}$ are given by:

$$\begin{aligned} w_{jk} &= \rho \eta^k + \frac{\eta^j}{T}, \quad k = 1, \dots, j \\ &= \frac{\eta^j}{T}, \quad k = j + 1, \dots, T \end{aligned} \quad (38)$$

By inspection of (38), $w_{jk} \geq 0$ and it is straightforward to show

that $\sum_{k=1}^T w_{jk} = 1$. Thus, each $p^*(t)$ is just a convex combination of the sequence of realized stock prices.

We can express (37) in matrix form as $\underline{p}^* = W\underline{p}$ where \underline{p}^* and \underline{p} are $T \times 1$ vectors and $W = [w_{jk}]$ is a $T \times T$ matrix. W can be rewritten as:

$$W = \frac{1}{T} X + \rho Y \quad (39)$$

where the matrices X and Y are given by:

$$X \equiv \begin{bmatrix} 1 & \dots & 1 \\ \eta & \dots & \eta \\ \eta^2 & \dots & \eta^2 \\ \vdots & & \vdots \\ \eta^{T-1} & \dots & \eta^{T-1} \end{bmatrix}; \quad Y \equiv \begin{bmatrix} 0 & 0 & \dots & 0 \\ \eta & 0 & \dots & 0 \\ \eta^2 & \eta & 0 & \dots & 0 \\ \vdots & \cdot & \vdots & \cdot & \vdots \\ \eta^{T-1} & \dots & \dots & \eta & 0 \end{bmatrix} \quad (40)$$

The X matrix contains that part of the weights on prices which enter $p^*(t)$ from the terminal condition $p^*(T)$, and the Y matrix contains the part which comes from dividends. From (39) and (40), we have that

$$\text{Var}(p^*) = [p - \bar{p}]' W' (I - \frac{1}{T} \underline{1} \underline{1}') W [p - \bar{p}] \quad (41)$$

where: $\underline{1}$ is a $T \times 1$ vector of ones; I is the $T \times T$ identity matrix; $[p - \bar{p}]$ is the vector of realized prices measured as deviations about their sample mean, \bar{p} . To establish $\text{Var}(p^*) \leq \text{Var}(p)$, we prove that the (inner product) norm of the quadratic form given by (41) is less than the norm of the deviations of sample prices around their mean, $\|p - \bar{p}\|$. By inspection of (40), the rows of the X matrix have identical elements, and hence, $X[p - \bar{p}] = 0$. It follows from (39) that

$$\|W[p - \bar{p}]\| = \|\rho Y[p - \bar{p}]\| \quad (42)$$

From (39), the matrix Y can be decomposed into the sum of T matrices given by:³³

$$Y = \sum_{i=1}^{T-1} \eta^i C_i \quad (43)$$

where C_i is a matrix with zeros everywhere except for ones on the i th subdiagonal. From (43), $\rho Y[p - \bar{p}] = \rho \sum_{i=1}^{T-1} \eta^i C_i [p - \bar{p}]$. It follows by the Minkowski inequality that

$$\|\rho Y[p - \bar{p}]\| \leq \rho \sum_{i=1}^{T-1} \eta^i \|C_i [p - \bar{p}]\| \quad (44)$$

Moreover, we have that

$$\|C_i [p - \bar{p}]\| \leq \|p - \bar{p}\| \quad (45)$$

because each vector $C_i[\underline{p} - \bar{p}]$ has at least one zero element, and hence has shorter length than the vector $[\underline{p} - \bar{p}]$ which it simply "shuffles." From (44) and (45), we have therefore that

$$\| \rho Y[\underline{p} - \bar{p}] \| \leq \rho \sum_{i=1}^{T-1} \eta^i \| \underline{p} - \bar{p} \| \quad (46)$$

But, $\eta = 1/(1 + \rho)$ and $\| \underline{p} - \bar{p} \|^2 = T \text{Var}(p)$. Hence, from (42) and (46), $\| W[\underline{p} - \bar{p}] \|^2 \leq [1 - (1 + \rho)^{1-T}]^2 T \text{Var}(p)$. With this established, it is trivial to show that

$$\| (I - \frac{1}{T} \underline{1} \underline{1}') W[\underline{p} - \bar{p}] \|^2 / T \leq [1 - (1 + \rho)^{1-T}]^2 \text{Var}(p). \text{ Thus,}$$

$\text{Var}(p^*) \leq \text{Var}(p)$ with equality holding only if all the elements of \underline{p} are the same and $\text{Var}(p) = 0$. Q.E.D.

