Terry A. Marsh  
*University of California, Berkeley, and Stanford University*

Robert C. Merton  
*Massachusetts Institute of Technology*

**Dividend Behavior for the Aggregate Stock Market***

I. Introduction

In this paper, we develop a model of the dividend process for the aggregate stock market. Previous research has focused almost exclusively on dividend behavior at the micro level of the individual firm. Hence, to motivate the focus here on aggregate-dividend behavior, we begin with a brief review of these earlier micro studies, which is followed by a discussion that locates the place of our aggregate analysis within this body of research. In Sections II–V, we derive and fit our econometric model of the dividend process. In Section VI, we compare the performance of the model with other models in the literature.

Although long a staple of financial management textbooks, corporate-dividend policy remains a topic on which the field has failed to arrive at even a local sense of closure.¹ Black

---

* We dedicate this paper to the scientific contributions and the memory of John V. Lintner, Jr. We are grateful to J. Hausman, M. Miller, S. Myers, R. Ruback, and especially F. Black for helpful comments and to the Institute for Quantitative Research in Finance for partial funding. The first author is also grateful to the Battymarch Fellowship Program, under which portions of the work presented here were completed.

¹ For a more complete summary of the dividend controversy, see Brealey and Myers (1984, ch. 16). The degree to which this controversy is unresolved is exemplified by its inclusion in Brealey and Myers’s list of 10 important unresolved problems in finance that “seem ripe for productive research” (p. 790).


© 1987 by The University of Chicago. All rights reserved.

0021-9398/87/6001-0002$01.50
(1976) has aptly described this lack of closure as the "dividend puzzle." The pivotal point in this puzzle is the classic work of Miller and Modigliani (1961), which demonstrated the irrelevance of dividend policy for determining the firm's cost of capital.

Miller and Modigliani showed that, when investors can create any payout pattern they want by selling and purchasing shares, the expected return required to induce them to hold these shares will be invariant to the way in which firms "package" gross-dividend payments and new issues of stock (and/or other zero-net-present-value transactions). Since neither the firm's expected future net cash flows nor its discount rate is affected by the choice of dividend policy per se, its current market value cannot be changed by a change in that policy. Thus dividend policy "does not matter." Although, under Miller and Modigliani's proposition, there are no a priori reasons for firms to follow any systematic dividend policy, there are also no penalties if they choose to do so.

Exceptions to Miller and Modigliani's 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 they provide some early 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). Litzenberger and Ramaswamy (1979) and Poterba and Summers (1984, 1985) offer empirical support for this view. On the other hand, Black and Scholes (1974), Miller and Scholes (1978, 1982), Hess (1983), and Eades, Hess, and Kim (1984) present analysis and evidence suggesting that, as an empirical matter, tax effects per se do not appear to affect the cost of capital. Easterbrook (1984) proposes an agency-theory explanation for dividends. Along different lines from these studies, Shefrin and Statman (1984) have used behavioral theories of individual choice to argue that investors will prefer cash dividends even if they are tax disadvantaged. Although some of these analyses might provide reasons to believe that investors are not indifferent between cash dividends and capital gains, the empirical evidence to date is still inconclusive for rejecting Miller and Modigliani's proposition.

Even with their view of investor indifference for dividends, Miller and Modigliani (1961, p. 431) do point out that dividend policy can matter if dividend changes are used by firms to convey information not otherwise known to the market. Bhattacharya (1979), Miller and Rock (1985), and John and Williams (1984) use a signaling-model approach to formalize this notion. Hakansson (1982) derives the additional general equilibrium conditions required for dividend signals to improve investor welfare. Aharony and Swary (1980) provide some empirical evi-
vidence on the informational content of dividends in their study of dividend-announcement events for 149 NYSE industrials, in which they find that, on the average, unexpected dividend and unexpected price changes are positively correlated around announcement dates. Asquith and Mullins (1983) find similar results for firms that initiate dividend payments for the first time. The evidence is, however, that the correlations, while statistically significant, are rather small, a conclusion also reached in empirical work by Watts (1973) and Gonendes (1978).

In summary, there are a number of conflicting theories of dividend behavior, and the empirical studies to date provide little compelling evidence for one over the others. The management of a firm is free to choose a dividend policy with virtually any time pattern it wants, subject only to the overall constraint that the present value of expected future distributions net of new stock offerings cannot exceed the present value of the firm’s expected net cash flows generated by its investments. Indeed, except for certain debt-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 managers of publicly traded and widely held corporations to follow any particular dividend policy.

With so much controversy surrounding the various normative theories of the dividend process, it is perhaps not surprising that empirical researchers have relied heavily on positive theories of dividend behavior to specify their models. The prototype for these models is Lintner’s (1956) model, which is based on stylized facts first established by him in a classic set of interviews with managers about their dividend policies. A similar model, motivated by Friedman’s Permanent Income Hypothesis, is proposed by Fisher (1957). From Lintner’s interviews, it was readily apparent that dividend policies across firms were hardly uniform. Lintner did, however, identify some common characteristics: namely, managers tend to change dividends primarily in response to an unanticipated and nontransitory change in their firm’s earnings, and they are guided by target payout ratios in making those changes. Using an econometric model based on these perceived patterns, Lintner found that he could explain a significant portion of annual dividend changes for a sample of companies over the period 1918–41. Using similar types of models, subsequent empirical work by Fama and Babiak (1968), Petit (1972), and Watts (1973) supports Lintner’s original findings.

With few exceptions (notably, Brittain [1966] and Shiller [1981a, 1981b]), research on both normative and positive models of dividends has focused on the micro behavior of individual firms. The relative lack of research on aggregate-dividend behavior is perhaps not surprising since many of the more interesting issues surrounding dividend policy are likely to be firm specific. For example, clientele effects and inden-
ture restrictions that could in principle affect an individual firm’s dividend policy are likely to "wash out" in any aggregate-dividend analysis. Similarly, issues involving the informational content of an individual firm’s dividends are 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 meaningful announcement dates for aggregate dividends to perform event studies along the lines of those by Gonedes (1978) and Aharony and Swary (1980).

If firms only changed their dividends to signal information, and if the only information worth signaling is specific to the firm, then changes in aggregate dividends would be random and their magnitudes small. If, however, firms change their dividends for reasons other than signaling, then the very fact that aggregate-dividend changes are unlikely to contain much signaling information may make them an especially useful series for measuring the informational content of an individual firm’s dividend announcements. To identify the signals or abnormal changes in a firm’s dividends, it is, of course, necessary to have a model of its "normal" dividend behavior. Watts (1973), for example, uses Lintner’s model for this purpose. Lintner’s model does not, however, take account of the cross-sectional dependencies among firms’ dividend policies. It is reasonable to expect that, in addition to its own economic circumstances, the firm would use the dividend behavior of other firms to calibrate its dividend policy—as, for example, observing industry practice in the selection of its target payout ratio. Moreover, these dependencies may be of considerable empirical significance in the light of the already-documented strong correlations among different firms’ contemporaneous stock-price changes. Removal of the aggregate-market component of a stock’s return to obtain a better estimate of its abnormal price change is commonplace. Just so, use of an aggregate-dividend model to remove the "systematic" component of an individual firm’s dividend policy would appear to provide a better estimate of its abnormal dividend changes.²

If individual firms follow reasonably stable dividend policies over time, then the aforementioned cross-sectional dependencies will induce systematic behavior in the time series of aggregate dividends. It is, however, possible for aggregate dividends to exhibit stable and consistent time-series properties even if no such stability were found for individual firms. For example, in a purely demand-driven model for

² In fact, in their events study of the dividend behavior of split stocks, Fama et al. (1969) did adjust individual firm dividend changes for marketwide dividend changes before classifying the former as increases or decreases. Fama et al. did this to be internally consistent in associating security-return residuals with dividend changes. Note that, as would be the case for our model at the micro level, an added variable might well be required to account for dividend increases or decreases associated with a firm’s "normal" secular progression through its "life cycle."
dividends, the demand for dividends is not firm specific because investors care only about the dividend–capital gain mix at the portfolio level. Hence, as in Miller’s (1977) theory of corporate debt, there will be, in general, many different allocations of dividend policies to individual firms that will support an equilibrium in the dividend market. Thus equilibrium aggregate dividends may be determinate, but which firms service this demand and the quantity each chooses to supply may not.\(^3\)

In the next section, we motivate the specification of our econometric model with a discussion of the descriptive facts established in Lintner’s interviews.\(^4\) This is followed in Sections III, IV, and V by the estimation and testing of the model.

Although sharing Lintner’s stylized facts with the previously cited empirical studies of micro dividend behavior, our model, in addition to being applied at the aggregate level, differs significantly from these earlier studies because it assumes that economic earnings, not accounting earnings, are the primary determinant of dividends. The analysis in Section VI compares the performance of our model with one that uses accounting earnings. Because the Brittain (1966) study of aggregate dividends relies on the relations between dividends and accounting variables, our analysis sheds light on his findings as well. In this same section, we also compare the relative performance of our model with the univariate trend-autoregressive model associated with Shiller’s (1981\(a, 1981b\)) model of aggregate dividends.

