A SPECIFICATION TEST FOR SPECULATIVE BUBBLES*

KENNETH D. WEST

The set of parameters needed to calculate the expected present discounted value of a stream of dividends can be estimated in two ways. One may test for speculative bubbles, or fads, by testing whether the two estimates are the same. When the test is applied to some annual U. S. stock market data, the data usually reject the null hypothesis of no bubbles. The test is of general interest, since it may be applied to a wide class of linear rational expectations models.

I. INTRODUCTION

The seeming tendency for self-fulfilling rumors about potential stock price fluctuations to result in actual stock price movements has long been noted by economists. In a famous passage Keynes, for example, described the stock market as a certain type of beauty contest, in which judges try to guess the winner of the contest: speculators devote their "intelligence to anticipating what average opinion expects average opinion to be" [1964, p. 156]. In recent rational expectations work this possibility has been rigorously formalized, and the self-fulfilling rumors dubbed speculative bubbles [Blanchard and Watson, 1982; Shiller, 1978; Taylor, 1977; Tirole, 1982, 1985]. Recent attempts to detect such bubbles with formal statistical tests have, however, met with mixed success [Blanchard and Watson, 1982; Diba and Grossman, 1984; Flood and Garber 1980; Flood, Garber, and Scott, 1984; Hamilton and Whiteman, 1984].

One possible reason for the inability of the empirical tests to detect the bubbles so often described is that the tests have been few and not very powerful. This paper develops and applies a test for speculative bubbles that (a) allows for a wider class of bubbles than did Flood and Garber [1980] and Flood, Garber, and Scott [1984]; (b) is specifically designed to test against the alternative that bubbles are present, in contrast to the volatility tests of Shiller [1981a, 1981b] and Leroy and Porter [1981]; and (c) may be applied even if prices and dividends are nonstationary, again in contrast to the volatility tests and to the tests in Flood and Garber [1980] and Flood, Garber, and Scott [1984].

*I thank Fischer Black, Whitney Newey, Lawrence Summers, an anonymous referee, and participants in various seminars for helpful comments and discussions, and the National Science Foundation for partial financial support. Responsibility for remaining errors is mine. This paper was revised while I was a National Fellow at the Hoover Institution.

The Quarterly Journal of Economics, August 1987

The basic idea of the present paper's test is very simple and was suggested by the specification test of Hausman [1978]. The test compares two sets of estimates of the parameters needed to calculate the expected present discounted value (PDV) of a given stock's dividend stream, with expectations conditional on current and all past dividends. In a constant discount rate model the two sets are obtained as follows. One set may be obtained simply by regressing the stock price on a suitable set of lagged dividends. The other set may be obtained indirectly from a pair of equations. One of the pair is an arbitrage equation yielding the discount rate, and the other is the ARIMA equation of the dividend process. The Hansen and Sargent [1981] formulas, familiar from rational expectations tests of cross-equation restrictions, may be applied to this pair of equations' coefficients to obtain a second set of estimates of the expected PDV parameters.

Under the null hypothesis that the stock price is set in accord with a standard efficient markets model [Brealey and Myers, 1981, pp. 42–45], the regression coefficients in all equations may be estimated consistently. When the two sets of estimates of the expected PDV parameters are compared, then, they should be the same, apart from sampling error.

But this equality of the two sets will not hold under the alternative hypothesis suggested by, e.g., Blanchard and Watson [1982], that the stock price equals the sum of two components: the price implied by the efficient markets model and a speculative bubble. In this case, the equation that relates price to a suitable set of dividends omits a relevant regressor—the bubble. As long as the bubble is correlated with the included regressors, the coefficients in this equation will be estimated inconsistently. The bubble will not, however, cause estimation of the other two equations to be inconsistent. So the coefficients in this pair of equations, as well as the implied value of the set of expected PDV parameters, will still be estimated consistently. Therefore, when the two estimates of the set of expected PDV parameters are compared, the two will be expected to be different.