Because the variance bound proposition applies to each and every time path of prices, the derived inequality, $\text{Var}(p^*) \leq \text{Var}(p)$, holds in-sample. A fortiori, it will obtain in the population for any distribution of prices. Thus, even for a "bad draw," $\text{Var}(p^*)$ will not exceed $\text{Var}(p)$.

While the proposition holds for any distribution of prices, its proof does use the hypothesized condition that $d(t) = \rho p(t + 1)$ -- a condition which is inconsistent with $d(t)$ having a stationary distribution with a positive variance. Our proposition, therefore, does not imply that Shiller's "reverse" condition, $\text{Var}(p^*) \geq \text{Var}(p)$, is mathematically incorrect under the posited conditions of his model. It does, however, imply that the p^* -test is very sensitive to the strict stationarity assumption for the detrended dividend process, and together with the time series analysis in Section IV, it does raise significant doubts about the empirical validity of this test.

As discussed in Section III, the " $d(t) = \rho p(t)$ " model is also not empirically valid (except in the very long run) because it implies that managers do not smooth the time path of dividends whereas the empirical evidence appears to be that they do. The effect of such dividend smoothing is to reduce the variance of dividends below that which would obtain if they were not smoothed and $d(t) = \rho p(t)$. The "ex post rational" price series $p^*(t)$ is, by definition, a weighted sum of these dividends (and through the terminal $p^*(T)$, the sample period's path of prices) where the weights are positive and sum to one.³⁴ The reduction in the variance of dividends from smoothing will, therefore, cause $\text{Var}(p^*)$ to be smaller than it would otherwise have been if $d(t) = \rho p(t)$. If stock prices are rationally determined, such "endogeneous" reductions in the variance of the dividend path will not, however, cause a corresponding reduction in $\text{Var}(p)$.³⁵ Hence, under our model of the dividend and stock price processes, it would not be surprising to find that our derived variance bound, $\text{Var}(p^*)/\text{Var}(p) \leq 1$ is not binding and that $\text{Var}(p^*)/\text{Var}(p) \ll 1$.

To provide a more-visual (if less-quantitatively precise) representation of the "excess volatility" of stock prices, Shiller (1981a, p. 422) plots the time series of the levels of actual stock prices and the computed "ex post rational" prices, $p^*(t)$. By inspection of his plots, it is readily apparent that $p(t)$ is more volatile than $p^*(t)$. Instead of implying "too much" stock price volatility, we interpret these plots as implying that p^* has "too little" volatility to be consistent with a dividend process which is not smoothed. They are, however entirely consistent with the ff of our equation (16), which shows that managers undertake considerable effort to smooth dividends and that the speed of convergence in the dividend-to-price ratio to

its "long-run" target is very slow. In Shiller's plots, it also appears that the levels of actual prices "revert" toward the p^* "trend line." This apparent correspondence in trend is not surprising since $p^*(t)$ is in effect a weighted sum of future actual prices which were, of course, not known to investors at time t . Thus, the ex post "mistakes" in forecasts by the market at time t are "corrected" when the subsequent "right" prices (which were already contained in $p^*(t)$) are revealed. The strength of this apparent reversion to trend is further accentuated by using the ex post or in-sample trend of stock prices to detrend both the actual stock price and the $p^*(t)$ time series.

In his latest published remarks on the plots of these time series, Shiller (1983, p. 237) concludes:

The near-total lack of correspondence, except for trend, between the aggregate stock price and its ex post rational counterpart (as shown in Figure 1 of my 1981a paper) means that essentially no observed movements in aggregate dividends were ever correctly forecast by movements in aggregate stock prices!

This conclusion does not conform to the empirical facts as shown by the fit of our equation (16) and the additional regression studies in Section IV. That is, the single variable which appears to provide, by far, the most significant forecasting power of the subsequent year's change in aggregate dividends is the previous year's unanticipated change in aggregate stock price.

To briefly analyze the findings of Shiller's innovation test within the context of our model, we once again use the polar case where $d(t) = \rho p(t)$. If, for analytical convenience, it is assumed that price innovations are measured continuously during the sample period $[0, T]$, then from (19) and Ito's Lemma, we have that:

$$\text{Var}(\delta p) = \frac{1}{T} \int_0^T \sigma^2 p^2(t) dt \quad (47)$$

and from $d(t) = \rho p(t)$, that

$$\text{Var}(d) = \frac{1}{T} \int_0^T \rho^2 [p(t) - \bar{p}]^2 dt \quad (48)$$

where $\bar{p} \equiv [\int_0^T p(t) dt]/T$. From (47) and (48), it follows that

$$\epsilon_0 [\text{Var}(\delta p)] = p^2(0) [e^{\sigma^2 T} - 1]/T \quad (49)$$

and

$$\epsilon_0 [\text{Var}(d)] = \rho^2 p^2(0) [e^{\sigma^2 T} - \sigma^2 T - 1]/\sigma^2 T \quad (50)$$