II. Model of Aggregate-Dividend Dynamics

Lintner found considerable heterogeneity among firm’s dividend policies in his interviews with corporate managers. However, he also found some characteristics to be common to many of these firms’ dividend policies. These stylized facts are summarized as follows. (i) Managers believe that firms should have some long-term target payout ratio. (ii) In setting dividends, they focus on the change in existing payouts, not on the level. (iii) A major unanticipated and nontransitory

\(^3\) Note, however, that with only slight embellishment to include transactions costs for either issuers or investors even this extreme demand-driven model would predict systematic micro dividend behavior. With issuing costs, e.g., firms in mature or declining industries with large cash flows relative to their new investment needs, would be marginally lower-cost producers of dividends than firms in growth industries. Since the position of a firm or an industry within its “life cycle” is hardly random from year to year, both individual firm and industry payout patterns are likely to exhibit serial dependencies. The empirical evidence for microstability in dividends is explored in Marsh and Merton (1985\(a\)).

\(^4\) The stylized facts distilled by Lintner from his interviews can be interpreted as a description of “average” or “systematic” dividend behavior. In this sense, these facts are macro rather than micro.
change in earnings would be an important reason to change dividends.
(iv) Most managers try to avoid making changes in dividends that stand
a good chance of having to be reversed within the near future.

Most textbook discussions seem to agree with the interpretation of
these stylized facts to the effect that it is changes in some measure of
long-run sustainable or "permanent" earnings, rather than current
earnings, that drive dividend decisions. That is, a change in current-
earnings flow that is viewed by management as essentially transitory
would not be likely to give rise to a noticeable change in dividends.
Unfortunately, except for the special case of a firm whose future earn-
ings are certain and generated without further net new investment, the
textbooks are not specific in defining permanent earnings. Our inter-
pretation (which is consistent with this special case) defines the perma-
nent earnings per share of a firm at time \( t \) as equal to the expectation as
of time \( t \) of that level of uniform payments that 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 cash flow 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 \( s \).
(\( I(s) = \text{gross new physical investment} + \text{purchases of financial assets} - \text{sales of physical and financial assets} \).) If \( \alpha \) denotes the firm’s real cost of
capital, then the discounted value of the expected cash flows avail-
able for distribution to each share outstanding at time \( t \) is given by

\[
V(t) = \epsilon_t \left[ \int_t^\infty [\Pi(s) - I(s)] e^{-\alpha(s-t)} ds \right] / N(t),
\]

where \( \epsilon_t \) denotes the expectation operator, conditional on information
available as of time \( t \), and \( N(t) \) denotes the number of shares outstand-
ing. The variable \( V(t) \) 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.5 That is, if \( E(t) \)

\[
\Pi(s) - I(s) = N(s)D(s) - [N(s+1) - N(s)]P(s),
\]

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

\[
D(s) = ([\Pi(s) - I(s) + [N(s+1) - N(s)]P(s)] / N(s).
\]

5. 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,

\[
\Pi(s) - I(s) = N(s)D(s) - [N(s+1) - N(s)]P(s),
\]

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 per share 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 issued, 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
denotes permanent earnings per share of the firm at time $t$, then

$$E(t) = \alpha V(t).$$ (2)

Since corporate managers set dividends for their firms, it is their assessments of permanent earnings that are relevant for the evolution of aggregate dividends. For this purpose, we denote managers’ determination of permanent earnings by $E''(t) = \alpha V''(t)$, where $V''(t)$ is given by (1) with the expectation operator $\mathbb{E}_t = \mathbb{E}_t''$ based on the probability distribution for future $I(s)$ and $I(s)$ generated by the managers’ information sets as of time $t$.

Although Lintner’s stylized facts suggest that dividend changes are related to permanent-earnings changes, the interview data on which they are based contain little information about the detailed functional form of that relation.\(^6\) In the absence of a specific structural model of that relation, we posit that logarithmic dividends can be expressed as the sum of a rational distributed lag of logarithmic permanent earnings, a drift term that is conditional on information known at time $t$, $\alpha(t)$, and a disturbance term $\eta(t)$. That is, we represent the aggregate dividend process as

$$(1 - \phi_1 L)\log[D(t)] = \alpha(t) + (\lambda - \theta_1 L)\log[E''(t - 1)] + \eta(t),$$ (3)

where $D(t)$ is the integral $\int_{t-1}^t D(s)ds$ of aggregate dividends paid per share of the market portfolio over the interval from time $t - 1$ to time $t$; $E''(t)$ is permanent earnings as defined in (2), per share of the market portfolio at time $t$; and the roots of the first-order polynomials in the lag operator $L$ are outside the unit circle.

As specified, (3) is consistent with the “short-run” dividend dynamics of Lintner’s model. It does not, however, capture his stylized fact i that firms typically set a long-run target for the dividend-payout ratio. In line with the discussion concerning the steady-state properties of long-run equilibrium dividend payout, we take account of this long-run objective in our model by requiring that dividend payouts converge to a

---

6. At least at the time of his survey 30 years ago, Lintner’s (1956) evidence indicated that dividend policy was not viewed by management as simply a balancing item in the flow of funds account: “Dividends [rather than retained earnings and savings] represent the primary and active decision variable in most situations” (p. 97). “In general, management’s standards with respect to its current liquidity position appeared to be very much more flexible than its standards with respect to dividend policy, and this flexibility provided by the buffer between reasonably definite dividend requirements in line with established policy and especially rich current investment opportunities” (p. 105). Other statements could be interpreted as hints regarding the loss function underlying dividend rules, but they are not very specific; e.g., by “stabilizing” dividends, managers can “minimize adverse shareholder reactions,” and “management can live more comfortably with its unavoidable uncertainties regarding future developments” (p. 100).
constant target ratio; that is, as $t \to \infty$ (and in the absence of any disturbances),
\[
\lim_{t \to \infty} \log \left[ \frac{D(t)}{E^m(t - 1)} \right] = \beta. \tag{4}
\]

This special assumption that the long-run target be literally constant is more stringent than necessary.\(^7\) Moreover, this assumption does not, of course, imply that dividends and permanent earnings follow (trend) stationary processes.\(^8\)

As shown in the Appendix, if the long-run steady-state condition (4) is imposed on the short-run dynamics (3), then (3) can be rewritten as
\[
\log[D(t + 1)] - \log[D(t)] = \frac{D(t)}{E^m(t - 1) - m(t - 1)} \tag{5}
\]
where $m(t - 1)$ is the time $t - 1$ expectation of the logarithmic change in permanent earnings, $\{\log[E^m(t)] - \log[E^m(t - 1)] - m(t - 1)\}$ and $g(t)$ is the expected logarithmic change in dividends, $\{\log[D(t)] - \log[D(t - 1)]\}$ if the time $t - 1$ logarithmic payout ratio $\log[D(t)/E^m(t - 1)]$ is equal to its long-run target $\beta$ and the unexpected change in logarithmic permanent earnings, $\{\log[E^m(t)] - \log[E^m(t - 1)] - m(t - 1)\}$, is zero.

The model described by (5) takes the form of the well-known "error-correction" model, which has been studied and applied by Sargan (1964), Davidson et al. (1978), Nickell (1980), Salmon (1982), and Ericsson and Hendry (1984). It obviously satisfies the condition for a long-run steady-state distribution for $D/E^m$ because, if $\eta(t + 1) = 0$, then $\Delta \log[E^m(t)] = m(t - 1)$ implies $\{\log[D(t)] - \log[E^m(t - 1)]\} = \beta$ as required. Given the specification (4) of the long-run equilibrium, the model's potential for describing the short-run dynamics of aggregate dividends depends on the appropriateness of the rational distributed

7. It may be verified that only trivial modifications in the dynamic equation developed below are required to account explicitly for any steady-state variation in the ratio. For example, if it is hypothesized that the equilibrium ratio depends on a vector of stationary stochastic variables $Z(t)$, (4) is replaced by
\[
\log \left[ \frac{D(t)}{E^m(t - 1)} \right] = \beta_0 + \beta'Z(t),
\]
with the only effect being a change in the regressivity term in (5).

8. For example, if the dynamics of the firm's intrinsic value follow a geometric Brownian motion, and if management pays out a constant proportion of permanent earnings as dividends, then the dividend dynamics will also be described by a geometric Brownian motion. No amount of time detrending will make either of these processes stationary. Nevertheless, the ratio of dividends to permanent earnings is (trivially) a stationary process. For further discussion, see Rubinstein (1976, pp. 409–11) and Marsh and Merton (1985b, app. A).
lag in (3). Since the error-correction model is applicable to a wide range of stochastic processes governing \( E^m(t) \), including the geometric random walk (cf. Nickell 1980), the major assumptions imbedded in (3) are those of symmetry in the responsiveness of dividend changes to permanent-earnings changes and, empirically, the constancy of coefficients.  

In economic terms, the "normal" or unconditional growth rate for dividends, \( g(t) \), equals \( \alpha r(t) \), which is the usual expression for the deterministic growth rate of dividends, where \( r(t) = 1 - D(t)/E^m(t - 1) \) is the retention rate (in terms of permanent earnings) at time \( t \), and \( \alpha \) is defined as the aggregate cost of capital. That is, specification of the deterministic component \( g(t) \) 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 (5) describe the deviation of the growth in dividends from this normal rate.