Speculative bubbles are tested for, then, by seeing whether the two sets of estimates are the same, apart from sampling error. I check for the equality of the two sets in long-term annual data on the Standard and Poor's 500 index (1871–1980) and the Dow Jones index (1928–1978). The data reject the null hypothesis of no bubbles. The rejection appears to result at least in part because the coefficients in the regression of price on dividends are biased upwards. As is explained in Section II, this is precisely what would be expected if, as is sometimes argued [Shiller, 1984], bubbles reflect an overreaction by the market to news about dividends. A small amount of investigation of a linearized time varying discount rate model suggests that such variation may also help explain the results.

Section II quickly reviews the standard constant discount rate efficient markets model and the definition of a speculative bubble and then explains how the test is performed. Section III presents empirical results from a constant discount rate model and then develops and applies the specification test for a linearized time varying discount rate model. Section IV discusses the empirical results. Some econometric and algebraic details are in an appendix available from the author.

II. The Model and Test

According to a standard efficient markets model, a stock price is determined by the arbitrage relationship (1) [Brealey and Myers, 1981, pp. 42–45]:

$$p_t = bE(p_{t+1} + d_{t+1})|I_t, \tag{1}$$

where p_t is the real stock price in period t, b the constant ex ante real discount rate, $0 < b = 1/(1 + r) < 1$, r the constant expected return, E denotes mathematical expectations, assumed to be equivalent to linear projections, d_{t+1} the real dividend paid to the owner of the stock period $t + 1$, and I_t information common to traders in period t. I_t is assumed to contain, at a minimum, current and past dividends, and, in general, other variables that are useful in forecasting dividends. Time variation in the ex ante discount rate b is briefly considered in subsection III.D.

Equation (1) may be solved recursively forward to get

$$p_t = \sum_{1}^{n} b^i Ed_{t+i}|I_t + b^n Ep_{t+n}|I_t. \tag{2}$$

If the transversality condition

$$\lim_{n\to\infty} b^n Ep_{t+n}|I_t = 0 \tag{3}$$

holds, then $p_t = p_t^*$, where

$$p_t^* = \sum_{1}^{\infty} b^i Ed_{t+i}|I_t. \tag{4}$$

Now, the p_t^* defined in (4) is the unique forward solution to (1) as long as the transversality condition (3) holds. But if this condition fails, there is a family of solutions to (1) [Blanchard and Watson, 1982; Shiller, 1978; Taylor, 1977]. Any p_t that satisfies

$$(5) \qquad p_t = p_t^* + c_t, \qquad Ec_t | I_{t-1} = b^{-1} c_{t-1},$$

is also a solution to (1). c_t is by definition a speculative bubble, an otherwise extraneous event that affects stock prices because everyone expects it to do so. An example of a stochastic process for c_t, similar to one described in Blanchard and Watson [1982], is

$$(6) \qquad c_t = \begin{cases} (c_{t-1} - \bar{c})/(\pi_t b) & \text{with probability } \pi_t \\ \bar{c}/[(1 - \pi_t)b] & \text{with probability } 1 - \pi_t \end{cases}$$

$$0 < \pi_t < 1, \qquad \bar{c} > 0.$$

According to (6), strictly positive bubbles grow and pop. In this example, the probability that a bubble grows is π_t, that it collapses is $1 - \pi_t$. The bubble may be intimately connected with fundamentals, with π_t dependent on news about fundamentals. A simple example is $\pi_t = ½$ for all t, with the bubble popping if and only if the innovation in dividends is negative. If π_t is constant ($\pi_t = \pi$ for all t), each bubble has an expected duration of $(1 - \pi)^{-1}$. (π is not an identifiable parameter.) Combination of several bubbles are possible, each with a different π_t and $\bar{c}$; the growth and collapse of the bubbles may be either tightly or loosely related. See Blanchard and Watson [1982] for further examples and discussion.

Our aim is to test $p_t = p_t^*$ versus $p_t = p_t^* + c_t$, for some nontrivial c_t (possibly one not following the stochastic process (6)). Consider first this wildly implausible case: (a) there is no doubt that p_t and d_t are such that equations (1) and (2) hold. (b) d_t is a zero mean white noise process. Then $Ed_{t+i}|I_t = 0$ for $i > 0$, and $p_t^* = 0$ for all t. It follows from equations (1) to (4), then, that $p_t = 0$ for all t if equation (3) holds: given that the stochastic difference equation (1) is solved in the forward direction (2), the terminal condition (3) insures that (4) is the unique solution to equation (1), for all t. In this blissfully simple environment where (a) there is no doubt about the rational expectations, constant discount rate specification, and (b) no statistical inference is necessary, then (c) the null hypothesis that there are no bubbles should be rejected if $p_t = 0$ for some t.