From (49) and (50), it follows that

$$\frac{\epsilon_0 [\text{Var}(\delta p)]}{\epsilon_0 [\text{Var}(d)]} \geq \frac{\sigma^2}{\rho^2} \quad (51)$$

with equality holding in the limit as $T \rightarrow \infty$.³⁶ Using the $\sigma = 0.176$ and $\rho = 0.048$ values reported by Shiller for his 1871-1979 S&P data set, we have from (51) that $\sqrt{\text{Var}(\delta p)/\text{Var}(d)} \geq 3.67$. If managers smooth dividends, then $\text{Var}(d)$ will be smaller than is given in (51) but $\text{Var}(\delta p)$ will not be reduced by this endogenous reduction in the variance of dividends. Hence, our model would predict that $\sqrt{\text{Var}(\delta p)} \gg 3.67 \sqrt{\text{Var}(d)}$ is not an unlikely finding. In contrast, under its posited stationarity assumption, the Shiller model would predict $\sqrt{\text{Var}(\delta p)} \leq 3.19 \sqrt{\text{Var}(d)}$.

In summary, the Shiller model predicts that $\sqrt{\text{Var}(p)/\text{Var}(p^*)}$ will be less than or equal to one while our model predicts the reverse inequality. As

reported in his Table 2 (1981a, p. 431), the sample value of this ratio is 5.59 for his 1871-1979 data set (and 13.28 for his 1928-1979 data set). For the 109-year sample period, the Shiller model predicts that $\sqrt{\text{Var}(\delta p)/\text{Var}(d)} \leq 3.19$ and our model predicts that it will be greater than 3.67. The reported sample value of this ratio is 17.27. While these empirical results clearly do not reject any of the assumptions of our model, they just as clearly do reject (at least) one of the assumptions of the Shiller model. That is, the test results reject either that model's assumption of rationally-determined stock prices or the assumption that dividends follow an exogeneous stationary process with a deterministic trend. If, as Shiller apparently does, one were to reject the former, then one must not only conclude that the market is inefficient--but that it is grossly inefficient--which would seem to require that efficient market (and related rational expectations) models be "scrapped" as reasonable approximations to real-world behavior. If, as we do, one instead interprets the data as rejecting the stationarity assumption, then one need only scrap an assumed process for dividends and rational stock prices for which there was little a priori theoretical or empirical evidence anyway.

In his previously-cited 1983 reply to Copeland, Shiller (p. 237) proclaims:

The challenge for advocates of the efficient markets model is to tell a convincing story which is consistent both with observed trendiness of dividends for a century and with the high volatility of stock prices. They can certainly tell a story which is within the realm of possibility, but it is hard to see how they could come up with inspiring evidence for the model.

We believe that the theoretical and empirical analysis presented here provides such "inspiring evidence."

We are well aware of the LeRoy and Porter (1981) study which claims to show that stock prices are "too volatile" by applying a Shiller-type methodology to accounting earnings instead of dividends. Although not as complete as their control over dividends, managers do have considerable control over accounting earnings as we have already noted.³⁷ We therefore conjecture that an analysis similar to the one presented here will also explain the LeRoy and Porter findings. Indeed, preliminary results from just such a study currently underway appear to confirm this conjecture.

Although lacking the quantitative precision of the Shiller and LeRoy-Porter results, economists have long known that fluctuations in stock prices are considerably larger than the fluctuations in aggregate consumption, national income, the money supply, or other similar variables whose expected values are presumably important to the rational determination of stock prices. Indeed, we suspect that the sympathetic view held by many economists toward the proposition of excess stock market volatility can, in large part, be traced to this long-established observation. However, the time series of the economic flow variables underlying this observation have in common that they are likely to be "endogeneously smoothed" either by the behavior of the economic agents that control them or by the statistical methods which are used to measure them. Our analysis therefore casts doubt over the use of such volatility comparisons in general, as a methodology to test stock market efficiency.

In keeping clear the distinction between "not rejecting" and "accepting" the efficient market theory, Paul Samuelson said it well when he addressed the practicing investment managers of the financial community almost a decade ago (1974):

Indeed, to reveal by bias, the ball is in the court of the practical men: it is the turn of the Mountain to take a first step toward the theoretical Mohammed...If you oversimplify the debate, it can be put in the form of the question,

Resolved, that the best of money managers cannot be demonstrated to be able to deliver the goods of superior portfolio-selection performance.

Any jury that reviews the evidence, and there is a great deal of relevant evidence, must at least come out with the Scottish verdict:

Superior Investment performance is unproved.