The second term in (5), which is multiplied by \( \lambda \), captures Lintner's stylized fact iii—that managers will change dividends away from the anticipated path in response to an unanticipated change in permanent earnings, \( [\log[E^m(t)/E^m(t - 1)] - m(t - 1)] \). The third term, which is multiplied by \( \gamma \), is the error-correction component that drags short-run dividends toward their long-run steady-state payout ratio, thus capturing Lintner's stylized fact i. The value of \( \gamma \), which should be positive,  

9. The interpretation of the parameter \( \gamma \) in (5) suffers from the absence of a more precise underlying structural model that leads to (3). To see this, suppose for simplicity that \( \Delta \log[E^m(t)] = m(t - 1) \), \( \eta(t) = 0 \), in (5) but that the current-dividend yield is out of equilibrium. Salmon (1982, p. 622) shows that a "proportional, integral, derivative" (PID) control rule,  

\[
\Delta \log[D(t)] = \beta - k_p \Delta [\log[D(t)] - \log[E^m(t - 1)]] - k_i \Delta \log[D(t)] - \log[E^m(t - 1)]
\]

\[
- k_d \Delta^2 [\log[D(t)] - \log[E^m(t - 1)]],
\]

leads to the following error correction term:

\[
\Delta \log[D(t)] = [(k_p + k_t + k_d) - (k_p + 2k_d)\lambda - k_d^2\lambda] [\log[D(t)] - \log[E^m(t - 1)]]
\]

\[
= \beta - A(L) [\log[D(t)] - \log[E^m(t - 1)]].
\]

In the absence of a more explicit model, one can only speculate about the need to include a term like \( \Delta^2 [\log[D(t)] - \log[E^m(t - 1)]] \) in (5). For example, in Salmon's general discussion, he argues that an error-correction rule might be appropriate when a decision maker faces an uncertain environment in which the control problem is itself changing over time. Terms like \( \Delta^2 [\log[D(t)] - \log[E^m(t - 1)]] \) may pick up such changes, especially if the type of model changes that occur result in a non-time-additive "control problem" to be solved for the short-run dynamics of dividends. Insofar as changes in the long-run equilibrium dividend-price ratio are concerned, many can be accommodated in the formulations in the text. Given these formulations, if uncertainties about the "short-run" dividend-control problem cannot be completely described by a linear-quadratic problem, a nonlinear model will generally be needed in place of (5)—the certainty-equivalence principle will no longer hold, and \( \eta(t) \) will not be an additive error.
measures the average speed of convergence of the payout ratio to its target.

The a priori reasons for choosing the lag specification in (3) and (5), in which an unanticipated change in permanent earnings from time \( t - 1 \) to time \( t \) causes a dividend change in the interval \( (t, t + 1) \), are as follows. 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 in response to new information. Second, although firms usually declare dividends once a quarter, many firms make significant changes only at the end of their fiscal year. Third, even if individual firms' managers did react instantaneously, the reaction in aggregate dividends will appear to be lagged because of different announcement dates and different speeds of reaction across firms.

In responding to an unanticipated change in permanent earnings, managers will change dividends in the same direction, which implies that \( \lambda \) in (5) should be positive. From stylized fact iv, managers prefer to avoid reversals in dividends, and it can be established that a partial adjustment policy with \( \lambda < 1 \) is optimal if reversals or changes are costly.

III. The Dividend Model Expressed as a Regression Equation

In the empirical studies of both Lintner's model and subsequent dividend models based on his original formulation, the equations corresponding to our (5) are treated as regression equations. We too assume that equation (5) 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 the next-period dividends, not the reverse.

Of course, in a complete general equilibrium model, dividend changes and intrinsic-value changes, along with other quantities and prices, are jointly endogenous. However, insofar as the bivariate series of dividend changes and intrinsic-value changes is concerned, there are persuasive grounds for treating the latter as a proper predetermined endogenous variable, particularly when the discount rate \( \alpha \) is assumed constant.

As already noted, there are no important legal or accounting constraints on dividend policy; hence managers have almost complete discretion and control over the choice of dividend policy. If, however, managers are not irrational, then they will, at least, choose a dividend policy that is feasible in both the short run and the long. Such 'feasible' 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 (1). Because managers set dividend policy, this constraint is properly specified in terms of their probability assessments. Hence, from (2), it follows that a rational dividend policy must satisfy

\[ e_t^m \left[ \int_t^\infty D(s)e^{-\alpha(s-t)}ds \right] = E^m(t)/\alpha. \] (6)

As discussed in Marsh and Merton (1986), this constraint on dividend choice is very much analogous to the intertemporal budget constraint on consumption choice in the basic lifetime-consumption decision problem for an individual. Like consumers in selecting their planned intertemporal expenditures for a given amount of wealth, managers, facing a given level of permanent earnings, have a great deal of latitude in their choice of dividend policy. The fact that individual firms pursue dividend policies that are vastly different from one another is empirical evidence consistent with this view.

It does not follow from (6) that a change in dividend policy by managers will cause a change in their current assessments of permanent earnings. For a fixed discount rate, \( \alpha \), it does, however, follow from (6) that an unanticipated change in permanent earnings must necessarily cause a change in expected future dividends. 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.

In making the case for causality in equation (5), we are not unaware of the possibility that there are other exogenous variables that may

10. Strictly, feasibility requires only 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.

11. Indeed, even in the restrictive context of our simple behavioral model, any values for \( \lambda \) and \( \gamma \) in (5) such that \( 0 \leq \lambda \leq 1 \) and \( \gamma > 0 \) are more than sufficient to ensure satisfaction of the rationality constraint.

12. 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, \( \alpha \). 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.

13. See, e.g., Hall (1978). We note further that, if consumer behavior is to smooth the time path of changes in consumption, then the dynamics for a change in the next-period consumption in response to an unanticipated change in this-period wealth may well be described by a partial adjustment process analogous to our eq. (5).
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 (5) 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 that 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 exogenous variables that explain the next-period change in dividends, it seems unlikely that they would be correlated with this-period unanticipated change in permanent earnings. Hence the assumption that equation (5) 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 (5) 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 identification problem—that statistical tests alone are not sufficient to establish causality and that, ultimately, this issue can be resolved only by a priori economic reasoning (cf. Zellner 1979).

IV. A Reduced Form for the Dynamic Model

Although we have proposed equation (5) for aggregate-dividend dynamics as both a structural and a regression equation, it cannot be estimated in its current form because management assessments of changes in firms' permanent earnings are not observable. In this section, we add the necessary further specification to estimate the model.

If managers are rational forecasters, and if the market is reasonably efficient, then the market's estimate of a firm's intrinsic value should on the average be equal to the intrinsic-value estimate made by that firm's management. We therefore assume that the discounted value $V^M(t)$ of the expected future aggregate net cash flows of all firms per market share, 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.\(^{14}\) This market efficiency condition can be written as

\(^{14}\) Note that the degree of market efficiency posited here is much weaker than would be implied by assuming that the market-information set 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 or her firm's intrinsic value.
Dividend Behavior for Aggregate Stock Market

\[ V^M(t) = V^m(t), \quad (7) \]

for all \( t \), where \( V^M(t) \) is given by (1), with \( \epsilon_i = \epsilon_i^M \).

We further assume that the stock-market price is equal to its intrinsic value, that is, that there are no speculative bubbles. From this assumption and (7), we can write the cum-dividend price of a share of the market portfolio at time \( t \):

\[ P_c(t) = \epsilon_i^M \left[ \int_t^\infty e^{-\alpha(s-t)}[\Pi(s) - I(s)]ds \right]/N(t). \quad (8) \]

Using the market efficiency condition (7) and the definition of permanent earnings in (2), we can rewrite (8) as

\[ P_c(t) = E^m(t)/\alpha. \quad (9) \]

If, as we have assumed, the expected real rate of return on the market, \( \alpha \), is a positive constant, it follows from (9) that the percentage change in stock-market price is equal to the percentage change in manager's assessment of permanent earnings. Substituting for \( E^m(t) \) from (9) and splitting the cum-dividend stock-price change \( P_c(t)/P_c(t - 1) \) into its two component parts—the ex dividend change \( P(t)/P(t - 1) \) and the dividend yield \( D(t)/P(t - 1) \)—we rewrite (5) as

\[
\log \left[ \frac{D(t + 1)}{D(t)} \right] = \left[ \alpha - \frac{D(t)}{P(t - 1)} \right] + \lambda \left[ \log \left[ \frac{P(t) + D(t)}{P(t - 1)} \right] - \alpha \right] \\
+ \gamma \left[ \rho - \log \left[ \frac{D(t)}{P(t - 1)} \right] \right] + u(t + 1),
\]

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 assessments of the individual firm's intrinsic value that arise from this source are likely to disappear (statistically) 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 or her own firm's intrinsic value. However, rationally behaving managers would, presumably, take this possibility into account when making their dividend decisions.