The basis of the empirical work in this paper is the simple logical proposition illustrated in the previous paragraph: if a univariate stochastic difference equation is solved in the forward direction, a single terminal condition ties down a unique solution. Let us now allow for (a) uncertainty about b and the parameters of the dividend process; (b) the possibility that dividends are an endogenous variable, e.g., because they are smoothed by management; (c) uncertainty about whether the rational expectations, constant discount rate specification (1) really characterizes the data.

(a) Suppose that the actual value of b is not known. In addition, suppose that it is known that dividends follow a zero mean, AR(1) process,

$$d_t = \phi d_{t-1} + v_t. \tag{7}$$

In (7), $|\phi| < 1$, and v_t is a finite variance white noise process. The value of ϕ is not known. It is easy to verify that $\Sigma_1^\infty b^i E d_{t+i}|I_t = \delta_1 d_t$, $\delta_1 = b\phi/(1 - b\phi)$. So if $p_t = p_t^*$,

$$p_t = \delta_1 d_t. \tag{8}$$

The logical proposition described above is applied in this environment by estimating (1), (7), and (8). Equations (7) and (8) may be estimated by OLS, yielding point estimates $\hat{\phi}$ and $\hat{\delta}_1$. Equation (1) may be estimated by rewriting it as

$$\begin{aligned} p_t &= b(p_{t+1} + d_{t+1}) - b[p_{t+1} + d_{t+1} - E(p_{t+1} + d_{t+1}|I_t)] \\ &= b(p_{t+1} + d_{t+1}) + u_{t+1}. \end{aligned} \tag{1$'$}$$

An instrumental variables estimator, using as instruments variables known at time t—say d_t—will now produce a $\hat{b}$ that is a consistent estimate of b.

To apply the specification test, we compare two estimates of δ_1, the parameter needed to calculate $\Sigma_1^\infty b^i E d_{t+i}|I_t$. That is, we test H_0: $\hat{\delta}_1 = \hat{b}\hat{\phi}/(1 - \hat{b}\hat{\phi})$, and reject the null hypothesis only if the resulting test statistic exceeds an appropriate critical value.

(b) Allowing for endogeneity of dividends [Marsh and Merton, 1984] causes no substantial complications. Let H_t be the set consisting of a constant and current and lagged dividends, $H_t = \{1, d_{t-i}|i \geqq 0\}$. Since H_t is a subset of I_t, equation (4) in conjunction with $p_t = p_t^*$ implies [Hansen and Sargent, 1981] that

$$\begin{aligned} p_t &= \sum_1^\infty b^i E d_{t+i}|H_t + z_t, \\ z_t &= \sum_1^\infty b^i (E d_{t+i}|I_t - E d_{t+i}|H_t), \\ &\quad z_t \text{ serially correlated in general,} \\ &E x_s z_t = 0 \text{ for } x_s \text{ an element of } H_t. \end{aligned} \tag{9}$$

To apply the specification test, it is necessary to turn (9) into a regression equation. This can be done conveniently if there is a closed-form expression for $\Sigma_1^\infty b^i Ed_{t+i}|H_t$. Now, $Ed_{t+i}|H_t$ is by definition the forecast of dividends given the past history of dividends. If d_t is stationary, perhaps after differencing, $Ed_{t+i}|H_t$ may be calculated as the usual ARIMA forecast of d_{t+i}. And if d_t is stationary, possibly after differencing, there is a closed-form expression for $\Sigma_1^\infty b^i Ed_{t+i}|H_t$ in the form of a distributed lag on current and past d_t [Hansen and Sargent, 1981]. As in the simple example (7) and (8), the coefficients of the distributed lag are functions of b and the parameters of d_t's univariate ARIMA process. Exact formulas are given in subsection III.A.

