The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors

John Y. Campbell
Princeton University

Robert J. Shiller
Yale University

A dividend-ratio model is introduced here that makes the log of the dividend-price ratio on a stock linear in optimally forecast future one-period real discount rates and future one-period growth rates of real dividends. If ex post discount rates are observable, this model can be tested by using vector autoregressive methods. Four versions of the linearized model, differing in the measure of discount rates, are tested for U.S. time series 1871–1986 and 1926–1986: a version that imposes constant real discount rates, and versions that measure discount rates from real interest rate data, aggregate real consumption data, and return variance data. The results yield a metric to judge the relative importance of real dividend growth, measured real discount rates, and unexplained factors in determining the dividend-price ratio.

What accounts for the variation through time in the dividend-price ratio on corporate stocks? The dividend-

An earlier version of this paper was circulated as NBER Working Paper 2100. This research was supported by the NSF and the John M. Olin Fellowship at the NBER (Campbell). Any opinions expressed here are those of the authors and do not necessarily represent the views of the institutions with which they are affiliated. We are grateful to Michael Brennan, Steve Brown, David Hendry, Greg Mankiw, Sam Ouliaris, Peter Phillips, Mark Watson, and anonymous referees for helpful comments. We are particularly grateful to Jim Poterba for assistance with the data and to Andrea Beltrati for correcting an error in our unit roots test program. Address reprint requests to John Y. Campbell, Woodrow Wilson School, Princeton University, Princeton, NJ 08544.

The Review of Financial Studies 1989, Volume 1, number 3, pp. 195–228

price ratio is often interpreted as reflecting the outlook for dividends: when dividends can be forecast to decrease or grow unusually slowly, the dividend-price ratio should be high. Alternatively, the ratio is interpreted as reflecting the rate at which future dividends are discounted to today's price: when discount rates are high, the dividend-price ratio is high. In principle, the dividend-price ratio ought to have both of these interpretations at once. Yet their relative importance has never been established, and it is not clear whether these two interpretations together can account for time variation in the dividend-price ratio if one assumes that market expectations are rational. We address these questions by using long historical time series on broad stock indexes in the United States.

Our method is to test a dividend-ratio model relating the dividend-price ratio D/P to the expected future values of the one-period rates of discount r and one-period growth rates of dividends g over succeeding periods. The model might be described as a dynamic version of the Gordon (1962) model, $D/P = r - g$, which was derived under the assumption that dividends will grow at a constant rate forever, and that the discount rate will never change. This article fills a significant gap in the literature by permitting an analysis of the variation *through time* in the dividend-price ratio in relation to predictable changes in discount rates and dividend growth rates. Most previous studies of the dividend-price ratio have been concerned with the *cross-sectional* relationship between dividend-price ratios and average returns [e.g., Black and Scholes (1974)], while our own previous work on the time-series behavior of dividends and stock prices [e.g., Shiller (1981) and Campbell and Shiller (1987)] relies for the most part on the assumption that discount rates are constant.

The dividend-ratio model opens up important new avenues for econometric work. In this article we use it as follows. We think of log dividends and discount rates as two elements in a possibly large vector of variables that summarize the state of the economy at any point in time. The state vector evolves through time as a multivariate linear stochastic process with constant coefficients.[1] Stock market participants observe the state vector contemporaneously and know the process that it follows; they use this knowledge to forecast future log dividends and discount rates.

If dividends and discount rates are observable ex post, then this structure, together with the dividend-ratio model, implies restrictions on the joint time-series behavior of dividends, discount rates, and stock prices. In particular, the difference between the ex post stock return and the ex post discount rate should not be predictable from a linear regression on infor-

[1] This structure is consistent with models in which managers determine dividends without reference to stock prices, and also with "dividend-smoothing" models in which managers react to prices in setting dividends [Marsh and Merton (1986, 1987)].

mation known in advance, and the log dividend-price ratio should be an optimal linear forecaster of the present value of future dividend growth rates and discount rates. These propositions can be tested formally, and they can also be evaluated informally: for example, by comparing the history of the actual log dividend-price ratio with that of an optimal forecast from a linear vector autoregressive model.

The measurement of dividends is straightforward, but the measurement of discount rates is not. Indeed, one view is that the only source of information on discount rates is the stock price itself. Our approach can be useful even if this view is correct; as discussed further below, we can use the dividend-ratio model to obtain a better estimate of the long-term discount rate by correcting the stock price for dividend expectations. However, we begin by using several simple models which imply that discount rates can be measured outside the stock market. We recognize that these models are unlikely to be able to account for all variation in stock prices, but it is worth knowing how far they can take us toward a complete explanation. We do not attempt to provide any formal theoretical justification for the measures we use, but we note that they have been the subject of some attention in the recent finance literature [see, for example, Fama and French (1988); French, Schwert, and Stambaugh (1987); Hansen and Singleton (1983); Marsh and Merton (1986); and Poterba and Summers (1986, 1988)].

We study several versions of the basic model, which differ in their measure of ex ante discount rates. In what we will call version 1 of the model, the one-period real discount rate on stock is assumed to be constant through time. In version 2, the discount rate is assumed to be the one-period ex ante real return on short debt (Treasury bills or commercial paper), plus a constant risk premium. In version 3, the ex ante discount rate is given by the expected growth rate of real aggregate consumption per capita multiplied by the coefficient of relative-risk aversion, plus a constant risk premium.[2] In these three versions of the model, the discount rate on stock varies because the riskless real rate of interest varies, while the risk premium on stock is assumed to be constant. In version 4, by contrast, the ex ante discount rate is the sum of a constant riskless rate and a time-varying risk premium given by the conditional variance of stock returns times the coefficient of relative-risk aversion.

All four versions of the model have implications for returns—version 1, for example, implies that expected real stock returns are constant, while version 2 implies that expected excess returns on stock over short debt

[2] For a theoretical justification, see Breeden (1979), Grossman and Shiller (1981), and Hansen and Singleton (1983).

are constant—and these implications have been studied in the literature.[3] The main contribution of this article is to derive the implications of these discount rate models for stock prices, using the dividend-ratio model.

We also use the dividend-ratio model in a slightly different way. The model allows us to study the term structure of expected real stock returns implied by aggregate stock prices. The dividend-price ratio is in effect a long-term expected real return on stock, but it is contaminated in that it is also influenced by expected changes in real dividends. We can use the dividend-ratio model to purge the dividend-price ratio of expected changes in dividends, so that we derive a sort of real consol yield. This is of interest whether or not our measures of one-period discount rates, discussed above, are satisfactory.

The organization of this article is as follows. Section 1 derives the dividend-ratio model as a linear approximation to an exact relationship between stock prices, stock returns, and dividends. Section 2 discusses the stock market data and discount rate data that we use. Section 3 outlines our vector autoregressive method for analyzing movements in the dividend-price ratio, and Section 4 applies it to the data. Section 5 concludes. In the Appendix we study the approximation error in the dividend-ratio model, finding that it appears to be small in practice.

1. The Dividend-Ratio Model

We start by writing the real price of a stock or stock portfolio, measured at the beginning of time period t, as P_t. The real dividend paid on the portfolio during period t will be written D_t. The realized log gross return on the portfolio, held from the beginning of time t to the beginning of time $t + 1$, is written

$$h_t \equiv \log (P_{t+1} + D_t) - \log (P_t) \tag{1}$$

We would like to obtain a linear relationship between log returns, log dividends, and log prices. The exact relationship in Equation (1) is nonlinear, since it involves the log of the *sum* of the price and the dividend. It turns out, however, that h_t can be well approximated by the variable ξ_t, $h_t \simeq \xi_t$, where ξ_t is defined as follows:

[3] Version 1 of the model has been the subject of considerable controversy. A partial list of references is: Campbell and Shiller (1987); Fama and French (1988); Keim and Stambaugh (1986); Kleidon (1986); LeRoy and Porter (1981); Mankiw, Romer, and Shapiro (1985); Marsh and Merton (1986); Poterba and Summers (1988); Shiller (1981); and West (1987, 1988). With regard to version 2, several of the above authors have asked whether the variance of short-term interest rates might help explain the variance of stock market prices. Version 3 of the model has been analyzed extensively, following the original theoretical work of Lucas (1978) and Breeden (1979), by Grossman and Shiller (1981); Grossman, Melino, and Shiller (1987); Hansen and Singleton (1983); Hall (1988); Mankiw, Rotemberg, and Summers (1985); and Mehra and Prescott (1985), among others. Version 4 has been proposed, following an exploratory analysis by Merton (1980), by Pindyck (1984, 1986), who argues that much of the variability in stock prices can be explained by the variability of the volatility of stock returns. Against this, Poterba and Summers (1986) have argued that volatility is not persistent enough to account for much variation in stock prices. French, Schwert, and Stambaugh (1987) and Campbell (1987) also examine the relationship between volatility and expected stock returns.

$$\begin{aligned}\xi_t &\equiv k + \rho \log(P_{t+1}) + (1-\rho)\log(D_t) - \log(P_t) \\ &= k + \rho p_{t+1} + (1-\rho)d_t - p_t \end{aligned} \tag{2}$$

Here, lowercase letters denote logs of the corresponding uppercase letters. The parameter ρ is close to but a little smaller than 1, and k is a constant term.

Equation (2) differs from Equation (1) in that the log of the sum of the price and the dividend is replaced by a constant k, plus a weighted average of the log price and the log dividend with weights ρ and $(1-\rho)$. Below, we will justify this approximation rigorously as a first-order Taylor expansion of Equation (1). But first we will explain intuitively why the approximation works.

It is easiest to begin by explaining why the *difference* $\rho\Delta\log(P_{t+1}) + (1-\rho)\Delta\log(D_t)$ approximates the *difference* $\Delta\log(P_{t+1} + D_t)$. Having done this, we can derive the constant k that makes the approximation hold in levels. By a standard argument, the change in the log of $(P_{t+1} + D_t)$ is approximately equal to the proportional change in the level:

$$\begin{aligned}\Delta\log(P_{t+1} + D_t) &\simeq \frac{P_{t+1} + D_t - P_t - D_{t-1}}{P_t + D_{t-1}} \\ &= \frac{P_{t+1} - P_t}{P_t + D_{t-1}} + \frac{D_t - D_{t-1}}{P_t + D_{t-1}}\end{aligned}$$

If now we suppose that the ratio of the price to the sum of price and dividend is approximately constant through time at the level ρ, whereby $P_t \simeq \rho(P_t + D_{t-1})$ and $D_{t-1} \simeq (1-\rho)(P_t + D_{t-1})$, then we have the relationship we need:

$$\begin{aligned}\Delta\log(P_{t+1} + D_t) &\simeq \frac{\rho(P_{t+1} - P_t)}{P_t} + \frac{(1-\rho)(D_t - D_{t-1})}{D_{t-1}} \\ &\simeq \rho\Delta\log(P_{t+1}) + (1-\rho)\Delta\log(D_t)\end{aligned}$$

This explanation makes it clear that ρ is the average ratio of the stock price to the sum of the stock price and the dividend. In the static Gordon (1962) world—where the log stock return $h_t = h$, a constant, and the dividend growth rate $\Delta d_t = g$, a constant—the ratio $P_t/(P_t + D_{t-1})$ is also constant and equals $\exp(g-h)$.[4] In our empirical work below we will construct ρ by using the formula $\rho = \exp(g-h)$, setting h equal to the sample mean stock return and g equal to the sample mean dividend growth rate.