15. As discussed in n. 5 above, because of transactions by the firm in its own liabilities, it is not the case that \( [\Pi(s) - I(s)]/N(s) = D(s) \) in (8). 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. The presence of such significant nonequity sources of cash flow causes practical difficulties in testing theories that make fine distinctions between the behavior of gross and net dividends.
where $\rho = \beta + \log \alpha$. By rearranging terms, we can rewrite (10) as

$$\log\left[\frac{D(t + 1)}{D(t)}\right] + \frac{D(t)}{P(t - 1)} = a_0 + a_1 \log\left[\frac{P(t) + D(t)}{P(t - 1)}\right] + a_2 \log\left[\frac{D(t)}{P(t - 1)}\right] + u(t + 1),$$

(11)

where $a_1 = \lambda$, $a_2 = -\gamma$, and $a_0 = (1 - \lambda)\alpha + \gamma\rho$.

Note that (9) is not an identity. It is a specification that is valid under the hypothesis that market prices are rational predictors of firms' future net cash flows. Thus (11) is a reduced-form equation and, as such, can be consistent with more than one set of structural hypotheses. If market price provides a good estimate of managers' assessments of permanent earnings, then (11) 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. But it would also be the case if the market is inefficient because it is moved by waves of optimism and pessimism and managers either rely on market prices for their estimates of permanent earnings or are influenced by the same irrational waves as investors in making their assessments of intrinsic values.

To investigate this identification matter further, consider the following model first suggested to us by Zvi Griliches. Managers are rational forecasters of permanent earnings, and they fully adjust dividends in response to changes in these earnings; that is, $\lambda = 1$ in (5) and

$$\log\left[\frac{D(t + 1)}{D(t)}\right] = \log\left[\frac{E^m(t)}{E^m(t - 1)}\right].$$

The market is assumed to be inefficient both because of random "animal spirits" and because investors tend to overreact to new information about the fundamentals. That is, replace (9) with the alternative hypothesis,

$$\log\left[\frac{P_c(t)}{P_c(t - 1)}\right] = \psi \log\left[\frac{E^m(t)}{E^m(t - 1)}\right] + \epsilon(t),$$

where $\psi > 1$. If equation (11) is fit under these assumptions, then the predicted value for $a_1$ is given by $a_1 = \psi/(\psi^2 + q)$, where $q$ is the ratio of the variance of animal-spirits-induced price changes to the variance of rational permanent-earnings changes.

As in our model, the Griliches model predicts that $0 < a_1 < 1$. If market price is a very noisy estimator of permanent earnings principally because of animal spirits (i.e., $q/\psi \gg 1$), then (11) should be a poor predictor of the evolution of dividends. Alternatively, the market could be efficient and (11) could be a poor predictor because management dividend decisions in the aggregate are not well described by behav-
ioral equation (5). If (11) exhibits strong explanatory power, it is still possible for the Griliches model to hold if overreaction is the primary source of market inefficiency (i.e., \( \eta/\psi \ll 1 \)). It would thus appear that the empirical properties of equation (11) are alone inadequate to distinguish between our model and Griliches's alternative.

The focus of our study is not to test the hypothesis of stock-market inefficiency but instead to develop a model of aggregate-dividend behavior that is consistent with that hypothesis.\(^{16}\) Nevertheless, we digress briefly from that focus to point out some ancillary conditions to equation (11) that do provide some discriminatory power between our model and Griliches's.

As already noted, unanticipated changes in rationally forecasted permanent earnings should have no serial dependencies. In Griliches's model, this implies that successive changes in dividends should be uncorrelated. As shown in Section VI, the empirical time series of dividend changes exhibits rather strong positive serial correlation. In contrast, such serial correlations have little effect on the robustness of our model.

If one assumes that even inefficient stock prices cannot wander arbitrarily far from their intrinsic values for an indefinite period of time, then with a constant discount rate the regressivity of price toward intrinsic value will induce negative serial dependency in stock-price changes. If this regression takes place in a series of small adjustments systematically over time, then one would expect to find significant negative serial correlation in the stock-return series. As is well-known, the empirical evidence does not support this prediction.\(^{17}\) If the regression takes place in the form of large adjustments at random and relatively infrequent points in time, then the standard estimates of serial

---

16. The joint hypothesis implied by the assumptions used to derive eq. (11) includes the dividend behavioral equation (5) and a constant real discount rate in addition to stock-market rationality. Thus the goodness of fit of (11) is hardly a meaningful test of stock-market rationality. This is, indeed, the central point in Marsh and Merton (1986), made there with respect to the joint hypothesis of Shiller's (1981a) variance-bound tests and their interpretation as tests of stock-market rationality. If the model of dividend behavior posited here is substituted for the Shiller assumption of a stationary process for dividends, then these variance bounds will be systematically violated even if stock prices are rationally determined. If, as the theoretical and empirical evidence presented here suggests, this model of dividends is a plausible alternative to the one posited by Shiller, then his variance bounds are not a reliable test of stock-market rationality. This is so whether or not stock prices are, in fact, rationally determined. See Merton (in press) for a related discussion of other variance-bound tests for this hypothesis.

17. Although the majority of studies using serial correlation, filtering, and spectral analysis tests supports this view, Summers (1986) shows that their power is low in detecting long-wave serial dependencies in stock returns. Moreover, Fama and French (1986) find evidence of a regressive or temporary component in 3–5 year stock returns. Such findings are consistent with the view that stock prices deviate significantly from fundamental values with a slow speed of adjustment toward these values. They are, however, also consistent with rational stock prices and time-varying equilibrium expected returns.
correlation may not detect this dependency. However, these "outliers" would tend to cause the empirical distribution of stock-price changes to exhibit higher kurtosis than the distribution of rational permanent-earnings changes. In the context of the Griliches model, stock-price changes should have higher kurtosis than dividend changes. As shown in Section VI, quite the opposite seems to be the case as an empirical matter. As discussed there, this finding is consistent with our model. Although hardly a complete analysis, these ancillary findings appear to support our model over Griliches's alternative. With this, we end the digression and return to the development of our model of dividend behavior in an environment with a rational stock market.

For (11) to be a proper reduced-form equation, its right-hand-side variables must be predetermined relative to its left-hand-side variables. As discussed at length in Section III, an unanticipated change in this-period permanent earnings, over which managers have no control, is an exogenous variable relative to the change in the next-period dividends, which managers control almost completely. From (7) and structural equation (9), it therefore follows that an unanticipated change in this-period price is exogenous relative to the next-period dividend change; hence (11) is a proper reduced-form equation. In this limited sense of a reduced form, an unanticipated change in this-period price "causes" a (predictable) change in the next-period dividends.

In specifying (11), our intent is to construct a simple model of the dividend process that nevertheless captures the basic stylized facts of management behavior. We have therefore assumed a simple 1-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 that enter into the process. As we will show, the empirical conclusions derived from this simple model are likely to be robust with respect to refinements that include such additional variables.

V. Model Estimation

To estimate the reduced-form equation (11), we use annual data constructed from monthly dividend and price series for the value-weighted NYSE index contained in the Center for Research in Security Prices (CRSP) data set over the period 1926–81.\(^\text{18}\) Over the period 1927–80,

---

18. Our data set is different from those used by Shiller (1981a). However, our data set produces 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 our model were fit to his data sets. Equation (17) was also fitted 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
the discrete-time version of (11) estimated by ordinary least squares (OLS) is given by

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

\[\bar{R}^2 = 0.47, \quad \text{D-W} = 1.53.\] (12)

In (12), \(D(t)\) refers to aggregate NYSE dividends totaled over year \(t\), and \(P(t)\) refers to price at the end of year \(t\). The numbers in parentheses under the coefficients are standard errors, 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 (12) indicate that the deviations of real dividend changes from their normal growth rate covary positively and strongly with the previous-year unexpected cum-dividend price changes and negatively with the previous-year dividend yield. The Durbin-Watson (D-W) statistic suggests that there is positive autocorrelation in the residuals of the OLS fit of (12). Disturbance correlation can arise in various ways. As already noted, our simple model assumes 1-period adjustment in dividends by management, whereas longer lags are entirely possible. Further, as we noted earlier, it is possible that the target dividend yield is not literally constant. For example, yield might change if tax rates, the technology of communications and trading, or the mix of institutional and individual ownership changes. It is highly likely that any yield changes induced by such factors, which will show up in the residuals in (12), are serially correlated. Indeed, any omitted variables that are serially correlated could be a potential source of residual autocorrelation.\(^{19}\)

In light of the autocorrelation in the residuals of (12) we reestimated (11) using generalized least squares (GLS), and the results are

\(^{19}\) 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.
\[
\log\left[ \frac{D(t + 1)}{D(t)} \right] + \frac{D(t)}{P(t - 1)} = -0.234 + 0.444 \log\left[ \frac{P(t) + D(t)}{P(t - 1)} \right] - 0.085 \log\left[ \frac{D(t)}{P(t - 1)} \right] + u'(t + 1),
\]

\[R^2 = 0.53, \quad D-W = 1.83.\] (13)

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 the GLS fits are essentially the same in that they explain about 50% of aggregate NYSE real dividend changes. As will be shown in Section VI, the explanatory power of our single-equation aggregate time-series model\(^20\) is considerably higher than that of univariate trend-autoregressive models such as the one fitted by Shiller (1981\textit{a}).