Just so, our evidence does not prove that the market is efficient, but it does at least warrant the Scottish verdict:

Excess stock price volatility is unproved.

The ball is once again in the court of those who doubt the Efficient Market Hypothesis.

FOOTNOTES

**We are pleased to acknowledge financial support from the First Atlanta Corporation for computer services. We thank G. Gennotte, S. Myers, and R. Ruback for their helpful comments, and J. Hausman for his advice on the econometric issues. We are especially grateful to F. Black, both for his initial suggestion to explore this topic, and for sharing with us his deep insights into the problem. Without either, the paper would never have been written.

1. There is, however, strong empirical evidence that the variance rate for stock returns is not constant over time, (e.g., Rosenberg (1972) and Merton (1980)), and hence, the strict random walk version of the Hypothesis is probably not valid.
2. While the prices of stocks of individual firms may become zero, we rule out the Armageddon case in which the prices of stocks in aggregate are zero.
3. For discussions of Ito $\hat{\sigma}$ type processes in financial economics, see, for example, Merton (1971; 1975; 1982).

4. There are other restrictions on dividend payouts such as those imposed in loan indentures. Because these restrictions are typically specified in terms of current and past (i.e., retained) accounting earnings, they provide a rather modest constraint on dividend policy. As noted in Shiller (1981a, p. 429), "Earnings...are statistics conceived by accountants...and there is a great deal of latitude for the definition of earnings... ." We agree. Indeed, while there are a wide range of views in the literature about the extent of control that managers have over economic earnings, there is almost uniform agreement that managers have considerable flexibility in controlling earnings within their statutory accounting definition. Moreover, empirical evidence suggests that managers make use of this flexibility by attempting to smooth the time path of accounting earnings. It would appear, however, that management's control over dividend policy exceeds even that of accounting earnings because in effect, the "rules" covering the time pattern of such distributions are "conceived" by the managers themselves.

5. There are, of course, Shiller's studies of the aggregate dividend series. Other such studies appear to be rare. We found only one: Darling (1955) referenced in Brealey's (1971) chapter on dividends.

6. Although $V(t)$ is the present value per share of the future cash flows available for distribution to the shares outstanding at time t , it does not follow that the dividend per share paid at time s must equal $[\Pi(s) - I(s)]/N(s)$. By the accounting identity, $\Pi(s) - I(s) \equiv N(s)D(s) - [N(s+1) - N(s)]P(s)$ where $[N(s+1) - N(s)]P(s)$ is the cash flow received from the issue of new shares of stock at time s , and therefore, $D(s) = [\Pi(s) - I(s) + (N(s+1) - N(s))P(s)]/N(s)$. If issues or purchases of shares are made at "fair" market prices, then such future transactions have a zero net present value, and therefore, have no effect on the current intrinsic value of the firm. If the firm has debt in its capital structure, then interest payments must be subtracted and net proceeds of new debt issues added to the cash flows of the firm. As with stock issues, if the debt is issued or retired at fair market prices, then such future debt transactions will also have no effect on the current intrinsic value of the firm. Although the additional future cash flows from new share and debt issues do not affect current permanent earnings, for a given value of permanent earnings, such financial transactions provide management with considerable flexibility to control the time path of dividends per share.
7. Tests of heteroscedasticity such as White's (1980) on our same data set support this contention.

8. Strictly, feasibility only requires the weaker "less than or equal to." If, however, dividends include all distributions to stockholders and if managers do not throw cash away, then strict equality is required. In contrast to the actual dividend payments made, the term "dividend policy" refers to the contingent schedule or plan for future dividend payments. A dividend policy is, thus, much like the state-contingent functions for optimal control variables which are derived from the solution of a stochastic dynamic programming problem.

9. Indeed, even in the restrictive context of our simple behavioral model, any values for λ and γ in (10) such that $0 \leq \lambda \leq 1$ and $\gamma > 0$ are more than sufficient to ensure satisfaction of the rationality constraint.

10. Even the strongest supporters of the view that "dividend policy matters" would agree that the only effect of a change in dividend policy on investment policy is through its effect on the firm's cost of capital, α . Although a change in dividend policy may "signal" a change in investment policy, one could hardly argue that such a dividend policy change "caused" the subsequent change in investment policy that it signaled.

11. See, for example, Hall (1978). We note further that if consumer behavior is to smooth the time path of consumption, then the dynamics for a change in next period's consumption in response to an unanticipated change in this period's wealth may well be described by a partial adjustment process analogous to our equation (10). If this is the case, then the variation in the consumption time path would be expected to be smaller than the variation in the time path of wealth.