The above argument shows that the change in $\log(P_{t+1} + D_t)$ is approximated by the change in $\rho\log(P_{t+1}) + (1-\rho)\log(D_t)$. But we want our approximation to work in levels as well as changes. The constant term k in Equation (2) ensures that our approximation holds exactly for levels in the static world of constant stock returns and dividend growth rates. The

[4] To see this, just note that $\exp(g) = D_t/D_{t-1} = P_t/P_{t-1}$ and that $\exp(h) = (P_t + D_{t-1})/P_{t-1}$, so that $\exp(g-h) = P_t/(P_t + D_{t-1})$. We must have $g < h$ if stock prices are to be finite.

value of k can be expressed most simply if we define $\delta_t \equiv d_{t-1} - p_t$, the log dividend-price ratio. In the static world δ_t is a constant: $\delta_t = \delta = \log(1/\rho - 1)$. Then we have

$$k = -\log(\rho) - (1 - \rho)\delta \tag{3}$$

With this definition of k, the approximate return (which is also constant in the static world) is

$$\begin{aligned}\xi &= (1 - \rho)(d_t - p_{t+1}) + (p_{t+1} - p_t) + k \\ &= (1 - \rho)\delta + g - \log(\rho) - (1 - \rho)\delta \\ &= g - \log(\rho) = h\end{aligned}$$

where the last equality follows from the formula for ρ given above. Thus, ξ and h are equal in the static world and Equation (2) holds exactly.

When stock returns and dividend growth rates are not constant, but vary through time, then Equation (2) does not hold exactly. It holds as a first-order Taylor approximation of Equation (1).[5] The higher-order terms in the Taylor expansion of Equation (1), which are neglected in Equation (2), create an approximation error. In the Appendix, however, we present evidence that in practice the error is small and almost constant. (It is worth noting that a constant approximation error would not affect any of our empirical results since we do not test any restrictions on the means of the data.)

So far we have written our equations in terms of the log levels of dividends and prices, d_t and p_t. It will be convenient to rewrite them in terms of the dividend-price ratio $\delta_t \equiv d_{t-1} - p_t$ and the dividend growth rate Δd_t. Rewriting Equation (2) and substituting h_t for ξ_t, we get

$$h_t \simeq k + \delta_t - \rho\delta_{t+1} + \Delta d_t \tag{2'}$$

Equation (2′) can be thought of as a difference equation relating δ_t to δ_{t+1}, Δd_t, and h_t. We can solve this equation forward, and if we impose the terminal condition that $\lim_{i\to\infty}\rho^i\delta_{t+i} = 0$, we obtain

$$\delta_t \simeq \sum_{j=0}^{\infty} \rho^j(h_{t+j} - \Delta d_{t+j}) - \frac{k}{1 - \rho} \tag{4}$$

This equation says that the log dividend-price ratio δ_t can be written as a discounted value of all future returns h_{t+j} and dividend growth rates Δd_{t+j}, discounted at the constant rate ρ less a constant $k/(1 - \rho)$. It is important to note that all the variables in Equation (4) are measured ex post; (4) has been obtained *only* by the linear approximation of h_t and the imposition of a condition that δ_{t+i} does not explode as i increases. There is no economic content to Equation (4).

We can obtain an economic model of the dividend-price ratio if we are

[5] More precisely, if we rewrite the right-hand side of Equation (1) as a nonlinear function of dividend-price ratios and dividend growth rates δ_t, δ_{t+1}, and Δd_t and take a first-order Taylor expansion around the point $\delta_t = \delta_{t+1} = \delta$ and $\Delta d_t = g$, then we obtain Equation (2).

willing to impose some restriction on the behavior of h_t. In particular, suppose that we have a theory that provides an "ex post discount rate" r_t satisfying

$$E_t h_t = E_t r_t + c \tag{5}$$

Here E_t denotes a rational expectation formed by using the information set I_t that is available to market participants at the *beginning* of period t, and h_t and r_t are measured at the end of period t. Equation (5) says that there is some variable whose beginning-of-period rational expectation, plus a constant term c, equals the ex ante return on stock over the period. As an example, consider the hypothesis that the expected real return on stock equals the expected real return on commercial paper, plus a constant. Then the ex post real return on commercial paper can be used as the ex post discount rate in Equation (5).

If we can observe the ex post discount rate r_t, then Equations (4) and (5) together yield a testable economic model of the dividend-price ratio.[6] To see this, note that we can take expectations of the left- and right-hand sides of Equation (4), conditional on agents' information I_t at the beginning of period t. The left-hand side of (4) is unchanged because δ_t is known at the beginning of period t (it is in I_t).[7] The right-hand side becomes the discounted value of all expected future h_{t+j} and Δd_{t+j}, conditional on I_t. But Equation (5) implies that $E_t h_{t+j} = E_t r_{t+j} + c$, so we can substitute in expected future discount rates r_{t+j} to obtain

$$\delta_t \simeq E_t \sum_{j=0}^{\infty} \rho^j (r_{t+j} - \Delta d_{t+j}) + \frac{c - k}{1 - \rho} \tag{6}$$

Equation (6) is what we will call the dividend-ratio model, or dynamic Gordon model. It explains the log dividend-price ratio as an expected discounted value of all future one-period "growth-adjusted discount rates," $r_{t+j} - \Delta d_{t+j}$. It represents the combined effect on the log dividend-price ratio of expected future discount rates and dividends that we noted in the opening paragraph of this article.

The original Gordon model, $D_t/P_t = r - g$, can be obtained as a special case of our dividend-ratio model when discount rates and dividend growth rates are constant through time and when the constant term c equals zero. Unlike Gordon, however, we will not use our model to try to explain the mean level of the dividend-price ratio; rather, we will allow a free constant term c (representing a constant risk premium in stock returns), which means that our model restricts only the dynamics of the dividend-price ratio and not its mean level.

The dividend-ratio model has some important advantages when com-

[6] In fact, we can also test the model if we observe not r_t, but some unknown coefficient times r_t. We show how to do this in Section 3, but at this stage we assume that r_t itself is observable.

[7] This is true because we defined the log dividend-price ratio δ as the difference between last year's log dividend and the log stock price at the beginning of the year.

pared with earlier empirical models. First, it is linear in logs. This makes it easy to combine with log-linear models of dividends and prices. As stressed by Kleidon (1986) and others, log-linear models are appealing on a priori grounds, and they appear to fit the data better than linear ones do.

Second, Δd_{t+j} and r_{t+j} enter symmetrically in Equation (6); all that matters for the dividend-price ratio is their difference, that is, the growth-adjusted discount rate. This offers a significant advantage when we come to do empirical work. The original model [Equations (1) and (5)] concerns *real* prices, *real* dividends, and *real* discount rates. However, price indices used to convert nominal values to real values are measured much more poorly than are nominal dividends, share prices, and interest rates. Neither δ_t nor $\Delta d_{t+j} - r_{t+j}$ depends on the price index used,[8] so if we are willing to treat $\Delta d_{t+j} - r_{t+j}$ as a single variable, we can work in nominal terms throughout and reduce our vulnerability to measurement error. Unfortunately, if we want to study forecasts of Δd_{t+j} and r_{t+j} separately, we must rely on a measured price deflator.

The usefulness of the dividend-ratio model depends on the quality of the approximation used to derive it. This, in turn, will vary from one data set to another. The next section discusses the data sets used in this article, and the Appendix presents some measures of the approximation error in Equations (2) and (4) for these data.

2. Data on Prices, Dividends, and Discount Rates

The two main data sets used in this article are described in Table 1. The first consists of annual observations on prices and dividends for the Standard & Poors Composite Stock Price Index (S&P 500), extended back to 1871 by using the data in Cowles (1939).[9] The corresponding discount rate measures and price deflators are summarized in Table 1.

Our second data set is taken from the Center for Research in Security Prices (CRSP) series of monthly returns on the value-weighted New York Stock Exchange (NYSE) index from 1926 to 1985. Returns are reported both inclusive and exclusive of dividends, and this makes it possible to compute the levels of dividends and prices up to an arbitrary scale factor. The CRSP data incorporate careful corrections for stock splits, noncash distributions, mergers, delisting, and other potential problems [Fisher and

[8] Strictly speaking, this result assumes (as implied by the timing convention in our definition of b_t) that dividends are paid at the end of the time period t and that real interest rates r_t are measured from the beginning of t to the end of t.

[9] This data set is also used in Campbell and Shiller (1987) and is very similar to the data used by Shiller (1981); Mankiw, Romer, and Shapiro (1985); West (1988); and other contributors to the "volatility" literature. Kleidon (1986) also studied the Standard & Poors data from 1926, and Wilson and Jones (1987) have analyzed in some detail the properties of the pre-1926 Cowles data.

Table 1
Description of data sets

Cowles/S&P 500, 1871–1986

Nominal stock price: January S&P Composite Index, spliced to series in Cowles (1939). 1871–1986.

Nominal dividend: Total dividend per share accruing to index. 1871–1985.

Price deflator, model versions 1, 2, and 4: January Producer Price Index (annual average before 1900). 1871–1986.

Price deflator, model version 3: Consumption deflator, as used in Grossman and Shiller (1981). 1889–1985.

Nominal discount rate, model version 2: Annual return on four- to six-month prime commercial paper (six-month starting in 1979), rolled over in January and July. Interest rate data starting in 1938 are from the Board of Governors of the Federal Reserve System, with pre-1938 data from Macaulay (1938). 1871–1985.

Nominal discount rate, model version 3: Nominal growth in aggregate per capita consumption of nondurables and services, as used in Grossman and Shiller (1981). 1889–1985.

Real discount rate, model version 4: Squared ex post annual real return on the stock index. 1871–1985.

Value-weighted NYSE index, 1926–1986

Nominal stock price: January 1 value-weighted NYSE index, from CRSP Stock Index File. 1926–1986.

Nominal dividend: Value-weighted NYSE index, total dividends for year, from CRSP Stock Index File. 1926–1985.

Price deflator: January consumer price index, from Ibbotson Associates (1987). 1926–1986.

Nominal discount rate, model version 2: Annual return on one-month Treasury bills, rolled over monthly, from Ibbotson Associates (1987). 1926–1985.

Nominal discount rate, model version 3: Nominal growth in aggregate per capita consumption of nondurables and services, as used in Grossman and Shiller (1981). 1926–1985.

Real discount rate, model version 4: Squared ex post annual real return on the stock index. 1926–1985.

Lorie (1977)].[10] Corresponding monthly nominal Treasury bill rates and CPI inflation rates are from Ibbotson Associates (1987).

Although the raw CRSP data are available monthly, we follow Marsh and Merton (1987) and aggregate to an annual data interval. The main reason for doing this is that individual firms tend to change dividends no more frequently than once a year, so that aggregate dividends display seasonals within the year.[11]

In Table 2 we present some statistics that summarize the behavior of the nominal stock market variables in our two data sets. The most striking result in the table is the similarity in the period of overlap, 1926–1986. For example, the log dividend-price ratios on Cowles/S&P 500 stocks and the value-weighted NYSE index have a correlation of 0.985 over this period.

[10] These data have been used by Marsh and Merton (1987) in their study of aggregate dividend behavior, by Fama and French (1988) and Poterba and Summers (1987) in analyses of mean reversion in stock returns, and by many other writers in finance. CRSP data are also available for the equal-weighted New York Stock Exchange index. We obtained empirical results for this series, but they are qualitatively similar to those for the value-weighted index and to save space we do not report them here. These results are available from the authors on request.

[11] In aggregating the data to an annual interval, we assumed that dividends paid each month are accumulated through the year without receiving interest. The annual dividend is then the sum of monthly dividend payments, while the annual price is formed as the previous year's price times the one-year return excluding dividends, compounded monthly.

Table 2
Summary statistics for stock market data

Statistic	Data set and sample period			Correlation of Cowles/S&P and NYSE, 1926–1986
	Cowles/S&P, 1871–1986	Cowles/S&P, 1926–1986	NYSE, 1926–1986	
Δp_t:				
Mean	0.032	0.044	0.042	
Standard deviation	0.178	0.200	0.208	0.972
Δd_t:				
Mean	0.030	0.041	0.040	
Standard deviation	0.132	0.131	0.134	0.958
δ_t:				
Mean	−3.053	−3.121	−3.143	
Standard deviation	0.277	0.294	0.290	0.985

All variables in this table are nominal and measured annually. p_t is the log stock price, d_t is the log dividend, and δ_t is the log dividend-price ratio $d_{t-1} - p_t$.