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.\(^21\)

Because the coefficient on percentage price changes is both significantly greater than zero and significantly less than one, this finding is also consistent with Lintner's stylized fact that managers smooth dividends by responding to 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 (12) and (13) 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-price ratio is negative in

\(^{20}\) Our \(R^2\) is below the 85% figure given in Lintner (1956). However, Lintner, who stressed that his results were preliminary, pooled his time-series observations from 1918 to 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--64, report (e.g., in their table 2) average \(R^2\) figures of (roughly) 40%--45%, which are comparable with ours.

\(^{21}\) By inspection of (12) and (13), the OLS and GLS estimates of the coefficient on percentage price changes are negligibly different. If the GLS transformation is interpreted as a quasi-differencing operator, and if a logic similar to that of Plosser, Schwert, and White (1982) is applied, then the invariance of the coefficient estimate to GLS suggests, in terms of Hausman's (1978) specification test, that there is no simultaneity bias in regressions (12) and (13).
both (12) and (13), 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, however, rather small, which at best suggests that a substantial period of time is required for the dividend-price ratio to converge to its steady-state distribution. Thus, using the $-0.085$ estimate from (13), a conventional “half-life” calculation shows that it takes more than 8 years for the expected value of this ratio to move halfway from its initial value to its expected steady-state value.\footnote{22}

To investigate further the extent of synchronization between dividend changes and price changes, we computed the leads and lags of the percentage changes in dividends regressed on percentage changes in prices estimated by the Hannan-efficient procedure, and these are plotted in figure 1. By inspection, the cross correlation at the lag in price change of 1 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% correlation between contemporaneous (i.e., year $t + 1$) unanticipated price changes and the residuals from our regression (13). As noted in the discussion in the introduction of the informational content of dividends, it is difficult to identify an announcement date for aggregate dividends in a meaningful way. 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% correlation is likely to be a significant overstatement of the announcement effect of aggregate dividends.

The efficient market hypothesis does not rule out the change in this-

\footnote{22. 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 reduces the absolute magnitude of the (negative) dividend-yield coefficient. Dynamic regression models with autoregressed 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 Shiller’s posit, the dividend-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.}
Fig. 1.—Lead and lag structure of deviations in annual percentage real dividend changes, around their expected growth rate, on percentage unexpected real cum dividend stock price changes, for the NYSE value-weighted companies over the period 1927-80, computed using the Hannan-efficient procedure. The equation used is

$$\log \frac{D(t + j)}{D(t + j - 1)} + \frac{D(t + j - 1)}{P(t + j - 2)} = a_0 + a_1 \log \frac{P(t)}{P(t - 1)} + \frac{D(t - 1)}{P(t - 2)},$$

for $j = -4, \ldots, 0, \ldots, 6$. Dividends, $D(t)$, are defined as year-$t$ cash dividend payments for all NYSE companies, and the index “price” $P(t)$ is the end-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 $t + j$.

period dividends being predicted by variables that are observable prior to this period, as it does for speculative price changes. Nevertheless, if the posited economic process underlying the specification of (11) is a reasonably accurate description of reality, then lagged price changes much beyond the 1-year lag specified in (11) might be expected to exhibit relatively little explanatory power in forecasting the dividend change this period. If an unanticipated change in permanent earnings causes managers to change dividends, then strict rational forecasting would seem to dictate that their decision be based on their most recent assessment of permanent earnings; hence earlier revisions in those assessments should have relatively little effect 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 perma-
Division's behavior for Aggregate Stock Market

ment-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 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, they can use the most up-to-date information. The partial dividend adjustment will be appropriate in a wide variety of contexts in which the policy problem facing managers does not admit a certainty-equivalence solution.

Even if managers forecast 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 equation (5), it is unlikely that the lag between a change in permanent earnings and the change it induces in subsequent aggregate dividends could be much shorter than a year.23 The correlations between the change in dividends and lagged price changes displayed in figure 1 are therefore consistent with this view of rational forecasting by managers. As a further, more quantitative, test, (11) was reestimated with 5 years of lagged unexpected price changes as additional variables. None of these additional lagged variables had a coefficient point estimate more than 1 standard error from zero, and the F-statistic for their inclusion is .361, which has a p value of .872.

If these empirical results had turned out differently, they would imply neither 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 that are not based on the most up-to-date information. It is, however, reassuring for the overall credibility of our model that the data tend to support such "superrational" forecasting 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 that exhibit high contemporaneous correlation with changes in aggregate-stock prices. It is therefore rather unlikely

23. In their classic events study of stock splits and the cash-dividend changes that often accompany them, Fama et al. (1969) report a result similar to ours: stock splits, and the increased dividends that typically accompany them, were on the average preceded by abnormal price increases. In their study, some of the average "run-up" in prices took place earlier than 12 months before the stock-split event, but Fama et al. deliberately select the individual companies that, ex post, split their stocks. The early small run-up in prices could easily wash out in our aggregate data, and in any case, Fama et al.'s study does not provide information on changes in cash dividends other than those associated with the stock split.
that the change in stock price is merely serving in (11) as a proxy for some other observable variables that, if included, would cause the significance of the coefficient on the price change to disappear.

We have investigated whether the dividend response to stock-price changes is symmetric with respect to negative and positive price changes. The point estimate of this elasticity with respect to the 20 annual negative-price-change observations in our sample is 0.597, while it is 0.271 with respect to the positive price changes. Although an asymmetry of dividend response to negative and positive price changes could be readily explained in a more complete dividend model, the small number of negative-price-change observations causes the standard error for the difference in point estimates to be so large that equality of the elasticity coefficients cannot be rejected. In addition, two of the negative-price-change observations are the 1929 and 1974 market "crashes" in which the constant-discount-rate assumption is surely strained.

In his discussion of the stability of Lintner's original regression results, Tarshis (1959, p. 118) writes, "Everything else in the economy changed in those years: it seems unreasonable that these many changes exerted no influence upon dividend policy." More recently, Brittain (1966) and Miller (1985) document apparent shifts in dividend-payout policies in the late 1930s and the 1940s, which they attribute to tax changes. As we have already discussed, the factors that historically could have caused dividend policy to be important may themselves have diminished in significance over the sample period and, thereby, induced secular changes in the dividend-response function. With these observations in mind, we examine the temporal stability of the coefficients in our regression model (11).

A plot of the elasticity coefficient $a_1$ in (11), estimated recursively forward and recursively backward, in presented in figure 2. The plot suggests a downward shift in the coefficient in the 1940s. This suggestion is confirmed by formal squared CUSUM, Quandt (1958, 1960), and Chow tests (not reported here).24 The absolute value of the coefficient of reversion of the dividend-price ratio, $a_2$, decreases steadily from approximately −0.49 at the beginning of the period to −0.03 for the full period up to 1979. When (11) is estimated using robust regression and the tests of coefficient stability repeated (see Kuh, Samarov, and Shell 1986), the change in the model coefficients appears to have begun as early as 1938. That instability might be attributed to the undistributed-profits tax in the years 1936 and 1937.25 A shift in dividend

24. Both these tests and the recursive estimation were performed using the TROLL program RECUR. Details of all results cited but not reported are available on request.

25. The usual problems of identifying the point of shift in the regression regime are exacerbated when the shift occurs at the beginning (or end) of the sample period, where the tax-change explanation would place it (e.g., Quandt 1958, pp. 877–78). Moreover, an advocate of Tarshis’s position might argue that it is at best only “half naive” to search for a single discrete shift in regimes over our long sample period.
Fig. 2.—Plot of the recursively estimated coefficient of elasticity of the deviation in annual percentage real dividend change around its expected growth rate, with respect to cum dividend prior-year stock price change, for the value-weighted NYSE companies over the period 1927–79. The plot is of the point estimate of the coefficient $a_1$, where

$$
\log \frac{D(t + 1)}{D(t)} + \frac{D(t)}{P(t - 1)} = a_0 + a_1 \log \frac{P(t) + D(t)}{P(t - 1)}
$$

$$
+ a_2 \log \frac{D(t)}{P(t - 1)} + u(t + 1)
$$

and where $a_1$ is estimated recursively forward (labeled “A”) and recursively backward (labeled “B”). Here, dividends, $D(t)$, are defined as year-$t$ cash dividend payments for all NYSE companies, and the index “price” $P(t)$ is the end-of-year-$t$ value-weighted NYSE index. The dividend-price ratio, $D(t)/P(t - 1)$, is, up to a constant, the expected rate of growth of dividends in year $t + 1$.

Policy in the late 1930s and the 1940s is, of course, not necessarily attributable to taxes. For example, Officer (1971) shows that stock-market volatility also shifts toward the end of the 1930s. We do find that if the elasticity coefficient $a_1$ in (11) is allowed to depend linearly on a naive measure of the market’s volatility—the square of the cum-dividend stock-price change $\log P(t) + D(t)/P(t - 1)$—the fit is marginally, but significantly, improved. However, the apparent shift in dividend policy remains.26 We consider and reject the hypothesis that variations in the discount rate are the reason for the 1930s–1940s shift in dividend policy.27

26. The significance of stock-price volatility changes in explaining dividend movements has implications for the stock-price rationality debate. If manager’s “smooth out” what they consider to be irrational fluctuations in stock prices as measures of permanent earnings, then it seems rather implausible that the magnitude of the response coefficient to stock-price changes would decrease, rather than increase, when stock-price volatility decreases.