12. Note that the degree of market efficiency posited here is much weaker than would be implied by assuming that the market information contains all the relevant information contained in managers' aggregated information sets. Under our assumption, a manager may have information relevant to the estimation of his firm's intrinsic value that is not available to the market. If, as would seem reasonable, such nonpublic information is firm-specific, then differences between the market's and the manager's assessment of the individual firm's intrinsic value that arise from this source, are likely to (statistically) disappear when these individual assessments are averaged over all firms. It is, of course, possible that the market's information set is richer than the individual manager's, even with respect to estimates of his own firm's intrinsic value. However, rationally-behaving managers would presumably take this possibility into account when making their dividend decisions.

13. As discussed in footnote 6, because of transactions by the firm in its own liabilities, it is not the case that $[\Pi(s) - I(s)]/N(s) \equiv D(s)$. Even without such transactions, managers can still implement virtually any change in dividends per share by the purchase or sale of financial assets held by the firm or by marginal changes in the amount of investment in any other "zero net present value" asset (e.g., inventories). While these latter transactions will change the time pattern of $\Pi(s) - I(s)$, they will not affect the present value of these future cash flows, and therefore, will not affect the current level of permanent earnings.

14. We have also fitted (14') using quarterly data. Although it might at first appear that the use of quarterly rather than annual data would quadruple the number of observations available, there are good reasons for doubting this. There is a distinct yearly (and half-yearly) seasonal in real quarterly dividends. If as this suggests, managers wait until the fourth (fiscal) quarter to "take a look at the year's performance" before deciding to raise or lower that year's dividend relative to the previous year's, then the last quarter's dividend contains effectively the same information as the annual dividend. Further, any "seasonal adjustment" of quarterly dividends not only runs the risk of smoothing away the very innovations in dividends in which we are interested, but also is doubly hazardous when autocorrelated disturbances or lagged dependent variables might be present as in (14'). We therefore report only the results for annual data.

15. As is well known, the log of the expectation of a random variable is greater than the expectation of the log of that random variable. For an Ito Process, the (instantaneous) difference between the two is equal to one-half of the variance rate. Thus, the constant term using the log changes will be equal to $a_0 - (1 - \lambda)\sigma^2/2$.

16. Any of the above-mentioned ways in which the adjustment lags could arise are also potential explanations for disturbance autocorrelation, because in dynamic regression models, lag structure and disturbance autocorrelation can act as proxies for each other.

17. Our R^2 is below the 85% figure given in the original Lintner (1956) study which we reported in Section II. However, Lintner (1956), who stressed that his results were preliminary, pooled his time series observations from 1918-1941 for his 28 companies to estimate his model. Fama and Babiak (1968), who reestimated Lintner's model separately for each firm over the years 1946-1964, report (e.g., in their Table 2) average R^2 figures of (roughly) 40%-45% which are comparable with ours.

18. The point estimate of the dividend yield coefficient is sensitive to how the autocorrelation is taken into account. The half-life calculation in the text uses the GLS rather than the OLS estimate because the positive autocorrelation in the OLS residuals reduce the absolute magnitude of the (negative) dividend yield coefficient. Dynamic regression models with autocorrelated disturbances cannot be easily distinguished from ones with lagged dependent variables, and the dividend yield will be correlated with the lagged dependent variable if, as both our model and the Shiller model posit, the dividend-to-price ratio is autoregressive. There is evidence that the residual autocorrelation estimate is somewhat sensitive to an outlier in 1951, but this outlier apparently has no effect on the estimates of the coefficients in our model. We therefore omit a more detailed analysis of the disturbance autocorrelation.

19. In their classic events study of stock splits and the cash dividend changes which often accompany them, Fama, Fisher, Jensen, and Roll (1969) report a result similar to our's: stock splits, and the increased dividends which typically accompany them, were on average preceded by abnormal price increases. In their study, some of the average "run-up" in prices took place earlier than twelve months before the stock split event, but FFJR deliberately select the individual companies which, ex post, split their stocks. The early small run-up in prices could easily wash out in our aggregate data, and in any case, the FFJR study does not provide information on changes in cash dividends other than those associated with the stock split.

20. See Merton (1975, Appendix B). The proof depends on the behavior of the dividend-to-price ratio in the asymptotic regions where P is either very large or very small. As a strict mathematical condition, it is not necessary that D/P be monotonically increasing in these regions for a nondegenerate steady-state distribution to exist. However, while "local" violations of monotonicity are permitted, the overall "trend" of the function in those regions must be positive. We see no theoretical or empirical foundation for assuming that D/P behaves in this way. It is possible that D/P is, in fact, an increasing function of P in these regions, but that the values of P in the 1926-1981 sample period were not observations from these regions. If so, then, of course, price will have a asymptotic stationary distribution. However, in that case, it is readily apparent that it may take hundreds or perhaps thousands of years to observe it.
21. For proofs and further discussion, see Myers and Turnbull (1977) and Fama (1977).