When dividends are endogenous and are characterized by an ARIMA process of known order (but unknown parameters), the test can proceed essentially as just described in case (a) above: estimate (1′) by instrumental variables; estimate d_t's univariate ARIMA equation; estimate a distributed lag of p_t on d_t; compare the estimates of the parameters of the distributed lag with those of (1′) and d_t's ARIMA equation. (Actually, if differencing is required to induce stationarity in d_t, it is more convenient to estimate a distributed lag of a difference of p_t on a difference of d_t. See subsection III.A.) So the basic difference from case (a) is that it is acknowledged that d_t's ARIMA equation is simply a convenient way to forecast dividends, and not a statement about the exogeneity of dividends.

It still remains to determine the order of the ARIMA process for d_t. To make the results of as general interest as possible, the empirical work does not assume any particular structural model for dividends. The order of the ARIMA process for d_t is data rather than theoretically determined, in the spirit of the usual Box-Jenkins [1970] analysis. Consistent with such an approach, a variety of ARIMA specifications are tried, to make sure that the results are not sensitive to the exact specification chosen.

It is to be noted that this discussion assumes that arithmetic differencing is sufficient to induce stationarity in d_t. This is because such a condition makes it possible to obtain a closed-form solution to $\Sigma_1^\infty b^i Ed_{t+i}|H_t$. While the usual Box-Jenkins [1970] diagnostics suggest that arithmetic differences suffice to induce stationarity in the data used in this paper (see subsection III.B), much research in finance assumes that log differences are required [Kleidon, 1985]. Since it is also possible to obtain a closed-form expression for $\Sigma_1^\infty b^i Ed_{t+i}|H_t$ when d_t follows a lognormal random walk [Kleidon, 1985], the empirical work (in subsection III.C) briefly considers this specification as well.

(c) Suppose that the specification test described in case (b) indicates that the difference between the two sets of estimates of the parameters needed to calculate $\Sigma_1^\infty b^i Ed_{t+i}|H_t$ is unlikely to result solely from sampling error. Clearly, this can happen for many reasons, in addition to the presence of bubbles.

The possibility that a discrepancy between the two sets of parameter estimates results from certain factors other than bubbles is handled in two ways. In subsection III.D a model with time varying discount rates is linearized as in Shiller [1981a]. It is shown that in such a model one can apply a somewhat more complicated version of the test just described.

The second way that shortcomings of the present value model are considered is by applying diagnostic tests to the estimates of (1′). The diagnostic tests are chosen in light of two alternatives that have figured prominently in related work, that expectations are not rational [Ackley, 1983; Shiller, 1984] and that discount rates are time varying [Leroy, 1984]. The particular tests used are described in Section III. The greater the extent to which these diagnostics suggest that equation (1) is consistent with the data, the more plausible it is to discount expectational irrationality and discount rate variation as significant sources of a discrepancy between the two sets of parameter estimates.

To sum up: the specification test proceeds by estimating (1′), a variety of specifications for the univariate ARIMA process for d_t, and for each such specification, the corresponding distributed lag of p_t on d_t. It applies a battery of diagnostic tests to equation (1′), to see whether equation (1) appears to be consistent with the data. For each specification of the dividend ARIMA process, it applies diagnostics of the sort often used in ARIMA estimation to check whether each specification seems to adequately capture the dynamics of the d_t process. The test then uses each estimate of (1′) and the parameters of the d_t process to calculate an implied value of the parameters that characterize the expected present discounted value of d_t, conditional on current and lagged d_t. It compares these implied values to the estimates directly obtained by a distributed lag regression of p_t on d_t. One possible explanation of any difference between the two sets of estimates is bubbles. This explanation is more compelling, the less likely is the difference to result from sampling error, and the greater the extent to which the diagnostic tests fail to reject (1′) and the specification of the univariate dividend process.