27. Tests for the effects of a nonconstant discount rate were performed using the program ADAPT, written by Craig F. Ansley.
As Miller and Modigliani (1961) make clear, the determination of gross dividends is the principal issue to be explained in resolving the dividend puzzle. As an empirical matter, we find that both the explanatory power and the coefficient estimates of our model remain virtually unchanged whether it is fit to gross dividends or to aggregate net dividends (measured by gross dividends less net new stock issues). This finding may be surprising in the light of the recent 5-year wave of mergers and acquisitions financed by cash and debt, which has been large enough to cause net dividends actually to exceed gross dividends by significant amounts. It does not appear, however, that such large additional “distributions” to equity holders were typical during our earlier sample period, although the CRSP data from which we constructed our net-dividend series may not account completely for all such transactions.

For an all-equity-financed firm, net operating cash flow minus investment equals the net dividend, as a cash-flow identity. Because net operating cash flow is largely uncontrollable by managers in the short run, it would appear from the cash-flow identity that a model of net-dividend policy cannot be distinguished from a model of investment choice. Thus the empirical invariance of our model to gross-versus-net dividends raises the question of whether it describes dividend policy or corporate investment behavior.

We believe that the aggregate-net-dividend series is not a reliable indicator of investment policy and therefore that our model should not be interpreted as one of investment behavior. As noted in note 5 above, an important source of slippage in the empirical application of the identity is the cash inflow and outflow of the firm from financing sources not accounted for in the standard measure of net dividends. These sources would include changes in publicly issued and privately placed corporate debt, bank loans, trade credit and other short-term accruals, lease contracts, and other legal liabilities such as pensions and customer warranties.

New stock issues that figure in the difference between gross- and net-dividend calculations are small by comparison with these sources and uses of funds. For example, a 5% shift in the approximately $120 billion of outstanding nonfinancial corporate net trade credit would alone amount to more than half the $8–$10 billion level of average annual common-stock offerings. Public offerings of corporate debt typically exceed annual new equity issues; hence changes in debt offerings could easily offset changes in equity financing without affecting investment flows. Changes in the level of interest rates on floating-rate debt will also cause changes in cash flow without a change in investment. Thus it seems to us that dividends net of new equity issues cannot be used reliably to infer variations in investment policy.

In summary, our main empirical results are (i) that past (or possibly
contemporaneous) changes in stock-market prices explain a significant portion of the change in aggregate dividends and (ii) that the (partial) elasticity of the dividend response to a change in price is positive and significantly less than one. Further, there are reasonable grounds for believing that these results will be robust to more refined versions of this dividend model. If, for example, log-price and log-dividend levels follow integrated stochastic processes, as they seem to do empirically (cf. Kleidon 1986), the target-payout ratio can satisfy the steady-state properties assumed in the reduced form of our error-correction model (11).

VI. Further Discussion and Comparison of Our Model with Others in the Literature

In this section, we compare the fit of the model of aggregate dividends that has been developed in previous sections with that of the “final form” trend-autoregressive model employed by Shiller (1981a) to describe aggregate dividends. We show that our model fits considerably better than does the trend-autoregressive model and that lagged dividends explain little, if any, of the variation in aggregate dividends once prior-period stock-price changes are taken into account. We also claim that the characteristics of the observed distributions of dividend and stock-price changes reported by Shiller can be readily interpreted within the context of our model.

In Section VI C, the explanatory power of our model of aggregate-dividend behavior, in which stock-price changes are used to measure changes in firms’ permanent earnings, is also compared with models such as those fitted by Lintner (1956), Brittain (1966), and Fama and Babiak (1968), in which dividend movements are explained by accounting-earnings changes. We find that our model, using only prior-period stock-price changes, fits at least as well as accounting-earnings-based models, which use contemporaneous accounting-earnings data. Moreover, our model significantly outperforms models that use lagged accounting earnings only, as would be required if these models were used to forecast future dividend changes.

A. A Trend-Autoregressive Model for Aggregate Dividends?

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 repeat essentially the same OLS regression on our data set, and the results are

\[
\Delta \log D(t) = 2.492 - 0.249 \log D(t) + 0.004t + u(t), \quad (14)
\]

\[
\begin{align*}
& (0.890) & & (0.089) & & (0.002) \\
\bar{R}^2 &= 0.130, & D-W &= 1.495.
\end{align*}
\]
Because the left side of (14) 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 one with a $-2.80$ $t$-statistic in our sample versus a $-3.33$ $t$-statistic in his. The $t$-statistic for the trend coefficient in (14) is 2.00. This finding serves to confirm our belief that the important empirical results derived from our 1926–81 data set are not likely to be significantly altered if fit to the longer 1871–1979 data set used by Shiller in his analysis of aggregate-dividend and stock-price behavior.

Because the Durbin-Watson statistic suggests positively autocorrelated residuals, the lagged endogenous variable in (14) may cause the OLS coefficients to be biased. We therefore reestimated the reported Shiller equation (14) using a GLS iterative technique, and the results are

$$
\Delta \log D(t) = 5.225 - 0.524 \log D(t) + 0.009t + u(t),
$$

$$
(1.211) \quad (0.121) \quad (0.003)
$$

$$
\bar{R}^2 = 0.243, \quad D-W = 1.85.
$$

Hence, when the OLS specification of Shiller's autoregressive model for dividends is correctly adjusted for autocorrelation, its measured explanatory power almost doubles. That is, only about half of the total 24% 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.

Given the apparent statistical significance of the coefficients in (14) and (15), it is perhaps tempting to some to conclude that dividends follow an autoregressive process that approaches a steady-state distribution (possibly around a positive trend). Such a conclusion, if true, has far-reaching implications for the whole of financial economics. For example, if (14) were to describe fully the "true" dividend process (i.e., $\bar{R}^2 = 1$), then this result, along with even a casual inspection of the stock-return time series, would surely imply that stock prices are

28. 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.

29. Comparing the OLS and GLS $\bar{R}^2$s provides only a heuristic measure of the incremental explanatory power afforded by the GLS regression because the OLS $\bar{R}^2$ is not a proper benchmark in light of autocorrelation in the residuals. The $\bar{R}^2$ 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 $\bar{R}^2$ 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 the actual dependent variables.
“too volatile.” If, as implied by (15), dividends were to regress over 90% 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 rational 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 this is perhaps an appealing hypothesis, the real world is not this way, as further analysis of (14) and (15) will clearly indicate.

The fit of the autoregressive model (14) and (15) is rather poor, with half the explanatory power of (15) represented by unspecified variables. With respect to a different regression on similar data, Shiller (1981a, p. 433) gives one possible explanation for low $\hat{R}^2$: 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 alternative explanation is simply that the autoregressive process posited in (14) and (15) is not an accurate specification of the dividend process.

Motivated by the analysis of our model of the dividend process, we add the 1-year lagged unanticipated change in the log of stock price to the specification of (15). By the same iterative GLS procedure used in (15), the results are

$$\Delta \log D(t) = 2.107 + 0.347 \log \left[ \frac{P(t)}{P(t-1)} + \frac{D(t)}{P(t-1)} \right]$$

$$- 0.213 \log D(t) + 0.004t + u(t),$$

(16)

$$\hat{R}^2 = 0.473, \quad \text{D-W} = 1.755.$$ 

By inspection, the addition in (16) of the unexpected price change in the previous year doubles the explanatory power of (15). This measured increase in $\hat{R}^2$ greatly understates the effect of this added variable because, in addition to increasing $\hat{R}^2$ by 100%, it also virtually eliminates the explanatory power of the remaining unspecified variables whose effects are captured by the GLS procedure. We would

30. As was the case for the OLS and GLS fits (12) and (13) of our model, the GLS fit of (16) does not significantly improve on its OLS fit. The $F$-statistic for the autoregressive correction in (16) is 2.85, which is not significant at the conventional 5% level. Thus the addition of log price change to (15) substantially increases its explanatory power and improves the autoregressive model’s specification (14) 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.
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 explore the relative importance of the specified variables further, equation (16) was reestimated—first, with the time trend deleted and, second, with both the time trend and log $D(t)$ removed. The results are

\[
\Delta \log D(t) = 0.566 + 0.388 \log \left( \frac{P(t)}{P(t-1)} \right) + \frac{D(t)}{P(t-1)}
\]

\[
- 0.055 \log D(t) + u(t),
\]

\[
\hat{R}^2 = 0.437, \quad D-W = 1.814,
\]

and

\[
\Delta \log D(t) = -0.014 + 0.402 \log \left( \frac{P(t)}{P(t-1)} \right) + \frac{D(t)}{P(t-1)} + u(t),
\]

\[
\hat{R}^2 = 0.435, \quad D-W = 1.82.
\]