22. For the period 1926-1981, Ibbotson and Sinquefeld (1982) report that the average annual real return on stocks in excess of the return on Treasury bills is 8.3 percent. About half of this return comes from dividends and half from price appreciation. For the 1871-1979 period, Shiller (1981a, p. 430) reports an annual real return from dividends of 4.8 percent and an estimated real growth in stock price of 1.5 percent which aggregates to an estimated total return of 6.3 percent. Shiller, however, uses the log of price to estimate the growth rate. As discussed in footnote 15, the logarithmic mean underestimates the arithmetic mean by about one half of the variance rate. For a stock return standard deviation of .20, the magnitude of the understatement of the expected total return is therefore about 2 percent. Since it is the arithmetic expected return, and not the geometric one, which is relevant to the substantive discussion, Shiller's "corrected" figure for the total return is about 8.3 percent. Moreover, with this correction, the mix between the dividend and price appreciation components of the total return is approximately 60-40 which is more in line with the cited Ibbotson and Sinquefeld results.

23. In previous estimations, we used a "one-step" GLS procedure. Because Maddala (1971) has shown that "iteration pays" in GLS estimation when a lagged dependent variable is present, we use the iterative approach for the equations in this section.

24. Comparing the OLS and GLS R^2 s provides only a heuristic measure of the incremental explanatory power afforded by the GLS regression, because the OLS R^2 is not a proper benchmark in light of autocorrelation in the residuals. The R^2 for the GLS regression, which, in this case and all others in the paper, we compute as (geometrically) the square of the cosine of the angle between the (centered) dependent variable and the (centered) fitted dependent variable, is also well known not to be uniquely defined for GLS and nonlinear regression models. However, we believe our statements in this paper concerning model fit are not sensitive to our R^2 measure, especially since the OLS and GLS fits of our model are essentially the same. Further, it is hard to think of a more "natural" way of generally measuring the tightness between the fitted and actual dependent variables.
25. As was the case for the OLS and GLS fits (15) and (16) of our model, the GLS fit of (25) does not significantly improve upon its OLS fit. The F statistic for the autoregressive correction in (25) is 2.85, which is not significant at the conventional 5 percent level. Thus, the addition of log price change to (24) substantially increases its explanatory power and improves the autoregressive model's specification (23) by eliminating the requirement that it be supplemented by more structure on the stochastic process for the "unknown" variables before it is a proper regression equation.
26. This regression is almost the equivalent of the Granger-Sims causality test referred to in the causality discussion in Section II.

27. It is all the more surprising because the model does not use contemporaneous price changes. Because all the variables in our model are lagged, equation (16) is a "true" forecast equation in the sense that at time t , it provides an unconditional forecast $D(t + 1)$. The relatively high \bar{R}^2 suggests that aggregate dividends may be forecasted rather successfully.
28. There is, of course, ample precedent in the economics literature for assuming that relative values such as the dividend-to-price and earnings-to-price ratios have steady-state or stationary distributions. As exemplified by our model, the existence of steady-state distributions for such relative values surely does not justify the assumption of stationarity distributions for the levels (or absolute values) of dividends, earnings, or prices.
29. As a brief sampling: Keynes (1936) speculated that, in investment decisions "...reasonable calculation is supplemented by animal spirits," while the presumably more calculating professional money managers engage in "...a game of Snap, of Old Maid, of Musical Chairs." Harberler (1964) refers to "excessive speculation on the stock exchanges" as a factor "in the financial sphere" which contributes to business cycles. In a recent Wall Street Journal editorial page article, Smith (1982) contends that "...stocks...rise and fall even more than the consensus explanations of security analysts, portfolio managers, and newspaper columnists."

30. (1981a, p. 424).
31. In presenting our earlier results, we defined $p(t)$ as the price at the end of year t , which is a slightly different timing convention from that used by Shiller, who defines price at the beginning of the year. To avoid confusion, we continue with our timing convention in the following analysis, where the beginning of sample price and dividend are denoted as being at the end of "year 0."
32. Shiller uses "large enough" in a "double sense." T must be long enough to permit the included dividends in each $p^*(t)$ to capture most of the value of subsequent dividends. It must also be long enough to provide a large number of $p^*(t)$ which are "independent" so as to avoid the necessity of deriving "small sample" statistics for his estimates. Flavin (1982) and Kleidon (1983) explore some of these sampling issues.
33. Our special thanks to G. Gennotte for pointing out this simplification of the proof.
34. With this condition on the weights, the same method of proof can also be used to show that $\text{Var}(p^*) \leq \text{Var}(p)$ if dividends are a weighted sum of the sample period's prices. If dividends are such a weighted sum of past prices, then, of course, some dividends in the sample will depend on some prices outside the sample period. If, however, the sample period is reasonably long, then the dependency of the "early" dividends on out-of-sample prices should have little effect on the estimate of $\text{Var}(p^*)$.