Four final comments are of interest before the empirical work is presented. The first comment concerns how reasonable it is to use the past history of the dividend process to forecast future dividends. It clearly is not reasonable at all in everyone's favorite example of a corporation that has yet to pay out any dividends. It also may not be reasonable if there is a "peso problem" and market participants are rationally considering a small probability event that has not occurred in the sample. There are three points to make. The first is that the best protection against such a problem is to use a long sample period, which is what I did. The second is that certain forms of the peso problem in fact are implicitly allowed under the null, by suitably reinterpreting the parameter b [Shiller, 1981b]. Finally, I tested for the stability of the dividend process; this can detect in-sample switches of the dividend process.

The second concerns the distribution of the estimates of the distributed lag of p_t on d_t when there is a bubble. This is conveniently illustrated when the univariate dividend process is as in (7). Then $p_t = \delta_1 d_t + z_t + c_t$, z_t defined in equation (9). When p_t is regressed on d_t, we have

$$\begin{aligned} \hat{\delta}_1 &= (T^{-1}\Sigma\, d_t^2)^{-1}(T^{-1}\Sigma\, d_t p_t) \\ &= \delta_1 + (T^{-1}\Sigma\, d_t^2)^{-1}(T^{-1}\Sigma\, d_t z_t) + (T^{-1}\Sigma\, d_t^2)^{-1}(T^{-1}\Sigma\, d_t c_t) \\ &\rightarrow p\lim \hat{\delta}_1 = \delta_1 + p\lim\,(T^{-1}\Sigma\, d_t^2)^{-1}(T^{-1}\Sigma\, d_t c_t). \end{aligned} \tag{10}$$

(Recall that $Ed_t z_t = 0$ by construction.) The asymptotic bias in $\hat{\delta}_1$, then, is equal to the asymptotic value of the coefficient of a regression of the bubble on d_t. An additional check on the plausibility of bubbles as the source of any discrepancy of the two estimates of δ_1 comes from looking at the value of the estimate of δ_1 that comes from the regression of p_t on d_t. It is often argued that bubbles result at least in part from an overreaction to news about fundamentals [Shiller, 1984]. If bubbles are present, then, one would expect the point estimate of δ_1 to be biased upwards. More generally, when $\Sigma_1^\infty b^i E d_{t+i}|H_t$ involves more than one lag of d_t, one might expect bubbles to cause the sum of coefficients in the distributed lag projection of p_t onto d_t to be biased upwards.[1]

The third comment is that this test has a substantial advantage over the tests undertaken in Flood and Garber [1980] and Flood,

1. Unfortunately, the discussion in this paragraph cannot, in general, be justified rigorously. In at least certain cases, $T^{-1}\Sigma\, d_t c_t$ will not converge in mean square to a constant. This results from the fact that c_t is growing on average at a rate faster than T^{-1}, i.e., at rate b^{-1}. This is briefly discussed in West [1985], as are the implications of the explosive growth of c_t for the distribution of $\hat{\delta}_1$ under the alternative that bubbles are present.

Garber, and Scott [1984], and that proposed in Sargent and Wallace [1984]. This is that the specification test does not require parametric specification of the bubble process. *Any* bubble that is correlated with dividends can be detected: the bubble described in (6); a bubble as in (6) whose probability of continuing to float π_t depends stochastically on events such as, say, money supply news, or GNP growth, or political events; and combinations of any and all such bubbles.

The fourth comment is that the specification test can be used to test for bubbles in other infinite horizon linear rational expectations models. The idea is to compare two sets of estimates. One set is obtained from the dynamic programming, or equilibrium, solution to the model (i.e., from the model's analogue to equation (12a) or (12b) below). The second set is obtained by applying the relevant Hansen and Sargent [1981] formulas to estimates obtained from two types of equations. The first is the model's Euler equations, or first-order conditions (i.e., the model's analogue to equation (1)). The second is ARIMA equations for the model's forcing variables (i.e., the model's analogue to equation (11a) or (11b) below). The null hypothesis of no bubbles should be rejected only if (a) diagnostic tests on the Euler and ARIMA equations suggest that these equations are acceptably specified, and (b) any difference between the two sets of estimates is unlikely to result from sampling error.

III. Empirical Results

Subsection A describes data and estimation technique. Subsection B presents empirical results. Subsection C extends the specification test to allow for a dividend process that follows a lognormal random walk. Subsection D extends it to test a model that allows discount rates to vary over time.