It would seem that these indices reflect the same broad movements in prices and dividends.

In Table 3 we complete our review of the data by testing for unit roots in the various nominal and real series. This is important for two related reasons. First, the standard theory of inference in regressions with stochastic regressors requires that all variables be stationary. If we regress stock returns on variables with unit roots, the conventional standard errors may be seriously misleading. Second, when we use Equation (6) to characterize the behavior of the dividend-price ratio that is implied by a model of stock-price behavior, the results are likely to be sensitive to the stationarity assumption we make.

The test statistic used in Table 3 is one of a class recently proposed by Phillips (1987) and Phillips and Perron (1988). It is a modification of the F-statistic in the Dickey-Fuller (1981) regression of the change in a variable on a constant, a time trend, and the lagged level of the variable. Under the null hypothesis that the variable has a unit root, this regression has no explanatory power asymptotically since the change in the variable is stationary while the trend and level are not. However, the F-statistic has a nonstandard distribution, which is calculated numerically by Dickey and Fuller for the case where the change in the variable is white noise. Phillips and Perron's modification to the F-statistic enables one to apply the same distribution even when the change in the variable is serially correlated.

The main results in Table 3 are as follows. The null hypothesis of a unit root is generally not rejected for levels of prices and dividends, whether these are measured in nominal terms or real terms. The exception is that the null of a unit root can be rejected for the real dividend on Cowles/

Table 3
Univariate tests for unit roots

Variable	Data set and sample period	
	Cowles/S&P, 1871–1986	NYSE, 1926–1986
Nominal		
p_t	2.890	3.622
d_t	3.194	4.972
Δp_t	49.578***	22.969***
Δd_t	33.012***	14.342***
r_t	3.916	4.173
$\Delta d_t - r_t$	30.794***	13.410***
δ_t	10.711***	6.250**
π_t	28.442***	7.359**
Real		
p_t	4.124	2.670
d_t	7.089**	3.532
Δp_t	47.831***	25.436***
Δd_t	33.438***	17.422***
r_t	26.113***	5.346*

Variables are defined as follows: p_t is the log stock price, d_t is the log dividend, δ_t is the log dividend-price ratio $d_{t-1} - p_t$, π_t is the measured inflation rate, and r_t is the commercial-paper rate (1871–1986) or Treasury-bill rate (1926–1986).

This table presents tests of the null hypothesis that a series has a unit root. The test statistic is $Z\Phi_3$ from Phillips and Perron (1988) and as used in Perron (1988). The statistic is formed from the F-statistic in the regression $\Delta y_t = \mu + \beta t + \alpha y_{t-1}$, corrected for serial correlation in the equation error by using a fourth-order Newey-West (1987) correction. The critical values for the statistic are as reported in Fuller (1976): 1% 8.27 (*** in the table), 2.5% 7.16, 5% 6.25 (** in the table), and 10% 5.34 (* in the table).

S&P 500 stocks at the 5 percent level. This presumably reflects some negative autocorrelations in the growth rate of this series.[12] The unit root null is strongly rejected for growth rates of stock market variables and for the log dividend-price ratio.

The unit root tests in Table 3 are univariate and do not take account of the "adding up" constraint that the sum of stationary processes must be stationary. Accordingly, there are some internally inconsistent results in the table. For example, in the Cowles/S&P data it cannot be true that the log dividend-price ratio and the log real dividend are stationary while the log real price has a unit root. It is also inconsistent that inflation and real interest rates, and nominal dividend growth corrected or uncorrected for nominal interest rates, seem to be stationary in all data sets, but the unit root null is not rejected for nominal interest rates.

We proceed under the assumption that the log dividend-price ratio and growth rates of real dividends and prices are stationary, so that log divi-

[12] It is possible that the series is genuinely stationary around a trend, but it is also possible that our tests are falsely rejecting the null hypothesis because the series has negative autocorrelations in growth rates arising from an autoregressive-moving-average (ARMA) representation in growth rates with a large (but not unit) moving-average root. Such series tend to revert, not to a trend, but to a long moving average of their past history. Schwert (1988) presents Monte Carlo results in which false rejection of a unit root null occurs for this reason.

dends and prices are cointegrated processes.[13] Econometric techniques have been developed for processes of this sort by Engle and Granger (1987), Phillips and Durlauf (1986), and Stock (1987). Our model is particularly straightforward to deal with since the cointegrating vector is specified in the model and does not require estimation. Ordinary theory of estimation of stationary vector autoregressions is applicable here.

3. Vector Autoregressions and the Dividend-Ratio Model

In this section we propose a method for analyzing the historical movements of stock prices in relation to dividends and alternative measures of discount rates. The method is an extension of that developed in Campbell and Shiller (1987).

Our approach uses the dividend-ratio model derived in Section 1. As we showed there, a linear approximation to the log stock return implies that the log dividend-price ratio can be written as a discounted value of expected future dividend growth rates and discount rates. We would like to test the adequacy of some popular measures of the ex post discount rate on stock. We can do this by comparing the log dividend-price ratio with the forecast of dividend growth and discount rates obtained from an unrestricted econometric model: in practice, we will use a log-linear vector autoregression, or VAR. We can make the comparison in a formal statistical way, or informally by looking at the historical movements of these two variables.

Our approach is different from the regression methods commonly used to test a model of expected stock returns. The standard way to proceed would be to regress the one-period ex post stock return less the ex post discount rate, $h_t - r_t$, on a constant term and on some variables known at the start of period t. If the coefficients on these variables are jointly significant, then the model is rejected statistically.

Our approach has two potential advantages over the standard approach. First, it may have more power to detect long-lived deviations of stock prices from the "fundamental value" implied by the model. As Shiller (1984) and Summers (1986) have argued, single-period returns regressions have extremely low power against this alternative. Second, even if the regression approach does reject the model, the rejection can be hard to interpret. The regression results do not tell us whether the behavior of the dividend-price ratio is quite different from that implied by the model or whether it is rather similar. We do not know whether we reject the model because it is entirely wrong or because the dividend-price ratio is affected by some economically minor, but statistically detectable, factor. Our approach explicitly compares movements in the dividend-price ratio with the movements that are implied by the model.

[13] We note that this is a conservative assumption in the sense that it leads to greater variability in the rational forecast of expected future dividends, and less evidence of excess volatility in stock prices, than does the assumption that dividends and prices are stationary around a deterministic trend.

It might seem that our approach does not take proper account of extra information that market participants may have. If the market knows a great deal about the future paths of dividends and discount rates, then how can our unrestricted econometric model reproduce the market's forecast?

The answer is that we include the log dividend-price ratio itself as one variable in the vector autoregression. This enables the VAR to generate a forecast of dividends and discount rates that exactly equals the log dividend-price ratio. Intuitively, even though we do not observe everything that market participants do, we do observe the log dividend-price ratio, and that variable summarizes the market's relevant information.

Another way to make the same point is that those restrictions on the VAR which ensure that the unrestricted forecast equals the dividend-price ratio are algebraically equivalent to those which ensure that the one-period stock return cannot be predicted by the lagged variables included in the VAR. Adding extra variables cannot make stock returns unpredictable if they were found to be predictable using fewer variables; similarly, adding extra variables cannot make the log dividend-price ratio equal the unrestricted forecast of dividends and discount rates if these series differ when fewer variables are used.

In order to explain our approach precisely, we will now introduce some more notation. To keep the exposition as simple as possible, we will redefine all variables as deviations from means; this enables us to drop constant terms, which, as explained in Section 1, are not the focus of our inquiry and are not restricted by our model.

We assume that at the start of period t, market participants observe a vector of state variables y_t; their information set I_t is just the history $\{y_t, y_{t-1}, \ldots\}$. We assume that y_t follows a linear stochastic process with constant coefficients that are known to market participants. It follows from this assumption that any subset of the variables in y_t (where linear combinations of variables can be included) also follows a linear stochastic process with constant coefficients.

We now define a vector x_t that includes the variables in y_t that we econometricians observe. Our information set H_t is the history $\{x_t, x_{t-1}, \ldots\}$. A parsimonious choice for x_t is the vector $[\delta_t, r_{t-1} - \Delta d_{t-1}]'$, where we remove the means from the data since these are unrestricted, and we lag the growth-adjusted discount rate by one period to ensure that it is known to the market by the start of period t. This vector x_t is the smallest that allows us to test the restrictions of the dividend-ratio model. We could of course choose a larger vector [see, for example, Campbell and Shiller (1988a)], but as noted above this could only strengthen a rejection of the model.

We assume that the linear process for x_t can be written as a vector autoregression (VAR) with p lags: $x_t = C_1 x_{t-1} + C_2 x_{t-2} + \cdots + C_p x_{t-p} + u_t$, where C_i for $i = 1 \cdots$ p are each 2×2 matrices. Since p can be large, this assumption involves little further loss of generality. We will write the (j, k) element of C_i as C_{ijk}; thus, C_{ijk} is the coefficient of the jth variable in x_t on the kth variable lagged i times.

It will be convenient for us to rewrite the VAR in first-order, or companion, form along lines first suggested for rational expectations models by Sargent (1979). Doing this enables us to convert a pth-order autoregression into a first-order autoregression, for which the formula for conditional expectations has a simple form. This is done by defining a new vector z_t, which includes $2p$ rather than 2 elements: δ_t and $(p-1)$ lags, and $r_{t-1} - \Delta d_{t-1}$ and $(p-1)$ lags. For example, if $p = 2$, then we can write $z_t = [\delta_t, \delta_{t-1}, r_{t-1} - \Delta d_{t-1}, r_{t-2} - \Delta d_{t-2}]'$, $v_t = [u_{1t}, 0, u_{2t}, 0]'$. The vector z_t then follows a first-order VAR, where the rows corresponding to δ_t and Δd_{t-1} are stochastic and the others are deterministic:

$$\begin{bmatrix} \delta_t \\ \delta_{t-1} \\ r_{t-1} - \Delta d_{t-1} \\ r_{t-2} - \Delta d_{t-2} \end{bmatrix} = \begin{bmatrix} C_{111} & C_{211} & C_{112} & C_{212} \\ 1 & 0 & 0 & 0 \\ C_{121} & C_{221} & C_{122} & C_{222} \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \delta_{t-1} \\ \delta_{t-2} \\ r_{t-2} - \Delta d_{t-2} \\ r_{t-3} - \Delta d_{t-3} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ 0 \\ u_{2t} \\ 0 \end{bmatrix}$$

The VAR system can be written more parsimoniously as

$$z_t = Az_{t-1} + v_t \tag{7}$$

The vector z_t has the useful property that to forecast it ahead k periods, given our information set H_t, we simply multiply z_t by the kth power of the matrix A: $E[z_{t+k} \mid H_t] = A^k z_t$.

As a final investment in notation, we define a vector $e1$ such that $e1'z_t \equiv \delta_t$, and a vector $e2$ such that $e2'z_t \equiv r_{t-1} - \Delta d_{t-1}$. That is, $e1$ and $e2$ "pick out" the elements δ_t and $r_{t-1} - \Delta d_{t-1}$ from the vector z_t. In the example above, $e1 = [1\ 0\ 0\ 0]'$ and $e2 = [0\ 0\ 1\ 0]'$.

We are now in a position to state the restrictions of the dividend-ratio model (6) on the vector autoregression (7). We note first that a very weak implication of the dividend-ratio model is that the log dividend-price ratio δ_t should Granger-cause $r_t - \Delta d_t$. The reason is that δ_t embodies the market's information about the full vector of state variables y_t. Unless y_t contains only $r_{t-1} - \Delta d_{t-1}$ and its lags, so that market forecasts are based only on the history of this variable, δ_t will have incremental explanatory power for $r_t - \Delta d_t$. And we can rule out the case in which y_t contains only $r_{t-1} - \Delta d_{t-1}$ and lags by noting that it implies, counterfactually, that δ_t should be an exact linear function of current and lagged $r_{t-1} - \Delta d_{t-1}$. This point is discussed at greater length in Campbell and Shiller (1987).