35. If the uncertainty about future permanent earnings is reduced either because managers change investment policy toward "safer" projects or because exogeneous economic changes make the future cash flows of existing projects more predictable, then the reduced variance in future dividends will be accompanied by a corresponding reduction in the variance of stock price. We distinguish such exogeneous reductions in the variance of dividends from the endogeneous reduction caused by managers smoothing of the dividend path for a given volatility of permanent earnings. This latter type reduction will not cause stock price to exhibit less volatility unless investors are "fooled" by managers' smoothing activities. By analogy with the lifecycle model described in Section II, if a consumer chooses to smooth the time path of his consumption, then it does not follow that this endogeneous reduction in the variance of consumption will cause his permanent income or wealth to exhibit less variability.
36. Actually, as $T \rightarrow \infty$, almost certainly, the ratio of the sample paths converges (from below) to the stated population limit.
37. See footnotes 4 and 13. It is also well known that managers exercise this control in an attempt to smooth the time path of earnings--a task made easier by generally accepted accounting rules and conventions which filter the changes in permanent economic earnings.

REFERENCES

- Bhattacharya, S, 1979, "Imperfect Information, Dividend Policy, and 'The Bird in the Hand' Fallacy," Bell Journal of Economics, Spring, 259-270.
- Black, F., and M.S. Scholes, 1974, "The Effects of Dividend Yield and Dividend Policy on Common Stock Prices and Returns," Journal of Financial Economics, 1 (May), 1-22.
- Brealey, R.A., 1971, Security Prices in a Competitive Market: More About Risk and Return From Common Stocks, MIT Press, Cambridge, MA.
- Darling, P.G., 1955, "A Surrogate Measure of Business Confidence and Its Relation to Stock Prices," Journal of Finance, 10 (December), 442-458.
- Davidson, R., and J.G. Mackinnon, 1981, "Several Tests for Model Specification in the Presence of Alternative Hypothesis," Econometrica, 49(3) (May), 781-793.
- Dickey, D.A., and W.A. Fuller, 1979, "Distribution of the Estimators for Autoregressive Time Series With a Unit Root," Journal of the American Statistical Association, 74(366) (June), 427-431.
- _____, 1981, "Likelihood Ratio Statistics for Autoregressive Time Series With a Unit Root," Econometrica, 49(4) (July), 1057-1072.
- Fama, E.F., 1977, "Risk-Adjusted Discount Rates and Capital Budgeting Under Uncertainty," Journal of Financial Economics, 5, 3-24.
- _____, and H. Babiak, 1968, "Dividend Policy: An Empirical Analysis," Journal of the American Statistical Association, 63 (December), 1132-1161.
- _____, L. Fisher, M. Jensen, and R. Roll, 1969, "The Adjustment of Stock Prices to New Information," International Economic Review, 10 (February), 1-21.
- Flavin, M., 1982, "Excess Volatility in the Financial Markets: A Reassessment of the Empirical Evidence," Unpublished Paper, University of Virginia, Economics Department, (January).
- Gordon, M.J., 1959, "Dividends, Earnings and Stock Prices," Review of Economics and Statistics, 41 (May), 99-105.
- _____, 1962, "The Savings, Investment and Valuation of a Corporation," Review of Economics and Statistics, (February), 37-51.
- Grossman, S.J., and R.J. Shiller, 1981, "The Determinants of the Variability of Stock Market Prices," The American Economic Review, 71(2), Papers and Proceedings (May), 222-227.