A. Data and Estimation Technique

The data used were those used by Shiller [1981a] in his study of stock price volatility, and were graciously supplied by him. There were two data sets, both containing annual aggregate price and dividend data. One had the Standard and Poor 500 for 1871–1980 (p_t = price in January divided by producer price index (1979 = 100), d_{t+1} = sum of dividends from that same January to the following December, deflated by the average of that year's producer price index). The other data set was a modified Dow Jones index, 1928–1978 (p_t,d_{t+1}, as above). See Shiller [1981a] for a discussion of the data.

Let me describe the following in turn: (i) identification of the order of d_t's ARIMA process; (ii) estimation of (1′), the d_t process, and the distributed lag of p_t on d_t; (iii) calculation of the variance-covariance matrix of the parameters; (iv) calculation of the basic test statistic; (v) diagnostic tests performed on the equations estimated.

(i) For each data set estimation was done with d_t in levels and with d_t in arithmetic first differences. In each case, only pure autoregressions were estimated, for computational simplicity:

$$d_{t+1} = \mu + \phi_1 d_t + \ldots + \phi_q d_{t-q+1} + v_{t+1} \tag{11a}$$

$$\Delta d_{t+1} = \mu + \phi_1 \Delta d_t + \ldots + \phi_q \Delta d_{t-q+1} + v_{t+1}. \tag{11b}$$

For each data set and for both d_t and Δd_t, two different values of the lag length q were used. One was arbitrarily selected as $q = 4$. The other was selected by the information criterion of Hannan and Quinn [1979]. This criterion chooses the value of q that minimizes a certain function of the estimated parameters, and asymptotically chooses the correct q if the process truly has a finite order autoregressive representation.[2] Thus, for each data set, up to four specifications were estimated: differenced and undifferenced, $q = 4$, and $q =$ lag length selected by the Hannan and Quinn [1979] criterion. In one case (Dow Jones, differenced) the Hannan and Quinn [1979] criterion chose $q = 4$. So for the Dow Jones, only three specifications were estimated.

(ii) If $d_t \sim AR(q)$, as in (11a), then

$$\text{(12a)} \qquad p_{t+1} = m + \delta_1 d_{t+1} + \ldots \delta_q d_{t-q+2} + w_{t+1}$$

$$m + \delta_1 d_{t+1} + \ldots + \delta_q d_{t-q+2} = \sum_1^\infty b^i E d_{t+i+1} | H_{t+1}$$

$$w_{t+1} = z_{t+1} + c_{t+1}$$

$$z_{t+1} = \sum_1^\infty b^i (E d_{t+i+1} | I_{t+1} - E d_{t+i+1} | H_{t+1}).$$

The formulas linking m and the δ_i, on the one hand, b, μ, and the ϕ_i, on the other, under the null, are given in equation (13a) below.

2. The Hannan and Quinn [1979] procedure selects the q that minimizes

$$\ln \hat{\sigma}_v^2 + T^{-1} 2qk \ln \ln T, \ \hat{\sigma}_v^2 = T^{-1} \Sigma_{t=1}^T \hat{v}_t^2,$$

for $q < Q$ for some fixed Q, with $k > 1$. I set $Q = 4$, $k = 1.001$.

If $\Delta d_t \sim AR(q)$, as in (11b), then projecting a first difference of $E\Sigma_1^\infty b^i d_{t+i+1}|I_{t+1}$ onto H_t yields

$$\Delta p_{t+1} = m + \delta_1\Delta d_t + \ldots + \delta_q\Delta d_{t-q+1} + w_{t+1}$$

$$\cdot\, m + \delta_1\Delta d_t + \ldots + \delta_q\Delta d_{t-q+1}$$

$$= \sum_1^\infty b^i E\Delta d_{t+i+1}|H_t \tag{12b}$$

$$w_{t+1} = z_t + \Delta c_{t+1}$$

$$z_t = \sum_1^\infty b^i(Ed_{t+i+1}|I_{t+1} - Ed_{t+i}|I_t) - \sum_1^\infty b^i E\Delta d_{t+i+1}|H_t.$$

The z_t variable is dated t rather than $t + 1$ to emphasize that it is orthogonal to H_t but not H_{t+1}. Under the null hypothesis that $c_t = 0$, the disturbances to (12a) and (12b) of course depend only a suitably dated z.