Comparing (16) with (17a), the elimination of the time-trend variable causes only a modest reduction in $\hat{R}^2$, and it has little effect on the estimated coefficient of the log price change. However, by eliminating the time-trend variable, the magnitude of the regressive coefficient on log $D(t)$ falls by 75%, and, with a $p$-value equal to .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 (15). If either variable is removed, then the magnitude and the statistical significance of the coefficient of the remaining variable are both nil. In this light, it is not surprising that the elimination of log $D(t)$ as a variable in (17b) has no effect on either the $\hat{R}^2$ 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 surely it belongs in the specification of the dividend process. Because it alone accounts for over 90% of the explanatory power of (16), its omission from (14) and (15) 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. Its inclusion produces an insignificant coefficient and actually causes the corrected OLS and GLS $\hat{R}^2$'s in (12) and (13) to fall. Our
model would, however, predict that changes in log dividends are related to log $D(t)$ through the dividend-price ratio, $\log[D(t)/P(t-1)]$. Although not explicit in our simple model, it is entirely consistent with the spirit of the model that lagged changes in log dividends explain part of the adjustment process used by managers to decide on subsequent dividend changes. If this were the case, then log $D(t)$ may be a proxy for these lagged changes. The inclusion of such lagged dividend changes would in no substantive way change the conclusions derived for the dividend and rational stock-price processes in Section V. To investigate this possibility, we reestimate equation (13) with the addition of log $D(t)$, and the fitted results are given by

$$\log \left[ \frac{D(t + 1)}{D(t)} \right] + \frac{D(t)}{P(t - 1)} = 1.550 + 0.441 \quad (1.923) \quad (0.061)$$

$$\log \left[ \frac{P(t)}{P(t - 1)} + \frac{D(t)}{P(t - 1)} \right] - 0.247 \log \left[ \frac{D(t)}{P(t - 1)} \right] - 0.220 \log D(t) + u(t + 1), \quad (0.267) \quad (0.256)$$

$$\overline{R^2} = 0.532, \quad \text{D-W} = 1.88.$$

The $F$-statistic for log $D(t)$ in (18) is 3.23, which is insignificant at the 5% level. If log $D(t - 1)$ is added to (18), its estimated coefficient is $-0.135$, while that of log $D(t)$ becomes 0.064. Although this result suggests that log $D(t)$ in (22) 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 (18) had no effect on the point estimate of the coefficient of lagged log-price change.31

In summary, adding the trend and lagged-dividend variables of Shiller’s autoregressive model to our specification does little to improve the explanatory power of our model in terms of $\overline{R^2}$. 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 (14) 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 it is a common belief that the dividend time series is quite smooth by comparison with the price series. Thus a distributed lag of past dividends together with a time trend might be expected to do a better

31. This regression is almost the equivalent of the Granger-Sims causality test referred to in the causality discussion in Section III.
job than stock-price changes in explaining subsequent dividend changes, almost independently of the "true" economic specification. 32

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 III, the resolution of such issues must ultimately come from economic reasoning. As Shiller (1983, p. 236) notes 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, there appears to be little theoretical structure to support the assumption that dividends follow a stationary process with a trend. In particular, there is neither an oral nor a written tradition in the financial economics literature that assumes dividends and rational stock prices have stationary distributions. 33

One could, of course, revive Malthus, or the more contemporary "limits-to-growth" view of economics, to justify the assumption of a steady-state distribution for the levels of real dividends and prices. This theory, however, also rules out an exponential growth trend in these levels. Refitting equation (14) without the trend, the OLS estimate is given by

$$\Delta \log D(t) = 0.802 - 0.076 \log D(t - 1) + u(t), \quad (19)$$

$$\begin{align*}
(0.576) & \quad (0.055) \\
R^2 &= 0.017, \quad D-W = 1.576.
\end{align*}$$

By inspection of (19), it would appear that, at least in the dividend series, there is no evidence at all to support this "zero-growth" model.

Perhaps notable by its absence from this section is any discussion of

32. It is all the more surprising because the model does not use contemporaneous price changes. Because all the variables in our model are lagged, eq. (18) is a "true" forecast equation in the sense that, at time $t$, it provides an unconditional forecast for $D(t + 1)$. The relatively high $R^2$ suggests that aggregate dividends may be forecasted rather successfully.

33. There is, of course, ample precedent in the economics literature for assuming that relative values such as the dividend-price and earnings-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. Further, unless investors are risk neutral and the riskless rate of interest is constant, the assumption of a constant expected return on the market, made in Shiller (1981a, 1981b) and in much of the finance literature, is inconsistent with the trend-autoregressive process for dividends and stock prices. For proofs and further discussion, see Fama (1977) and Myers and Turnbull (1977).
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 that have much power in forecasting future stock-price changes. Moreover, using autocorrelation andDickey-Fuller (1979, 1981) tests, Kleidon (1986) finds that neither the arithmetic nor the geometric Brownian motion models can be rejected against the trend model for the Standard and Poors 500 annual composite index over the period 1926–79. We would expect that these same results would obtain for our data set.

Shiller (1981a, pp. 432–33) does report that the dividend-price ratio appears to forecast next-period 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 explain the positive relation between stock returns and lagged dividend yield. 34 In addition, Miller and Scholes (1982) show that holding-period returns for individual firms can be forecasted just as accurately using the reciprocal of stock price as they can using dividend yield, which suggests that the numerator of the dividend-price ratio does not play an important role in its predictive power for future stock-price changes. Of course, forecastability of discrete-period returns is compatible with market efficiency if it reflects no more than forecastable variations in expectations of those returns. However, as the size of Shiller’s $R^2$ estimate would seem to indicate, forecastable changes in expected returns probably account for only a small amount of the variation in 1-year holding-period returns.

B. The Distribution of Rational Stock Prices and Dividends in Our Model

In our dividend model, the dividend-to-price ratio has a stationary distribution with a finite variance. It follows that, for $T$ large, $\text{var} [\log(P(T)/P(0))]$ will be proportional to $T$. Hence, for large $T$, the cumulative dynamics for $P$ can be well approximated as having come from a geometric Brownian motion with an instantaneous expected rate of growth equal to $(\alpha - \rho)$, where $\rho$ is the expected “long-run” dividend-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 a reply to a comment on his work by Basil Copeland (1983),

34. Study of our sample using blunt interocular analysis suggests that there are a few influential “outliers” in the annual data that cause the correlation, and the correlation disappears all together with monthly data.
Shiller (1983, p. 237) notes, “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; that is (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 Poors 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% level, he concludes that prices are too volatile to be consistent with this model. In our much shorter 1926–81 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.

The extreme polar case of our model is $D(t) = \rho P(t)$, where managers do not attempt to smooth dividends at all and fully and immediately adjust dividends to reflect unanticipated changes in permanent earnings. If the “short-run” instantaneous variance rate for logarithmic price changes is $\sigma^2$, then in our model the corresponding short-run 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 observation 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 (13) of 0.44 for $\lambda$ would imply that the ratio of the variances of annual log price changes to log dividend changes would exceed five. Because the model explains only about 50% of the variation in dividends, the actual ratio is reduced to 2.64.

As is generally true of “smoothed” processes that 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 $\lambda < 1$,

$$\frac{\text{var}\{\log[D(T)/D(0)]\}}{T} \leq \frac{\text{var}\{\log[P(T)/P(0)]\}}{T},$$

for any interval $T$, with equality holding only in the limit as $T \to \infty$. Our model of rationally determined prices and rationally determined divi-
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.

Shiller (1981a, p. 428) also 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 in which 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 that 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 log 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 that 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 reflect fully 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 long-run 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, \( \sigma^2/n \), is smaller than the variance of stock-price changes around trend, \( \sigma^2 \). If \( m_4 \) denotes the fourth central moment of the stock-price-change distribution, then the fourth moment of the dividend-change distribution is \( km_4/n \). 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, \( m_4/\sigma^4 \). 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.

The basic reasoning underlying our claim that a large kurtosis of
dividend changes relative to stock-price changes is evidence in support
of dividend changes being a (short-run) controlled process is hardly
new. Over a half century ago, Means (1935) used it to contend that
many product prices are "administered." By comparing the frequency
of product-price changes in "administered" and competitive markets,
he found that administered-market price changes were much less fre-
quent and that, when they did occur, the changes were much larger in
magnitude than they were in competitive markets. That is, the distribu-
tion of administered price changes has fatter tails than the one for
competitive prices. Quandt and Ramsey (1978), among others, have
developed techniques that take account of other information in addi-
tion to kurtosis to estimate the parameters of the mixed distributions
that arise from such administered price processes.

In the light of these results, we estimate 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.\(^ {35} \) As it happens, the sample kurtosis for stock
prices is not much different than the kurtosis of three for a normal
distribution, whereas the sample kurtosis for dividend changes is more
than two times larger.

C. Stock Price versus Accounting Earnings as a Measure of
   Permanent Earnings

In deriving our reduced-form dividend model (11), we adopted the
specification (9) that stock prices embody rational predictions of firms’
future net cash flows and thus permanent earnings. In past empirical
work based on Lintner’s model, the practice is to use accounting earn-
ings, modified accounting earnings, or cash-flow data to measure firms’
permanent earnings. Hence we compare the performance of our model
to these alternative measures of permanent earnings.