Of course, the dividend-ratio model also imposes a tight set of cross-equation restrictions on the VAR. To derive these, we take expectations of Equation (6), conditional on the VAR information set H_t. The left-hand side is unchanged since δ_t is in H_t. The right-hand side becomes an expected discounted value conditional on H_t [since H_t is a subset of the market's information set I_t that defines the expectations in Equation (6)]. Thus, we have

$$\delta_t \simeq E\left[\sum_{j=0}^{\infty} \rho^j (r_{t+j} - \Delta d_{t+j}) \mid H_t\right] \equiv \delta_t' \tag{8}$$

where we have dropped constant terms, as discussed above. Equation (8) says that δ_t should equal the unrestricted VAR forecast of the discounted value of $r_{t+j} - \Delta d_{t+j}$, which we call δ'_t. We can rewrite this equation, using the multiperiod VAR forecasting formula given above, as

$$\delta_t = e1'z_t = \sum_{j=0}^{\infty} \rho^j e2'A^{j+1}z_t \equiv \delta'_t \tag{9}$$

Since Equation (9) is to hold for all realizations of z_t, we must have

$$e1' = \sum_{j=0}^{\infty} \rho^j e2'A^{j+1} = e2'A(I - \rho A)^{-1} \tag{10}$$

where the second equality follows by evaluating the infinite sum, noting that it must converge because the elements of z_t are stationary.

Equation (10) defines a set of $2p$ nonlinear restrictions on the VAR coefficients. These can be tested by using a nonlinear Wald test. If we write the estimated vector of VAR coefficients as γ, the estimated variance-covariance matrix of these coefficients as Θ, and the vector of deviations of the estimated system from the model as λ, then the Wald test statistic is $\lambda'(\partial\lambda/\partial\gamma'\Theta\partial\lambda/\partial\gamma)^{-1}\lambda$. Under the null hypothesis, it is distributed χ^2, with degrees of freedom equal to the number of restrictions (the number of elements of λ). In the case of Equation (10), $\lambda = e1' - e2'A(I - \rho A)^{-1}$ and has $2p$ elements. The derivatives of λ with respect to the VAR parameters can be calculated numerically.[14]

The more usual regression approach to the model can also be expressed as a Wald test of restrictions on the VAR. If we postmultiply the restrictions in Equation (10) by $(I - \rho A)$, we get $2p$ linear restrictions, one for each column of the matrix A:

$$e1'(I - \rho A) - e2'A = 0 \tag{11}$$

One can show that Equation (11) states the restrictions of the model in returns form: $E(\xi_t - r_t \mid H_t) = 0$, where ξ_t is as defined in Equation (2).[15] A Wald test of Equation (11) on the VAR, using the formula above with $\lambda = e1'(I - \rho A) - e2'A$, is numerically equivalent to a test that there are $2p$ zero coefficients when $\xi_t - r_t$ is regressed on the variables in z_t.

It is important to note, however, that a Wald test of the restrictions (10) is not equivalent to a test of (11), even though (10) and (11) are algebraically equivalent. Formally, the reasons are that (10) and (11) are related

[14] A similar approach can be used to calculate standard errors for diagnostic statistics, such as the correlation between δ'_t and δ_t. This correlation is a function of the VAR coefficients γ (and also of the variance-covariance matrix of the VAR explanatory variables, but we treat this as fixed). For any such function $f(\gamma)$, we can obtain a standard error as $\sqrt{\partial f/\partial\gamma'\Theta\partial f/\partial\gamma}$, where the derivatives are calculated numerically.

[15] The constant term c does not appear here because in this section we have defined all variables as deviations from their means. Means are not restricted by our model. To see how Equation (11) is equivalent to the restriction on returns, rewrite it as $e1'z_t - e1'\rho(Az_t) - e2'(Az_t) = 0$. Note that $Az_t = E_t z_{t+1}$. Then $e1'z_t = \delta_t$, $e1'\rho(Az_t) = \rho E_t\delta_{t+1}$, and $e2'(Az_t) = E_t(r_t - \Delta d_t)$. Combining these elements and using the definition of ξ_t in Equation (2) gives us the restriction on returns.

by a nonlinear transformation and that Wald tests are not invariant to such transformations. As we have already argued, it may be easier to detect departures from the null by looking at the behavior of the dividend-price ratio than by looking at one-period stock returns. In our empirical work we will present tests of both the dividend-price ratio restriction (10) and the one-period stock return restriction (11).

The approach outlined above is appealingly simple, since it involves a VAR with only two variables. It enables us to test a model of expected stock returns by using Equation (10) or (11) and to compute the implications of predictable excess returns for the log dividend-price ratio by using Equation (9).

Unfortunately, the two-variable approach does not allow us to judge the relative importance for the log dividend-price ratio of expectations of future dividends and discount factors, since it treats the discount rate adjusted for dividend growth as a single variable. To address this question, we need to expand the vector of variables we observe, x_t, to include Δd_{t-1} and r_{t-1} separately. Redefining z_t and A in the obvious manner, and defining $e1$, $e2$, and $e3$ to pick out δ_t, Δd_{t-1}, and r_{t-1}, respectively, we now have the following equation instead of (9):

$$e1'z_t = \sum_{j=0}^{\infty} \rho^j (e3' - e2')A^{j+1}z_t \tag{12}$$

or $\delta_t = \delta'_t \equiv \delta'_{rt} + \delta'_{dt}$. δ'_t is now defined to equal the right-hand side of Equation (12), while δ'_{rt} is the component of δ'_t that forecasts future discount rates and δ'_{dt} is the component that forecasts (the negative of) dividend growth rates: $\delta'_{rt} \equiv e3'A(I - \rho A)^{-1}z_t$, and $\delta'_{dt} \equiv -e2'A(I - \rho A)^{-1}z_t$.

We can use the three-variable system to see whether expectations of dividend growth, measured by δ'_{dt}, or of discount rates, measured by δ'_{rt}, have historically been more important in determining the dividend-price ratio. As discussed in the introduction, we can also construct an implicit long-term expected real return on stock by purging the log dividend-price ratio of the influence of expected future dividends. From Equation (4), if we remove the present value of expected future dividend growth, δ'_{dt}, from δ_t, what we are left with is a present value of expected future stock returns, $E_t \Sigma_{j=0}^{\infty} \rho^j h_{t+j}$. The sum of the weights in this expression is $1/(1 - \rho)$, so if we multiply by $(1 - \rho)$ we get a weighted average of expected future stock returns. This has the same form as Shiller's (1979) expression for a consol yield as a weighted average of expected future short-term interest rates. Accordingly, we will call $(1 - \rho)(\delta_t - \delta'_{dt})$ a long-term expected real stock return.[16]

The restrictions (12) can again be rewritten in returns form to get

[16] The discussion here omits constant terms. The long-term expected real stock return can be adjusted to have the correct mean by adding the unconditional mean log stock return.

$e1'(I - \rho A) - (e3' - e2')A = 0$, which corresponds to $E(\xi_t - r_t \mid H_t) = 0$, as before.

Our discussion so far has assumed that we observe the ex post discount rate r_t itself. But in version 3 of the model we observe instead real consumption growth Δc_t, which is related to r_t by $r_t = \alpha \Delta c_t$, where α is the coefficient of relative-risk aversion.[17] Similarly, in version 4 we observe V_t, the squared ex post stock return, and our model is that $r_t = \alpha V_t$. Since α is not known, but must be estimated from the data, our methods require some modification in this case.

Taking version 3 as an example, we define $x_t = [\delta_t, \Delta d_{t-1}, \Delta c_{t-1}]'$. We redefine z_t, A, $e1$, $e2$, and $e3$ appropriately ($e3$ picks out Δc_{t-1} from z_t). Then the model implies

$$\delta_t = \delta_t' \equiv \delta_{rt}' + \delta_{dt}' \equiv \alpha e3'A(I - \rho A)^{-1}z_t - e2'A(I - \rho A)^{-1}z_t \quad (13)$$

and

$$e1'(I - \rho A) - (\alpha e3' - e2')A = 0 \quad (14)$$

We can estimate α, the coefficient of relative-risk aversion, by using the restrictions (14). One might at first think that a unique value for α could be found by postmultiplying (14) by $A^{-1}e3$ and solving the resulting expression for α in terms of estimated coefficients. However, the restrictions (14) imply that α is overidentified. When $p > 1$, the matrix A is singular because of its special structure. Defining $e4$ as the vector that is 0 except for the second element, which is 1, then $(\rho e1' + \alpha e3' - e2' - e4')A = 0$.

Our approach was, instead, to use a method-of-moments estimator for α [Hansen (1982)]. Recall that we write γ as the vector of VAR coefficients, Θ as its variance-covariance matrix, and λ as the vector of deviations of the estimated VAR from the model. In this case, $\lambda = \lambda(\alpha, \gamma) = e1 - (\rho A)'e1 - A'(\alpha e3 - e2)$. We choose α to minimize the Wald test statistic for the model $\lambda(\alpha, \gamma)'(\partial\lambda/\partial\gamma'\Theta\,\partial\lambda/\partial\gamma)^{-1}\lambda(\alpha, \gamma)$. We do this in two steps, first evaluating the derivatives $\partial\lambda/\partial\gamma$ at $\alpha = 1$, and then evaluating them at the first round estimate of α. The minimized Wald statistic is distributed χ^2 under the null, with $3p - 1$ degrees of freedom.

The resulting estimate of α has the following interpretation. Equation (14) asserts that the prediction at time t of the linearized return ξ_t equals (a constant plus) α times the predicted change in log consumption. Our estimate of α is thus analogous to other estimates in the literature that rely on making forecast returns correspond to forecast changes in consumption. In Grossman and Shiller (1981), estimation of α along these lines was suggested (in the context of a plot of stock prices and their ex post rational

[17] Once again we have dropped constant terms. The implicit assumption here is that the consumption data for each year represent consumption on December 31 of the year. Thus, in January of each year (the month in which our stock-price data are drawn), Δc_{t-1} is known but Δc_t is not. There is no fully satisfactory way to handle the unit-averaged consumption data in the context of a theoretical model involving point-of-time consumption data without going to the continuous-time econometrics format, as in Grossman, Melino, and Shiller (1987). We did experiment with the model $r_t = \alpha \Delta c_{t+1}$, with results that are discussed below.

counterpart), but the discussion was couched in levels; the simple method used here of dealing with nonstationarity (dividing by lagged dividend) was not used, and formal estimation in such terms was not attempted.

4. Empirical Results

In this section we apply the methods worked out in the previous section to our two stock market data sets. We begin by analyzing versions 1 and 2 of the model—in which expected real stock returns and expected excess returns on stock over short debt, respectively, are constant through time—since in these versions no unknown parameter needs to be estimated from the VAR system.

We study version 1 of the model by using a two-variable system that includes the log dividend-price ratio and the real dividend growth rate. (In this version the dividend growth rate is the only component of the growth-adjusted discount rate that varies through time, so it is the only component that needs to be included in the VAR.) We study version 2 by using a similar two-variable system that includes the log dividend-price ratio and the return on short debt less the dividend growth rate. We then move to a three-variable system in which real dividend growth and the real return on short debt are included separately.

For each data set, the parameter ρ, which is taken as known in our analysis, was formed as the exponential of the difference between the sample average change in log dividends and the log of the sample average real return on stocks. The parameter ρ therefore differs slightly across data sets; it is 0.937 for the Cowles/S&P data and 0.933 for the NYSE data. In order to check for robustness, we also estimated models with ρ fixed at 0.90, 0.95, and 0.975. The results were very similar to those reported.[18]

Table 4 displays detailed results for version 1 of the model, based on a VAR system with a single lag ($p = 1$). While this lag length may be too short (we try longer lags below), it has the advantage that we can give full details of the results in a single table. We report results for the Cowles/S&P and New York Stock Exchange data in separate panels.