The trivariate system estimated for undifferenced specifications therefore was (1′), (11a), and (12a). For differenced specifications the system estimated was (1′), (11b), and (12b). The discount rate b was estimated from equation (1′) by two-step, two-stage least squares [Hansen, 1982]. The first step was standard two-stage least squares. The second step obtained the optimal, heteroskedasticity consistent estimate. The instruments used were the variables on the right-hand side of the dividend equation (11a) or (11b).

Equations (11a), (11b), (12a), and (12b) were estimated by OLS, with the covariance matrix of the parameters adjusted as described in (iii). Under the null, OLS may be used in (12a) and (12b), since $Ez_{t+1}|H_{t+1} = 0$ in (12a), $Ez_t|H_t = 0$ in (12b).

(iii) For both undifferenced and differenced specifications, the parameter vector estimated was thus $\hat\theta = (\hat b,\hat\mu,\hat\phi_1, \ldots, \hat\phi_q,\hat m,\hat\delta_1, \ldots, \hat\delta_q)$. $\hat\theta$ is asymptotically normal with a $(2q + 3) \times (2q + 3)$ asymptotic variance-covariance matrix V. V was calculated by the methods of Hansen [1982], Newey and West [1986], and West [1986a]. This allows for arbitrary heteroskedasticity conditional on the instruments. It also allows for an *arbitrary* ARMA process for the disturbance to equations (12a) and (12b). An appendix available from the author describes in detail the calculation of V.

(iv) The relationship between the parameters in (12a) and (12b), on the one hand, and b and the parameters of (11a) and (11b), on the other, may be derived in a straightforward fashion from the formulas in Hansen and Sargent [1981]. The corresponding constraints that are implied for stationary specifications are

$$
\begin{aligned}
&0 = m - b(1-b)^{-1}\Phi(b)^{-1}\mu \\
&0 = \delta_1 - [\Phi(b)^{-1} - 1] \\
\text{(13a)}\qquad &0 = \delta_j - \Phi(b)^{-1}\sum_{k=j}^{q} b^{k-j+1}\Phi_k \qquad j = 2, \ldots, q \\
&\Phi(b)^{-1} = \left[1 - \sum_{i=1}^{q} b^i\phi_i\right]^{-1}.
\end{aligned}
$$

The constraints for differenced specifications are

$$
\begin{aligned}
&0 = m - [b(1-b)^{-1}\Phi(b)^{-1} + \Phi(b)^{-1} - 1]\mu \\
&0 = \delta_j - \left\{\Phi(b)^{-1}\sum_{k=j+1}^{q} b^{k-j}\phi_k \right. \\
\text{(13b)}\qquad &\qquad \left. + [\Phi(b)^{-1} - 1]\phi_j\right\} \qquad j = 1, \ldots, q-1 \\
&0 = \delta_q - [\Phi(b)^{-1} - 1]\phi_q \\
&\Phi(b)^{-1} = \left[1 - \sum_{i=1}^{q} b^i\phi_i\right]^{-1}.
\end{aligned}
$$

Let $R(\theta)$ denote either of these $(q+1) \times 1$ constraints. The null hypothesis is that $R(\theta) = 0$. The test statistic was calculated as

$$
\text{(14)}\qquad R(\hat{\theta})'\left[\left(\frac{\partial R}{\partial \hat{\theta}}\right) V \left(\frac{\partial R}{\partial \hat{\theta}}\right)'\right]^{-1} R(\hat{\theta}).
$$

The derivative of $R(\hat{\theta})$ was calculated analytically. Under the null hypothesis, the statistic (14) is asymptotically distributed as a chi-squared random variable with $q + 1$ degrees of freedom.[3]

(v) The final item discussed before results are presented is diagnostic tests on the estimated equations.[4] As explained in the previous section of the paper, a significant value of the test statistic (14) is more compelling as evidence of bubbles the less the extent to which diagnostic tests on (1′), (11a), and (11b) indicate that other source of misspecification are present. Possible sources that have been suggested include failure to allow for expectational irrational-

3. One troublesome aspect of the distribution of the test statistic should be noted. This is that the test may not be consistent: if there are bubbles, the asymptotic probability that the test will reject the null may not be unity, even though the two sets of parameter estimates will be different with probability one in an infinite sized sample. See West [1985] for further discussion.