Since the readily available data on aggregate earnings are for the
Standard and Poors 500 companies, the results reported in this section
pertain to this index. The fit of our model (11) for the Standard and

\(^ {35} \) Although not presented here, the sample distributions of dividend and stock-price
changes were plotted using Tukey’s (1970) robust statistics. These plots are consistent
with the relatively fat tails of the dividend process implied by our reported kurtosis
statistics.
Poors index is virtually identical to that reported in (12) or (13) for the NYSE. The GLS fit for the period 1928–80 is given by

\[
\log\left[ \frac{D(t + 1)}{D(t)} \right] + \frac{D(t)}{P(t - 1)} = -0.140 + 0.426 \log\left[ \frac{P(t + D(t))}{P(t - 1)} \right] \\
- 0.056 \log\left[ \frac{D(t)}{P(t - 1)} \right] + u'(t + 1),
\]

\[ R^2 = 0.41, \quad D-W = 1.94. \] (20)

If the change in accounting earnings from year \( t \) to year \( t + 1 \) is used to explain (not predict) the dividend change from year \( t \) to year \( t + 1 \), and if \( (\alpha - D/P) \) is used to account for the expected growth in accounting earnings, then the GLS fit of this contemporaneous-accounting-earnings model can be expressed as

\[
\log\left[ \frac{D(t + 1)}{D(t)} \right] + \frac{D(t)}{P(t - 1)} = 0.146 \\
+ 0.518 \log\left[ \frac{E(t + 1)}{E(t)} + \frac{D(t)}{P(t - 1)} \right] \\
+ 0.039 \log\left[ \frac{D(t)}{P(t - 1)} \right] + u'(t + 1),
\]

\[ R^2 = 0.44, \quad D-W = 2.06, \] (21)

where \( E(t) \) refers to the aggregate accounting earnings for the Standard and Poors companies over year \( t \). These results suggest that data on year \( t + 1 \) accounting earnings, which only become available at the end of year \( t + 1 \), explain no larger a percentage of aggregate dividend changes in year \( t + 1 \) than is explained by using price data available at the beginning of year \( t + 1 \). Moreover, the coefficient on the dividend-price ratio (i.e., the dividend–permanent earnings ratio) has the wrong sign for regressivity, although that coefficient is imprecisely measured with only 50 years of data. Except for the regressivity coefficient, the contemporaneous-accounting-earnings model has roughly the same fitted characteristics as the lagged stock-price model.

We examined the temporal stability of the coefficients in regression model (22) and found a pattern of secular decline in their magnitudes quite similar to the one reported for our model in Section V. When backward recursion is used over the period 1926–80, the full-sample estimate of the coefficient of the earnings-change variable declines by about 75% by the end of the sample period. The point estimate for the sign on the dividend-price variable decreases in magnitude but continues to have the “wrong” sign. As Brittain (1966) found, the standard
diagnostic tests (Chow, Quandt likelihood, CUSUM squared) indicate that the instability occurs around the years 1939–42.

Empirical analysis of the elasticity of response in accounting earnings to lagged stock-price changes shows that it remains reasonably stable over the sample period. Thus it appears that the secularly reduced responsiveness of dividend changes to permanent-earnings changes is essentially the same, whether permanent earnings are measured by stock price or by accounting earnings. Therefore it would seem unlikely that this seemingly secular change in dividend behavior is the result of a temporal pattern of increasing disparity between managements’ assessments of permanent earnings and those implied by stock prices.

If the accounting earnings change from period \( t \) to \( t + 1 \) is substituted for the period \( t \) to \( t + 1 \) change, the accounting model’s explanatory power drops to about 20% (and to only 10% if OLS is used). An examination of the distributed leads and lags of dividend changes and accounting-earnings changes confirms that the substantial portion of their association is, in fact, contemporaneous. This result is consistent with Fama and Babiak’s (1968) finding that their “best” accounting-earnings-based model of dividend changes includes contemporaneous and lagged earnings levels since our accounting-earnings-change variable involves both contemporaneous and lagged accounting-earnings levels.

Contemporaneous accounting earnings are roughly on a par with lagged stock-price changes in explaining aggregate-dividend changes, in part because lagged stock-price changes themselves provide reasonably good forecasts of the subsequent earnings changes. If the component of contemporaneous-accounting-earnings changes that could have been predicted from the lagged stock-price change is removed, then the unpredictable component of contemporaneous earnings, which we denote by \( UE(t) \), does add significantly (at the 95% level) to past price changes in explaining aggregate-dividend movements. Using GLS over the 1929–79 period,

\[
\log \left( \frac{D(t + 1)}{D(t)} \right) + \frac{D(t)}{P(t - 1)} = -0.069 + 0.426 \log \left( \frac{P(t) + D(t)}{P(t - 1)} \right) \\
+ 0.246UE(t) - 0.032 \log \left( \frac{D(t)}{P(t - 1)} \right) + u'(t + 1), \tag{22}
\]

\[R^2 = 0.47, \quad D-W = 1.88.\]

The \( F \)-statistic for inclusion of the earnings forecast error \( UE(t) \) is about 7.96, which is significant at the 5% level. The OLS and GLS estimates of (22) are virtually identical, which is consistent with the earnings-forecast error eliminating part of the serial dependence in the
disturbances. The measured correlation between the variable \( \log[P(t + 1)/P(t) + D(t + 1)/P(t)] \), which, up to a constant, is the unexpected price change from the end of period \( t \) to the end of period \( t + 1 \), and the (GLS) residual \( u(t + 1) \) in (19) is about 10%. By including contemporaneous unanticipated earnings, this contemporaneous correlation between unexpected price changes and the (GLS) residual \( u(t + 1) \) in (22) is reduced to about 3.3%. Thus the previously noted small amount of "information content" in aggregate dividends is further reduced by taking account of earnings changes. We pursue this no further because, as we noted earlier, the information content of dividends and earnings is probably not well defined for an aggregate of corporations over a coarse annual grid.

VII. Conclusion

In this paper, we have shown that aggregate dividends exhibit a systematic time-series behavior that can be well described by an error-correction model in which aggregate-real-dividend changes are driven by the 1-period lagged real changes in stock prices. Although the time series of aggregate dividends is considerably less volatile than the stock-price series, this model significantly outperforms the trend-autoregressive model in which a distributed lag of past dividends together with a time trend is used to explain subsequent dividend changes.

The stock-price model performs on a par with dividend models that use contemporaneous- and lagged accounting-earnings variables, as in previous studies of Lintner's (1956) model by Brittain (1966), Fama and Babiak (1968), and Watts (1973). However, because the stock-price model uses only lagged prices, it can be used to forecast future dividend changes, whereas the accounting-earnings model that employs contemporaneous earnings cannot be used to forecast. The version of the accounting-earnings model that uses only lagged accounting earnings to forecast dividends significantly underperforms the stock-price model.

As noted at the outset, a wide range of possible micro theories of dividend behavior could be consistent with the observed systematic behavior of aggregate dividends. However, our finding that aggregate dividends do exhibit systematic time-series behavior provides evidence that strictly firm-specific theories of dividends such as signaling cannot by themselves explain the dividend puzzle.

Appendix

The long-run steady state (4) in the text can be imposed on the short-run dynamics (3) by the further coefficient restrictions

\[ \lambda - \theta_1 = 1 - \phi_1 = \gamma, \]  

(A1)
that is,
\[ \lambda - \theta_1 = 1 - \phi_1, \]
and by setting the constant in the deterministic function \( a(t) \) in (3) equal to \( \gamma \beta \). Denoting the function \( a(t) \) without the constant by \( a'(t) \) and incorporating (A1), we can rewrite (3) as follows:
\[
\log[D(t + 1)] - \log[D(t)] = \gamma \beta + a'(t) + \lambda \{ \log[E^m(t)] - \log[E^m(t - 1)] \} \\
- \gamma \{ \log[D(t)] - \log[E^m(t - 1)] \} + \eta(t + 1)
\] (A2)
or, by rearrangement,
\[
\log[D(t + 1)] - \log[D(t)] = a'(t) + \lambda \{ \log[E^m(t)] - \log[E^m(t - 1)] \} \\
+ \gamma (\beta - \{ \log[D(t)] - \log[E^m(t - 1)] \}) + \eta(t + 1).
\] (A3)

To derive (5), define the expected logarithmic change in permanent earnings,
\[ \{ \log[E^m(t)] - \log[E^m(t - 1)] \}, \]
as \( m(t - 1) \). Then (A3) can be rewritten as
\[
\log[D(t + 1)] - \log[D(t)] = a'(t) + \lambda m(t - 1) \\
+ \lambda \{ \log[E^m(t)] - \log[E^m(t - 1)] - m(t - 1) \} \\
+ \gamma (\beta - \{ \log[D(t)] - \log[E^m(t - 1)] \}) + \eta(t + 1).
\] (A4)
Substituting \( g(t) = a'(t) + \lambda m(t) \) gives (5) in the text.

References


Merton, R. C. In press. On the current state of the stock market rationality hypothesis.