At the top of each panel of Table 4, we report regressions of the exact log stock return h_t and the approximate log stock return ξ_t on two variables known at the start of period t: the dividend-price ratio δ_t and the lagged dividend growth rate Δd_{t-1}. These are the variables that enter the first-order VAR system. If version 1 of our model is correct, so that expected real stock returns are constant through time, then the regression coefficients of returns on these variables should equal zero. We can test the

[18] One might think that changing the discount factor ρ would dramatically affect the standard deviation of the series δ'_t. The discount rate did indeed have an important effect in our earlier paper [Campbell and Shiller (1987)]. But recall that here δ_t and δ'_t are defined in log terms, so their movements represent percentage changes in the dividend-price ratio. As the discount rate changes, it affects both the absolute variability of the dividend-price ratio and its absolute mean level. The net effect on the percentage variability of the dividend-price ratio is small.

Table 4
Testing constant expected real returns

Dependent variable	Explanatory variable: δ_t	Explanatory variable: Δd_{t-1}	R^2	Joint significance of coefficients
Cowles/S&P, 1871–1986				
Regressions of returns on information:				
h_t	0.129	−0.013	0.043	0.078
	(0.057)	(0.121)		
ξ_t	0.141	−0.012	0.053	0.045
	(0.057)	(0.120)		
VAR estimation:				
δ_{t+1}	0.706	0.259	0.515	0.000
	(0.066)	(0.139)		
Δd_t	−0.197	0.231	0.227	0.000
	(0.039)	(0.083)		

Implications of VAR estimates:

$\delta'_t = \underset{(0.123)}{0.636}\delta_t - \underset{(0.106)}{0.097}\Delta d_{t-1}$

$\sigma(\delta'_t)/\sigma(\delta_t) = \underset{(0.124)}{0.637}$ corr $(\delta'_t, \delta_t) = \underset{(0.006)}{0.997}$

Significance level for Wald test that $\delta'_t = \delta_t$: 0.005

Dependent variable	Explanatory variable: δ_t	Explanatory variable: Δd_{t-1}	R^2	Joint significance of coefficients
Value-weighted NYSE, 1926–1986				
Regressions of returns on information:				
h_t	0.154	−0.385	0.108	0.031
	(0.087)	(0.213)		
ξ_t	0.170	−0.374	0.120	0.021
	(0.086)	(0.209)		
VAR estimation:				
δ_{t+1}	0.689	0.570	0.500	0.000
	(0.094)	(0.229)		
Δd_t	−0.187	0.157	0.246	0.000
	(0.047)	(0.115)		

Implications of VAR estimates:

$\delta'_t = \underset{(0.114)}{0.471}\delta_t + \underset{(0.131)}{0.109}\Delta d_{t-1}$

$\sigma(\delta'_t)/\sigma(\delta_t) = \underset{(0.114)}{0.470}$ corr $(\delta'_t, \delta_t) = \underset{(0.011)}{0.995}$

Significance level for Wald test that $\delta'_t = \delta_t$: 0.000

Variables are defined as follows: p_t is the log stock price, d_t is the log dividend, δ_t is the log dividend-price ratio $d_{t-1} - p_t$, h_t is the log one-period stock return, and ξ_t is the approximation of h_t defined in Equation (2). The variable δ'_t is the unrestricted forecast of the present value of future growth-adjusted discount rates from a VAR (equivalent to the negative of the present value of future dividend growth rates in this version of the model), defined in Equations (8) and (9). The VAR includes the log dividend-price ratio and the dividend growth rate.

Standard errors are reported in parentheses under each coefficient. The matrix of VAR coefficients is the matrix A defined in Equation (7), except for a sign switch in the off-diagonal elements due to the fact that results are reported for dividend growth rather than for the negative of dividend growth. The constant-expected-return model is tested in two different ways: by regressing the exact and approximate log stock returns h_t and ξ_t on the lagged variables that appear in the VAR and testing the joint significance of the coefficients, and by a VAR Wald test of the hypothesis that $\delta'_t = \delta_t$ [Equation (10)]. The approximate stock-return regression test is equivalent to a VAR Wald test of Equation (11).

model by testing the joint significance of these coefficients.[19] We use both exact and approximate returns as a way to evaluate the accuracy of our approximation. If the approximation error is constant, then the regression test of the model will give the same result whether exact or approximate returns are used.

The returns regressions show that stock returns are somewhat predictable. The lagged log dividend-price ratio has a positive effect on stock returns, while the lagged real dividend growth rate has a negative effect. This pattern holds across both data sets, although in the Cowles/S&P data the log dividend-price ratio is more highly significant, while in the NYSE data both variables are about equally significant. The two variables are jointly significant at about the 5 percent level in each data set, which rejects the constancy of expected real stock returns at this level. The results are similar whether exact or approximate returns are used, although the rejections are slightly stronger with approximate returns.

The ability of the dividend-price ratio to predict returns has been noted before [for example, by Shiller (1984) and by Flood, Hodrick, and Kaplan (1986)]. The special feature of our approach is that we can use the dividend-ratio model to compute the implications of this predictability for the behavior of the dividend-price ratio. To do this, we start from the VAR estimates, which are reported in Table 4 below the returns regressions. The coefficients reported are the elements of the matrix A (except that the off-diagonal elements have a sign switch because results are reported for dividend growth rather than for the negative of dividend growth). We use these coefficients, and Equation (9), to compute the variable δ'_t, a linear combination of the explanatory variables in the VAR.

If the constant-expected-return model were true, the variable δ'_t would place a unit weight on δ_t and a zero weight on Δd_{t-1}. In fact, we can reject at the 0.5 percent level the hypothesis that δ'_t equals δ_t.[20] In both data sets the weight of δ'_t on Δd_{t-1} is close to zero, as it should be under the null, but the weight on δ_t is considerably less than unity. In the Cowles/S&P data, for example, it is 0.636 with an asymptotic standard error of 0.13. This means that δ'_t, the unrestricted forecast of the present value of future real dividend growth, has a standard deviation about two-thirds that of the log dividend-price ratio δ_t; in other words, the log dividend-price ratio moves about 50 percent too much.

One way to understand this result is to consider what it means for the dividend-price ratio to have a positive effect on subsequent stock returns. The dividend-price ratio is high when prices are low, and the effect on returns implies that prices tend to rise subsequently. To eliminate the

[19] As we noted in the previous section, the regression test that uses the approximate stock return is numerically equivalent to a VAR Wald test of Equation (11). The reported standard errors and test statistics are not corrected for heteroskedasticity. We also computed standard errors using White's (1984) heteroskedasticity correction and found them to be similar or slightly smaller.

[20] We test this hypothesis by using a nonlinear Wald test of the restrictions given in Equation (10).

Table 5
Testing constant expected real and excess returns

	Model version 1 (constant expected real returns)			Model version 2 (constant expected excess returns)		
	Lag length			Lag length		
	1	3	5	1	3	5
Cowles/S&P, 1871–1986						
h_t regression test	0.078	0.056	0.083	0.100	0.007	0.014
ξ_t regression test	0.045	0.035	0.055	0.073	0.005	0.009
Test that $\delta'_t = \delta_t$	0.005	0.000	0.000	0.009	0.000	0.000
$\sigma(\delta'_t)/\sigma(\delta_t)$	0.637	0.370	0.382	0.674	0.434	0.402
	(0.124)	(0.129)	(0.073)	(0.111)	(0.143)	(0.060)
corr (δ'_t, δ_t)	0.997	0.837	0.326	0.999	0.856	0.431
	(0.006)	(0.246)	(0.565)	(0.004)	(0.200)	(0.583)
Value-weighted NYSE, 1926–1986						
h_t regression test	0.031	0.063	0.042	0.111	0.097	0.042
ξ_t regression test	0.021	0.043	0.024	0.081	0.069	0.023
Test that $\delta'_t = \delta_t$	0.000	0.000	0.000	0.005	0.001	0.000
$\sigma(\delta'_t)/\sigma(\delta_t)$	0.470	0.290	0.365	0.544	0.360	0.528
	(0.114)	(0.079)	(0.130)	(0.140)	(0.115)	(0.284)
corr (δ'_t, δ_t)	0.995	0.616	−0.089	1.000	0.119	−0.353
	(0.011)	(0.432)	(0.505)	(0.002)	(0.767)	(0.392)

Variables are defined as follows: δ_t is the log dividend-price ratio, h_t is the log one-period stock return, and ξ_t is the approximation to h_t defined in Equation (2). The variable δ'_t is the unrestricted forecast of the present value of future growth-adjusted discount rates from a VAR, defined in Equations (8) and (9). The VAR includes the log dividend-price ratio and the dividend growth rate (for model version 1) or the short-term real interest rate less the dividend growth rate (for model version 2).

The constant-expected-return model is tested in two different ways: by regressing the exact and approximate log stock returns h_t and ξ_t (or excess returns $h_t - r_t$ and $\xi_t - r_t$ in model version 2) on the lagged variables that appear in the VAR and testing the joint significance of the coefficients, and by a VAR Wald test of the hypothesis that $\delta'_t = \delta_t$ [Equation (10)]. The approximate stock-return regression test is equivalent to a VAR Wald test of Equation (11).

predictability of returns, the stock price would have to be less variable around the dividend.

It is striking that the model is rejected at about the 5 percent level when using a one-period-return regression test [equivalently, a Wald test of Equation (11) on the VAR], but is rejected at the 0.5 percent level when using a test of the hypothesis that $\delta'_t = \delta_t$ [a nonlinear Wald test of Equation (10) on the VAR]. Campbell and Shiller (1988a) relate this fact to the observation made by Fama and French (1988) and others that returns are more predictable over many periods than over a single period.

The results in Table 4 are conditional on the one-lag specification of the system (although, of course, adding lags in a fixed sample can never reduce the R^2 of the equation explaining stock returns). In the left-hand part of Table 5 (the "Model version 1" columns) we summarize the results for VAR lag lengths 1, 3, and 5. (In the pth-order VAR, the independent variables are $\delta_t, \ldots, \delta_{t-p+1}$ and $\Delta d_{t-1}, \ldots, \Delta d_{t-p}$.) In both data sets the second lag variables raise the R^2 for the dividend growth equation (not reported in the table) by at least 10 percentage points. Little further improvement occurs thereafter. The significance level at which we reject the constant-

expected-real-return model fluctuates between 2 and 10 percent when we use a one-period-return regression test, and it is always 0.5 percent or better when we test that $\delta'_t = \delta_t$. The result that the dividend-price ratio moves too much seems very robust; indeed, as the lag length increases, the ratio $\sigma(\delta'_t)/\sigma(\delta_t)$ tends to fall. It is estimated quite precisely, even in the high-order VAR systems.

The main effect of increasing lag length is that the estimated correlation of δ_t and δ'_t falls. It seems likely that the extremely high correlation in the first-order model is an artifact of the information set, which contains only δ_t and Δd_{t-1}. The variable Δd_{t-1} is not highly persistent or smooth, so δ'_t does not place a large weight on it and instead moves closely with δ_t. It is also possible that the higher-order models are picking up the tendency of real dividends to revert to a long average of past dividends; when we compute the correlation of δ'_t with a detrended real dividend, it rises with lag length. However, the correlation coefficients are very imprecisely estimated, so strong conclusions are unwarranted.

The estimates in Tables 4 and 5 are derived from vector autoregressions over the whole sample period for each data set. We also estimated VAR systems with one and two lags over subsamples 1871–1925, 1926–1955, and 1956–1986. The constant-real-returns model is rejected more strongly in the later subsamples. Thus, in the Cowles/S&P data, the approximate return regression test (with lag length 2) rejects at the 41 percent level in 1871–1925, the 28 percent level in 1926–1955, and the 7.1 percent level in 1956–86. The test that $\delta'_t = \delta_t$ rejects at the 29 percent level in 1871–1925, the 5.1 percent level in 1926–1955, and the 0.1 percent level in 1956–1986. In every case the ratio $\sigma(\delta'_t)/\sigma(\delta_t)$ is estimated to be less than 1.