- Hall, R.E., 1978, "Stochastic Implications of the Life Cycle--Permanent Income Hypothesis: Theory and Evidence," Journal of Political Economy, 86(6), 971-987.
- Harberler, G., 1964, Prosperity and Depression, London, George Allen and Unwin Ltd., 5th ed.
- Hess, P.J., 1983, "Tests for Tax-Effects in the Pricing of Financial Assets," Journal of Business, Forthcoming.
- Ibbotson, R.G., and R.A. Sinquefeld, 1982, Stocks, Bonds, Bills, and Inflation: The Past and The Future, The Financial Analysts Research Foundation, Charlottesville, VA.
- Keynes, J.M., 1936, The General Theory of Employment Interest and Money, Harcourt, Brance & Co., New York.
- Kleidon, A.W., 1983, "Variance Bounds Tests and Stock Price Valuation Models," Unpublished Working Paper, Graduate School of Business, Stanford University, (January).
- Lintner, J., 1956, "Distribution of Incomes of Corporations Among Dividends, Retained Earnings, and Taxes," The American Economic Review, XLVI(2) (May), 97-113.
- _____, 1962, "Dividends, Earnings, Leverage, Stock Prices and The Supply of Capital to Corporations," Review of Economics and Statistics, XLIV(3) (August), 243-269.
- Litzenberger, R.H., and K. Ramaswamy, 1979, "The Effect of Personal Taxes and Dividends on Capital Asset Prices," Journal of Financial Economics, 7, 163-195.
- Maddala, G.S., 1971, "Generalized Least Squares With An Estimated Variance Covariance Matrix," Econometrica, 39(1), (January), 23-33.
- Malinvaud, E., 1966, Statistical Methods of Econometrics, Amsterdam: North-Holland Publishing Company.
- Merton, R.C., 1971, "Optimum Consumption and Portfolio Rules in a Continuous Time Model," Journal of Economic Theory, 3, (December), 373-413.
- _____, 1975, "An Asymptotic Theory of Growth Under Uncertainty," The Review of Economic Studies, 62(3), July, 375-393.
- Merton, R.C., 1980, "On Estimating the Expected Return on the Market." Journal of Financial Economics, 8, 323-361.

- _____, 1982, "On the Mathematics and Economic Assumptions of Continuous Time Models," W.F. Sharpe, Ed., Financial Economics: Essays in Honor of Paul Cootner, North-Holland, Amsterdam.
- Miller, M.H., and F. Modigliani, 1961, "Dividend Policy, Growth and the Valuation of Shares," Journal of Business, October, 411-433.
- _____, and M.S. Scholes, 1982, "Dividends and Taxes: Some Empirical Evidence," Journal of Political Economy, 90, 1118-1142.
- Myers, S.C., and S.M. Turnbull, 1977, "Capital Budgeting and The Capital Asset Pricing Model: Good News and Bad News," Journal of Finance, 32(2), 321-333.
- Petit, R.R., 1972, "Dividend Announcements, Security Performance, and Capital Market Efficiency," Journal of Finance, December, 993-1007.
- Rosenberg, B., 1972, "The Behavior of Random Variables With Nonstationary Variance and the Distribution of Security Prices," Working Paper, University of California at Berkeley.
- Ross, S.A., 1977, "The Determination of Financial Structure: The Incentive-Signalling Approach," Bell Journal of Economics, Spring, 23-40.
- Samuelson, P.A., 1965, "Proof That Properly Anticipated Prices Fluctuate Randomly," Industrial Management Review, 6, Spring, 41-49.
- _____, 1974, "Challenge to Judgment," Journal of Portfolio Management, 1, pp. 17-19.
- _____, 1976, "Optimality of Sluggish Predictors Under Ergodic Probabilities," International Economic Review, 17(1), February, 1-7.
- Shiller, Robert J., 1981a, "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?" The American Economic Review, 71(3), June, 421-436.
- _____, 1981b, "The Use of Volatility Measures in Assessing Market Efficiency," The Journal of Finance, XXXVI(2), May, 291-311.
- _____, 1982, "Consumption, Asset Markets, and Macroeconomic Fluctuations," Carnegie-Rochester Conference Series on Public Policy, 17, 203-238.
- _____, 1983, "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?: Reply," The American Economic Review, 73(1), 236-237.
- Smith, A., 1982, "Stock Values: Notes on the Lemming Syndrome," The Wall Street Journal, Monday, March 26, p. 26.

- Telser, L.G., 1967, "Discrete Samples and Moving Sums in Stationary Stochastic Processes," Journal of the American Statistical Association, 62(318), June, 484-499.
- Tukey, J.W., 1970, Explanatory Data Analysis, Addison-Wesley, Reading, MA.
- Watts, R., 1973, "The Information Content of Dividends," The Journal of Business, April, 191-211.
- White, R., 1980, "A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity," Econometrica, 48(4), May, 817-838.
- Zellner, A., 1979, "Causality and Econometrics," in K. Brunner and A.H. Meltzer, eds., Three Aspects of Policy and Policy Making, Carnegie-Rochester Conference Series, Volume 10, Amsterdam, North-Holland Publishing Co.

4221, 008^{LC}

MIT LIBRARIES



3 9080 004 512 122

Date Due

APR 05 '88

NOV 30 1990

Lib-26-67

Bar Code On
Last Page