In the right-hand part of Table 5 we move on to consider version 2 of the model, in which expected excess returns on stock over commercial paper or Treasury bills are constant through time. We begin with a two-variable system including δ_t and $r_{t-1} - \Delta d_{t-1}$; this has the advantage that the price deflator cancels from both variables, so our results are not dependent on the accuracy of the measured deflator.

The results for excess returns in Table 5 are qualitatively very similar to those for real returns. Excess stock returns are slightly less predictable than real returns in low-order systems, but as we increase lag length this difference disappears. Once again the test that $\delta'_t = \delta_t$ rejects more strongly than the return regression test, and the dividend-price ratio seems to move too much. The ratio $\sigma(\delta'_t)/\sigma(\delta_t)$ is never higher than 0.678 in any of the systems we estimate, and the difference between this ratio and unity is almost always statistically significant at the 5 percent level or better. The correlation between δ_t and δ'_t again falls with lag length, but it is imprecisely estimated.[21]

[21] When we estimated VAR systems over the subsamples used for model version 1, we again found the strongest evidence against the model in the period 1956–1986. The dividend-price ratio again appeared to "move too much" in every subsample.

Table 6
Testing constant expected excess returns by using interest rates and dividend growth rates

	Lag length	
	1	3
Cowles/S&P, 1871–1986		
h_t regression test	0.179	0.011
ξ_t regression test	0.135	0.008
Test that $\delta'_t = \delta_t$	0.019	0.000
$\sigma(\delta_t - \delta'_{dt})/\sigma(\delta_t)$	0.395	0.761
	(0.113)	(0.201)
$\sigma(\delta'_{rt})/\sigma(\delta_t)$	0.186	0.253
	(0.062)	(0.111)
corr $(\delta_t - \delta'_{dt}, \delta'_{rt})$	0.395	0.383
	(0.538)	(0.703)
Value-weighted NYSE 1926–1986		
h_t regression test	0.055	0.051
ξ_t regression test	0.036	0.032
Test that $\delta'_t = \delta_t$	0.003	0.000
$\sigma(\delta_t - \delta'_{dt})/\sigma(\delta_t)$	0.537	0.946
	(0.114)	(0.171)
$\sigma(\delta'_{rt})/\sigma(\delta_t)$	0.237	0.245
	(0.091)	(0.150)
corr $(\delta_t - \delta'_{dt}, \delta'_{rt})$	0.001	−0.707
	(0.434)	(0.346)

Variables are defined as follows: δ_t is the log dividend-price ratio, h_t is the log one-period stock return, and ξ_t is the approximation to h_t defined in Equation (2). The variable δ'_t is the unrestricted forecast of the present value of future growth-adjusted discount rates from a VAR, defined in Equations (8) and (9). The VAR includes the log dividend-price ratio, the dividend growth rate, and the short-term real interest rate. The variable δ'_{dt} is the unrestricted forecast of the present value of the negative of future dividend growth rates from the VAR, defined in Equation (12). The variable δ'_{rt} is the unrestricted forecast of the present value of future discount rates from the VAR, defined in Equation (12).

The constant-expected-excess-return model is tested in two different ways: by regressing the exact and approximate excess log stock returns $h_t - r_t$ and $\xi_t - r_t$ on the lagged variables that appear in the VAR and testing the joint significance of the coefficients, and by a VAR Wald test of the hypothesis that $\delta'_t = \delta_t$ [Equation (10)]. The approximate stock-return regression test is equivalent to a VAR Wald test of Equation (11).

The similarity between the results for excess and real stock returns in Table 5 suggests that time variation in short-term real interest rates is not particularly helpful in explaining the movements of the dividend-price ratio. In Table 6 we present some results, based on a three-variable system including real dividend growth and real interest rates separately, that confirm this view. Systems of lag length 1 and 3 are estimated; we do not go up to lag 5 since the number of variables in the system now grows more rapidly with lag length. For each data set, the table gives the following numbers: significance levels for the predictability of exact and approximate excess returns; a rejection significance level for the hypothesis that $\delta'_t = \delta_t$; the standard deviation of the implied long-term expected real return $\delta_t - \delta'_{dt}$ as a ratio to the standard deviation of δ_t; the standard deviation of the unrestricted forecast of the discounted value of future short-term real interest rates δ'_{rt} as a fraction of the standard deviation of δ_t; and the correlation

between $\delta_t - \delta'_{dt}$ and δ'_{rt}. If the predictability of real returns were entirely due to time-varying short-term real interest rates, so that excess stock returns were unpredictable, then we should not reject that $\delta'_t = \delta_t$, and we should find $\delta_t - \delta'_{dt}$ equal to δ'_{rt}, with the same standard deviation and a correlation of unity.

In fact, the hypothesis that $\delta'_t = \delta_t$ is strongly rejected. The results in Table 6 show that δ'_{rt} has a much smaller standard deviation than $\delta_t - \delta'_{dt}$ does (the difference in variability increases as we increase lag length) and that the correlation between the two variables is small. In about half the systems we estimate, in fact, the correlation is actually negative.

This finding can be traced back to the following features of the data. The ex post short-term real interest rate, used here as a measure of the discount rate r_t, is not highly variable; in the Cowles/S&P data, for example, it has a standard deviation of 0.091, while real dividend growth and the log dividend-price ratio have standard deviations of 0.278 and 0.132, respectively. In the NYSE data the standard deviation of the ex post real interest rate is even lower, at 0.045, while the other two series are about as variable as in the Cowles/S&P data. Furthermore, in the VAR systems we estimate, the real rate is forecastable largely because of its own serial correlation. We find that it is not even Granger-caused by the log dividend-price ratio at the 10 percent level. It seems that short-term real interest rates are not sufficiently variable, and do not have the appropriate correlation with stock prices, to explain big movements in the log dividend-price ratio.

In Table 7 we move on to evaluate real consumption growth and the volatility of stock returns as measures of time-varying discount rates on stock (versions 3 and 4 of the model, respectively). As discussed in the previous section, these versions of the model have a free parameter α which can be interpreted as the coefficient of relative-risk aversion and which we estimate from the unrestricted VAR coefficients by using the method of moments. The format of Table 7 is similar to that of Table 6, except that we report the estimate of α with its standard error.

The results in Table 7 are discouraging for the view that real consumption growth is an adequate measure of the one-period discount rate on stock. The estimates of the coefficient of relative-risk aversion always have the wrong sign, and this version of the model is rejected about as strongly as are previous versions in VAR systems that have more than one lag. Once again the expected present value of future discount rates, δ'_{rt}, has too little variability and a correlation with $\delta_t - \delta'_{dt}$ that falls with lag length.

Inspection of the equation for consumption growth in the underlying VAR system reveals why α is always estimated to be negative. In all the equations we estimate, the log dividend-price ratio Granger-causes consumption growth at conventional significance levels, but a high ratio at the start of a year forecasts low consumption growth over the year. This

Table 7
Testing consumption- and volatility-based models of the discount rate

	Model version 3 (consumption)		Model version 4 (volatility)	
	Lag length		Lag length	
	1	3	1	3
Cowles/S&P, 1871–1986				
Estimate of α	−2.191	−0.423	2.552	0.960
(standard error)	(1.399)	(0.818)	(2.549)	(0.985)
ξ_t regression test	0.072	0.038	0.170	0.096
Test that $\delta'_t = \delta_t$	0.012	0.000	0.039	0.007
$\sigma(\delta_t - \delta'_{dt})/\sigma(\delta_t)$	0.484	0.711	0.368	0.724
	(0.100)	(0.216)	(0.122)	(0.211)
$\sigma(\delta'_{rt})/\sigma(\delta_t)$	0.227	0.038	0.077	0.095
	(0.149)	(0.081)	(0.116)	(0.148)
corr $(\delta_t - \delta'_{dt}, \delta'_{rt})$	0.910	−0.024	0.339	0.660
	(0.073)	(0.788)	(1.415)	(0.525)
Value-weighted NYSE, 1926–1986				
Estimate of α	−3.423	−0.332	1.810	1.318
(standard error)	(1.926)	(1.067)	(2.614)	(1.261)
ξ_t regression test	0.259	0.046	0.052	0.162
Test that $\delta'_t = \delta_t$	0.168	0.000	0.000	0.000
$\sigma(\delta_t - \delta'_{dt})/\sigma(\delta_t)$	0.526	0.811	0.538	0.857
	(0.108)	(0.175)	(0.111)	(0.178)
$\sigma(\delta'_{rt})/\sigma(\delta_t)$	0.396	0.035	0.107	0.151
	(0.170)	(0.120)	(0.322)	(0.215)
corr $(\delta_t - \delta'_{dt}, \delta'_{rt})$	0.930	0.190	0.731	0.187
	(0.087)	(0.709)	(0.453)	(1.140)

Variables are defined as follows: δ_t is the log dividend-price ratio, h_t is the log one-period stock return, and ξ_t is the approximation to h_t defined in Equation (2). The variable δ'_t is the unrestricted forecast of the present value of future growth-adjusted discount rates from a VAR, defined in Equation (13). The VAR includes the log dividend-price ratio, the dividend growth rate, and the consumption growth rate (for model version 3) or the squared ex post stock return (for model version 4). The variable δ'_{dt} is the unrestricted forecast of the present value of the negative of future dividend growth rates from the VAR, and the variable δ'_{rt} is the unrestricted forecast of the present value of future discount rates from the VAR; both are defined in Equation (13).

In these versions of the model the discount rate includes a free parameter α. This is estimated from the VAR by using Equation (14) and a method-of-moments estimator. The model is tested in two ways: by using a Wald test of Equation (14) (ξ_t regression test) and by using a nonlinear Wald test of Equation (13) (test that $\delta'_t = \delta_t$).

means that low consumption growth is associated with a high one-period discount rate on stock, which requires α to be negative.[22]

The results for version 4 of the model are somewhat better, but still discouraging. The estimates of α now have the right sign, but they are very imprecisely estimated. The variable δ'_{rt} has very low variability, and in the

[22] We also estimated version 3 of the model assuming that consumption is measured at the beginning of the year, so that $r_t = \alpha \Delta c_{t+1}$. With this timing assumption, the estimate of α is no longer negative, but it is still insignificantly different from zero. The dividend-price ratio no longer Granger-causes the discount rate measure, and the model is more strongly rejected.

VAR forecasting equation for volatility, the squared ex post stock return is not Granger-caused by the log dividend-price ratio at the 10 percent level. The model is rejected almost as strongly as before in systems with more than one lag.

5. Conclusion

This article has examined time variation in corporate stock prices relative to dividends. As a framework for analyzing stock-price movements, we have proposed a dividend-ratio model that expresses the log dividend-price ratio as the rational expectation of the present value of future dividend growth rates and discount rates. We have used the equation in combination with a vector autoregression to break down movements in the log dividend-price ratio into components attributable to expected future dividend growth, expected future discount rates, and unexplained factors.

Our main results are three. First, there is some evidence that the log dividend-price ratio does move with rationally expected future growth in dividends. The log dividend-price ratio Granger-causes real dividend growth in all the systems we estimated, and the unrestricted forecast of the present value of future dividend growth rates from a VAR, δ'_{dt}, has a standard deviation that is generally about half that of the actual log dividend-price ratio δ_t. The correlation between δ'_{dt} and δ_t is extremely high in first-order VAR systems, falling dramatically as we increase the VAR lag length. It generally remains positive in high-order systems but is imprecisely estimated.

Second, the various measures of short-term discount rates that we used—short-term interest rates, consumption growth, and the volatility of stock returns themselves—are unhelpful in explaining stock-price movements. One of the weakest implications of the model is that the log dividend-price ratio should help forecast measured discount rates if, in fact, expectations of future discount rates drive stock prices. But neither short-term real interest rates nor squared ex post real stock returns are Granger-caused by the log dividend-price ratio at conventional significance levels. The log dividend-price ratio does Granger-cause real consumption growth, but the correlation between Δc_t and δ_t has the wrong sign. When we compute the rational expectation of the present value of future discount rates, δ'_{rt}, we find that it is far less variable than the component of the log dividend-price ratio that is not explained by dividends, $\delta_t - \delta'_{dt}$.

Third, there is substantial unexplained variation in the log dividend-price ratio. The unexplained part of δ_t is roughly equal to $\delta_t - \delta'_{dt}$, since measured discount rates contribute little to the explanation of δ_t, and the variable $\delta_t - \delta'_{dt}$ has a standard deviation about half that of δ_t.

To give an idea of what remains to be explained, we present an estimate of the long-term expected real stock return in Table 8. The variable reported is the unconditional mean log stock return plus $(1 - \rho)(\delta_t - \delta'_{dt})$. As discussed in Section 3, this variable is a weighted average of expected future discount rates on stock; it is the stock market equivalent of a consol

Table 8
Dividend-price ratio and long-term expected real stock return

Date	D_{t-1}/P_t	Long-term expected return	Date	D_{t-1}/P_t	Long-term expected return
1929	0.034	0.068	1958	0.047	0.081
1930	0.046	0.078	1959	0.033	0.068
1931	0.063	0.096	1960	0.031	0.061
1932	0.093	0.118	1961	0.033	0.061
1933	0.067	0.109	1962	0.028	0.056
1934	0.035	0.079	1963	0.033	0.060
1935	0.041	0.077	1964	0.030	0.059
1936	0.033	0.067	1965	0.029	0.055
1937	0.040	0.076	1966	0.029	0.054
1938	0.071	0.097	1967	0.034	0.061
1939	0.037	0.076	1968	0.029	0.057
1940	0.046	0.084	1969	0.028	0.052
1941	0.058	0.093	1970	0.033	0.059
1942	0.076	0.109	1971	0.033	0.061
1943	0.061	0.102	1972	0.029	0.056
1944	0.051	0.094	1973	0.026	0.049
1945	0.047	0.087	1974	0.034	0.059
1946	0.036	0.072	1975	0.053	0.083
1947	0.045	0.076	1976	0.041	0.081
1948	0.055	0.092	1977	0.038	0.074
1949	0.065	0.102	1978	0.048	0.081
1950	0.063	0.104	1979	0.052	0.089
1951	0.064	0.105	1980	0.050	0.089
1952	0.054	0.095	1981	0.045	0.085
1953	0.051	0.092	1982	0.052	0.087
1954	0.055	0.094	1983	0.048	0.088
1955	0.041	0.082	1984	0.042	0.080
1956	0.038	0.074	1985	0.045	0.081
1957	0.039	0.072			

This table reports the dividend-price ratio D_{t-1}/P_t for each year in our sample, together with an estimate of the required long-term real rate of return on stocks for that year. This estimate is formed as the unconditional mean real rate of return on stocks, plus $(1 - \rho)$ times $(\delta_t - \delta'_{dt})$, the component of the demeaned log dividend-price ratio that cannot be accounted for by expected dividend growth. The variable δ'_{dt} is the unrestricted forecast of the present value of the negative of future dividend growth rates from a VAR, defined in Equation (12). The estimated VAR includes two lags of the log dividend-price ratio and the log real dividend growth rate.

yield in the bond market. The estimate is based on a VAR system that includes two lags of the log dividend-price ratio and the real dividend growth rate. (The system is similar to the ones reported in Table 5.) Prices and dividends are taken from the CRSP value-weighted New York Stock Exchange index. After allowing for the lags in the model, our series runs from 1929 to 1985. For comparison, we also report the raw-dividend-price ratio D_{t-1}/P_t.

It is clear from Table 8 that the long-term expected real return on stock is highly variable (recall that it is not a one-period expected return, but a long-term average of expected returns). Also, it does not move in parallel with the short-term real interest rate. Short-term real rates were unusually low in the late 1970s, but there is no sign of matching behavior in $(1 - \rho) \cdot (\delta_t - \delta'_{dt})$—which, instead, was unusually low throughout the 1960s.

We have reached these conclusions by using a methodology that is significantly more general and robust than any previously available. In

particular, we have not assumed that dividends are stationary around a fixed time trend [as Shiller (1981) did] or that they follow a linear process with a unit root [as Mankiw, Romer, and Shapiro (1985) and West (1987, 1988) did]. Rather, we have been able to model dividends as a *log*-linear process with a unit root. This incorporates the geometric random walk model of Kleidon (1986) and LeRoy and Parke (1987) and the dividend-smoothing model of Marsh and Merton (1986, 1987) as special cases. We have also been able to incorporate some simple and popular models of discount rate variation into our analysis.

There are, of course, a number of caveats that need to be taken into account in evaluating our results. Our approach relies on the accuracy of a first-order Taylor approximation to the log return on stocks. This approximation is essential if we are to be able to solve forward for the price implications of a returns model. We have presented evidence in the Appendix that the approximation is quite accurate for our data, but approximation error could have some influence on our results.

Our measures of discount rates are applicable only under certain assumptions, which might be questioned. In one version of the model a constant risk premium is assumed; in another, a constant riskless rate—rather extreme assumptions. And of course, even under these assumptions, there is a question whether we measure the risk-free rate or the risk premium accurately with our data. Theory suggests that the discount rate should be a function of the leverage of the firm, which changes through time as the price of the firm changes [Black (1976)], a factor not taken into account in our analysis.

Throughout, we have also assumed that stable, linear stochastic processes drive log dividends, log prices, and discount rates. Some of our procedures are robust to failures of this assumption. Notably, when we test the hypothesis that dividends and measured discount rates can fully explain stock-price movements, our test is either a regression of ex post stock returns on information or a nonlinear test of algebraically equivalent restrictions. The return regression should give zero coefficients whether the information variables follow stable or unstable, linear or nonlinear processes. The nonlinear test will find that $\delta_t' = \delta_t$ if the return regression gives zero coefficients, so it is equally robust.

On the other hand, the qualitative comparison of δ_t and δ_t', which we use to characterize a failure of the null hypothesis, may be vulnerable to a misspecification of the processes for prices, dividends, and discount rates. However, we believe that this comparison is fairly reliable in our data, since we obtain similar results over various subsamples.

Our econometric methods are based on asymptotic distribution theory. It is possible that they are subject to some bias in finite samples. Mankiw and Shapiro (1986) and Stambaugh (1986) have studied bias in regression tests, while Flavin (1983), Kleidon (1986), and Mattey and Meese (1986) have pointed out that volatility tests can suffer from this problem. We have conducted our own Monte Carlo study of the methods used in this article

[Campbell and Shiller (1988b)]. While we do find some bias toward rejection of the model, and some downward bias in the estimates of $\sigma(\delta'_t)/\sigma(\delta_t)$, the bias is not sufficient to explain our empirical findings.

Our results are also conditional on an adequate specification of lag length in the VAR systems we estimate. We try lag lengths up to 5 in two-variable systems and 3 in three-variable systems. The main effect of altering lag length is that the correlation between δ_t and δ'_{dt} falls, particularly when we move from one to two lags. As argued above, we suspect that the extremely high correlation in the first-order system is to some extent spurious.

Our conclusions might also be affected by adding more lagged variables to our information set H_t. We note, however, that adding more variables could only increase the explanatory power of regressions explaining stock returns. Given the risk of overfitting regressions in finite samples, we have chosen to consider a relatively small set of information variables in this article. In Campbell and Shiller (1988a), we have extended the set of variables somewhat by including corporate earnings data. We obtain results for one-lag systems including earnings that are similar to the results reported here for higher-order systems.

There is an interesting parallel between our results and those of Fama and French (1988) and Flood, Hodrick, and Kaplan (1986). These authors find that stock returns are more highly predictable when measured over several years than when measured over one year. The predictability of returns seems to cumulate over time. Our dividend-ratio model can be seen as a way to compute the effects of single-period predictability of returns when they are cumulated over infinite time. We find that moderate predictability of one-year stock returns can have dramatic implications for the log dividend-price ratio. In particular, the log dividend-price ratio has a standard deviation that is at least 50 percent higher than it would be if stock returns were unpredictable.

Appendix: Approximation Error of the Dividend-Ratio Model

In this Appendix we evaluate in several different ways the approximation error in the dividend-ratio model, Equation (6). We do this by studying the error in the underlying Equations (2) and (4) and by directly comparing δ'_t as given in Equation (9) with an exact log dividend-price ratio in a simple model. The evaluations proceed as follows:

1. We compare exact return h_t to approximate return ξ_t.
2. We compare the log dividend-price ratio δ_t with the right-hand side of Equation (4), using a terminal condition.
3. We compare an exact log dividend-price ratio where the price is the present value of expected dividends with the approximate log dividend-price ratio δ'_t.
4. We compare regression results when h_t or ξ_t are regressed on information.

Table A1
Evaluation of approximation error: Log-linear approximation to stock return

Data set, sample period, and statistic	Variable: Exact return h_t	Approximate return ξ_t	Error, $\xi_t - h_t$
Cowles/S&P, 1871–1986:			
Mean	0.081	0.076	−0.005
Standard deviation	0.170	0.169	0.005
Correlation with exact return	1.0000	0.9996	−0.3108
Value-weighted NYSE, 1926–1986:			
Mean	0.089	0.081	−0.008
Standard deviation	0.201	0.199	0.006
Correlation with exact return	1.0000	0.9995	−0.3394

All variables in this table are nominal. h_t is the log one-period stock return, and ξ_t is the approximation to h_t defined in Equation (2).

The first, second, and fourth comparisons use actual data, the third uses a Monte Carlo experiment in which data are generated by a known vector autoregressive model. The first two comparisons relate to ex post data, and the second two relate to conditional expectations.

1. Table A1 compares the approximate nominal stock return ξ_t, defined in Equation (2), with the exact nominal return h_t.[23] The approximation error is quite small and, most important for our purposes, is almost constant. (Constant approximation error will not affect our results since our models do not restrict mean returns.)

2. Even though the approximation error is small for one-period returns, it might cumulate when we solve forward to obtain Equation (4). To check this possibility, we constructed an approximate log dividend-price ratio δ_t^A by using Equation (4), the time series of ex post stock returns h_t and dividend growth rates Δd_t, and a terminal condition $\delta_T^A = \delta_T$.[24] In Table A2 we compare this variable with the actual log dividend-price ratio δ_t. If our approximation held exactly in Table A1 (that is, if we had $\xi_t = h_t$ for all t), then there would also be no error in Table A2. In order to reduce the influence of the terminal condition, we use only the first 30 years of each sample in computing the summary statistics. The approximation error is again quite small and not highly variable.

3. For our Monte Carlo experiment, we first generated 1000 replications of the vector z_t by using a normal random-number generator and the estimated parameters of the one-lag vector autoregressive model estimated here in Table 4, top panel. We then computed the corresponding 1000

[23] Real stock returns are obtained by subtracting the inflation rate from nominal returns, and excess stock returns are obtained by subtracting ex post nominal discount rates from nominal returns. Therefore, the approximation error for these return concepts is the same as for nominal stock returns.

[24] This terminal condition is used only for evaluating the approximation in Equation (4), not in the empirical work of the article.

Table A2
Evaluation of approximation error: Log-linear approximation to dividend-price ratio

Data set, sample period, and statistic	Variable δ_t	δ_t^A	$\delta_t^A - \delta_t$
Cowles/S&P, 1871–1986:			
Mean	−3.007	−2.935	0.072
Standard deviation	0.242	0.238	0.011
Correlation with δ_t	1.000	0.982	0.044
Value-weighted NYSE, 1926–1986:			
Mean	−2.989	−2.877	0.112
Standard deviation	0.263	0.266	0.027
Correlation with δ_t	1.000	0.995	0.063

Summary statistics are computed only over the first 30 years of each sample. δ_t is the log dividend-price ratio, and δ_t^A is an approximate log dividend-price ratio constructed by using Equation (4), the time series of ex post stock returns h_t and dividend growth rates Δd_t, and a terminal condition $\delta_T^A = \delta_T$.

observations of δ_t' by using the right-hand side of Equation (9). The variable δ_t' is what the log dividend-price ratio should be if discount rates are constant through time (version 1 of our model) and if our approximation holds exactly.

If dividends are generated by the lognormal VAR system and if discount rates are constant through time, then, without using any approximation, the log dividend-price ratio should equal δ_t^{LN} as defined by

$$\delta_t^{LN} = -\log\left(\sum_{j=1}^{\infty} \exp(m_j)\right) \tag{A1}$$

where

$$m_j = e2'\left\{j\mu + (I - A)^{-1}(A - A^{j+1})(z_t - \mu) + 0.5\sum_{k=1}^{j}(I - A)^{-1}(I - A^k)\Omega(I - A^k)'(I - A)^{-1\prime}e2\right\}$$

Table A3
Monte Carlo evaluation of approximation error: Log-linear approximation to dividend-price ratio

Statistic	Variable: Exact log dividend-price ratio δ_t^{LN}	Approximate log dividend-price ratio δ_t'	Error, $\delta_t' - \delta_t^{LN}$
Mean	−2.867	−2.710	0.157
Standard deviation	0.179	0.176	0.003
Correlation with δ_t^{LN}	1.000	0.99997	−0.896

The results in this table are based on simulated, not actual, data. We generated 1000 replications of a vector z_t including the log dividend-price ratio and the dividend growth rate, by using a normal random-number generator and the estimated parameters of the one-lag vector autoregressive model estimated in Table 4, top panel. We then computed the corresponding observations of δ_t' by using the right-hand side of Equation (9). These are compared with observations of δ_t^{LN}, defined in Equation (A1). If our approximation (2) is accurate, then δ_t' should equal δ_t^{LN}.

where μ is defined as the mean of the vector $z_t = x_t$, where Ω is the variance matrix of the error term $u_t = v_t$, and where the other symbols are as described in the text of this article.[25]

It follows that we can test the accuracy of our approximation by comparing δ'_t with δ_t^{LN}. Comparisons of this sort are given in Table A3. The correlation of δ'_t with δ_t^{LN} is extremely high at 0.99997, and the other measures also show very close correspondence. There is thus no need to use the more complicated nonlinear expression (A1) instead of (9) for the log dividend-price ratio.

4. Even though ξ_t and h_t are highly correlated, it does not automatically follow that the approximation does not pose problems for our Wald tests of the dividend-ratio model. We therefore compared regressions of ξ_t and h_t on information. A detailed comparison is presented in Table 4; comparative significance levels for rejection of the dividend-ratio model are presented also in Tables 5 and 6. The significance tends to be slightly stronger for approximate returns, indicating that there is some correlation of the approximation error with the explanatory variables. However, the difference is small and should not affect the broad conclusions we get from our approximate model.

References

Black, F., 1976, "Studies of Stock Price Volatility Changes," *Proceedings of the Business and Economic Statistics Section, American Statistical Association,* 177–181.

Black, F., and M. Scholes, 1974, "The Effects of Dividend Yield and Dividend Policy on Common Stock Prices and Returns," *Journal of Financial Economics,* 1, 1–22.

Breeden, D. T., 1979, "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities," *Journal of Financial Economics,* 7, 265–296.

Campbell, J. Y., 1987, "Stock Returns and the Term Structure," *Journal of Financial Economics,* 18, 373–399.

Campbell, J. Y., and R. J. Shiller, 1987, "Cointegration and Tests of Present Value Models," *Journal of Political Economy,* 95, 1062–1088.

Campbell, J. Y., and R. J. Shiller, 1988a, "Stock Prices, Earnings, and Expected Dividends," *Journal of Finance,* 43, 661–676.

Campbell, J. Y., and R. J. Shiller, 1988b, "The Dividend Ratio Model and Small Sample Bias: A Monte Carlo Study," forthcoming in *Economics Letters.*

Cowles, A., 1939, *Common Stock Indexes* (2d ed.), Principia Press, Bloomington, Ind.

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

Engle, R. F., and C. W. J. Granger, 1987, "Cointegration and Error-Correction: Representation, Estimation and Testing," *Econometrica,* 55, 251–276.

Fama, E. F., and K. R. French, 1988, "Permanent and Temporary Components of Stock Prices," *Journal of Political Economy,* 96, 246–273.

[25] The first element δ_t in the vector z_t need not satisfy Equation (A1), of course, since the estimated VAR model from Table 4 does not exactly satisfy the assumptions of the nonlinear expectations model.

Fisher, L., and J. H. Lorie, 1977, *A Half Century of Returns on Stocks and Bonds,* Graduate School of Business, University of Chicago.

Flavin, M. A., 1983, "Excess Volatility in the Financial Markets: A Reassessment of the Empirical Evidence," *Journal of Political Economy,* 91, 929–956.

Flood, R. P., R. J. Hodrick, and P. Kaplan, 1986, "An Evaluation of Recent Evidence on Stock Market Bubbles," Working Paper 1971, National Bureau of Economic Research, Cambridge, Mass.

French, K. R., G. W. Schwert, and R. F. Stambaugh, 1987, "Expected Stock Returns and Volatility," *Journal of Financial Economics,* 19, 3–29.

Fuller, W. A., 1976, *Introduction to Statistical Time Series,* Wiley, New York.

Gordon, M. J., 1962, *The Investment, Financing and Valuation of the Corporation,* Irwin, Homewood, Ill.

Grossman, S. J., A. Melino, and R. J. Shiller, 1987, "Estimating the Continuous-Time Consumption-Based Asset-Pricing Model," *Journal of Business and Economic Statistics,* 5, 315–327.

Grossman, S. J., and R. J. Shiller, 1981, "The Determinants of the Variability of Stock Market Prices," *American Economic Review,* 71, 222–227.

Hall, R. E., 1988, "Intertemporal Substitution in Consumption," *Journal of Political Economy,* 96, 339–357.

Hansen, L. P., 1982, "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica,* 50, 1029–1054.

Hansen, L. P., and K. J. Singleton, 1983, "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns," *Journal of Political Economy,* 91, 249–265.

Ibbotson Associates, 1987, *Stocks, Bonds, Bills and Inflation: 1987 Yearbook,* Ibbotson Associates, Chicago.

Keim, D. B., and R. F. Stambaugh, 1986, "Predicting Returns in the Stock and Bond Markets," *Journal of Financial Economics,* 17, 357–390.

Kleidon, A. W., 1986, "Variance Bounds Tests and Stock Price Valuation Models," *Journal of Political Economy,* 94, 953–1001.

LeRoy, S. F., and W. R. Parke, 1987, "Stock Price Volatility: A Test Based on the Geometric Random Walk," unpublished paper, University of California, Santa Barbara.

LeRoy, S. F., and R. D. Porter, 1981, "The Present-Value Relation: Tests Based on Implied Variance Bounds," *Econometrica,* 49, 97–113.

Lucas, R. E., 1978, "Asset Prices in an Exchange Economy," *Econometrica,* 46, 1429–1445.

Macaulay, F., 1938, *Some Theoretical Problems Suggested by the Movements of Interest Rates, Bond Yields and Stock Prices in the United States Since 1856,* National Bureau of Economic Research, New York.

Mankiw, N. G., D. Romer, and M. D. Shapiro, 1985, "An Unbiassed Reexamination of Stock Market Volatility," *Journal of Finance,* 40, 677–687.

Mankiw, N. G., J. Rotemberg, and L. Summers, 1985, "Intertemporal Substitution in Macroeconomics," *Quarterly Journal of Economics,* 100, 225–251.

Mankiw, N. G., and M. D. Shapiro, 1986, "Do We Reject Too Often? Small Sample Properties of Tests of Rational Expectations Models," *Economics Letters,* 20, 139–145.

Marsh, T. A., and R. C. Merton, 1986, "Dividend Variability and Variance Bounds Tests for the Rationality of Stock Market Prices," *American Economic Review,* 76, 483–498.

Marsh, T. A., and R. C. Merton, 1987, "Dividend Behavior for the Aggregate Stock Market," *Journal of Business,* 60, 1–40.

Mattey, J., and R. Meese, 1986, "Empirical Assessment of Present Value Relations," *Econometric Reviews,* 5, 171–234.

Mehra, R., and E. Prescott, 1985, "The Equity Premium: A Puzzle," *Journal of Monetary Economics,* 15, 145–162.

Merton, R. C., 1980, "On Estimating the Expected Return on the Market," *Journal of Financial Economics,* 8, 323–361.

Newey, W. K., and K. D. West, 1987, "A Simple, Positive Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica,* 55, 703–708.

Perron, P., 1988, "Trends and Random Walks in Macroeconomic Time Series: Further Evidence from a New Approach," *Journal of Economic Dynamics and Control,* 12, 297–332.

Phillips, P. C. B., 1987, "Time Series Regression with a Unit Root," *Econometrica,* 55, 277–302.

Phillips, P. C. B., and S. N. Durlauf, 1986, "Multiple Time Series with Integrated Variables," *Review of Economic Studies,* 53, 473–496.

Phillips, P. C. B., and P. Perron, 1988, "Testing for Unit Roots in Time Series Regression," forthcoming in *Biometrika.*

Pindyck, R. S., 1984, "Risk, Inflation, and the Stock Market," *American Economic Review,* 74, 335–351.

Pindyck, R. S., 1986, "Risk Aversion and the Determinants of Stock Market Behavior," working paper 1921, National Bureau of Economic Research, New York.

Poterba, J. M., and L. H. Summers, 1986, "The Persistence of Volatility and Stock Market Fluctuations," *American Economic Review,* 76, 1142–1151.

Poterba, J. M., and L. H. Summers, 1988, "Mean Reversion in Stock Prices: Evidence and Implications," forthcoming in *Journal of Financial Economics.*

Sargent, T. J., 1979, "A Note on the Estimation of the Rational Expectations Model of the Term Structure," *Journal of Monetary Economics,* 5, 133–143.

Schwert, G. W., 1988, "Tests for Unit Roots: A Monte Carlo Investigation," forthcoming in *Journal of Business and Economic Statistics.*

Shiller, R. J., 1979, "The Volatility of Long-Term Interest Rates and Expectations Models of the Term Structure," *Journal of Political Economy,* 87, 1190–1219.

Shiller, R. J., 1981, "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?" *American Economic Review,* 71, 421–436.

Shiller, R. J., 1984, "Stock Prices and Social Dynamics," *Brookings Papers on Economic Activity,* 457–498.

Stambaugh, R. F., 1986, "Bias in Regressions with Lagged Stochastic Regressors," unpublished paper, Graduate School of Business, University of Chicago.

Stock, J. H., 1987, "Asymptotic Properties of Least Squares Estimates of Cointegrating Vectors," *Econometrica,* 55, 1035–1056.

Summers, L. H., 1986, "Does the Stock Market Rationally Reflect Fundamental Values?" *Journal of Finance,* 41, 591–601.

West, K. D., 1987, "A Specification Test for Speculative Bubbles," *Quarterly Journal of Economics,* 102, 553–580.

West, K. D., 1988, "Dividend Innovations and Stock Price Volatility," *Econometrica,* 56, 37–61.

White, H., 1984, *Asymptotic Theory for Econometricians,* Academic Press, Orlando, Fla.

Wilson, J. W., and C. P. Jones, 1987, "A Comparison of Annual Common Stock Returns: 1871–1925 with 1926–85," *Journal of Business,* 60, 239–258.