4. These same diagnostic tests were performed in West [1986c], and the discussion that follows is an abbreviated version of the discussion in section IVA of that paper.

ity [Ackley, 1983] and for time variation in discount rates [Leroy, 1984].

Four diagnostic checks were therefore performed on equations (1′), (11a), and (11b). The first checked for serial correlation in the residuals to the equations, using a pair of tests. Under rational expectations the expectational error u_{t+1} should be serially uncorrelated. If the ARIMA process for d_t is properly specified, so, too, should v_{t+1}, since v_{t+1} is the innovation in the process. The first of the pair of serial correlation tests checked for first-order serial correlation in u_{t+1} and v_{t+1}, using the techniques described in Pagan and Hall [1983, pp. 170, 191]. The second serial correlation test, performed only for v_{t+1}, calculated the Box-Pierce Q statistic for the residuals. This statistic tests for first- and higher order serial correlation [Granger and Newbold, 1977, p. 93].

The second of the four diagnostic checks, performed only on equation (1′), was Hansen's [1982] test of instrument-residual orthogonality. Under the null hypothesis that equation (1) is correctly specified, the test statistic is asymptotically distributed as a chi-squared random variable with q degrees of freedom. This test has the power to detect failures of equation (1) such as expectational irrationality and time variation in discount rates that is correlated with dividends.

The third of the four diagnostic checks tested for the stability of the regression coefficients in (1′), (11a), and (11b). This was done by testing for a midsample shift of the coefficients in these equations. The relevant statistic is asymptotically distributed as a chi-squared random variable, with one degree of freedom for (1′), $q + 1$ degrees of freedom for (11a) and (11b). This test clearly has the power to detect shifts in the discount rate, as well as in the dividend process.

The fourth and final diagnostic check performed is implicit in the estimation procedure described above. Several specifications of the dividend process were used—differenced and undifferenced—with a variety of lag lengths. Since the results did not prove very sensitive to the specification of the dividend process, it appears unlikely that small changes in the specification of the dividend process will affect the results.

B. Empirical Results

Regression results for (1′) are reported in Table IA.[5] The results in Table IA suggest that the basic arbitrage equation (1) is a

5. Tables IA and IB are identical to Tables IA and IB in West [1986c], so the discussion that follows is very similar to the discussion in section IVB of that paper.

TABLE IA
REGRESSION RESULTS: EQUATION (1′)

Data set	(1) Differenced	(2) q	(3) b	(4) ρ	(5) H/sig	(6) Stability/sig
S and P						
1873–1980	no	2[a]	0.9311 (0.0186)	0.0695 (0.0766)	5.50/0.064	4.55/0.033
1874–1980	yes	2[a]	0.9413 (0.0170)	0.0670 (0.0974)	2.87/0.238	0.33/0.566
1875–1980	no	4	0.9315 (0.0158)	0.0661 (0.0754)	6.96/0.138	3.69/0.055
1876–1980	yes	4	0.9449 (0.0136)	0.0671 (0.0984)	3.15/0.533	0.28/0.594
Modified Dow Jones						
1931–1978	no	3[a]	0.9402 (0.0301)	−0.1040 (0.0806)	5.42/0.144	1.56/0.211
1933–1978	yes	4[a]	0.9379 (0.0188)	−0.1182 (0.0752)	5.20/0.267	2.02/0.154
1932–1978	no	4	0.9271 (0.0253)	−0.1112 (0.1493)	6.08/0.108	0.49/0.483

See notes to Table IC.

sensible one. Consider first two diagnostic tests. Column (4) reports the estimates of the first-order serial correlation coefficient of the disturbance to (1′). Since the entries in the column are far from significant at the 0.05 level, there is little evidence of serial correlation in this disturbance. In addition, the entries in column (5), which report the Hansen [1982] test of instrument residual orthogonality, does not reject the null hypothesis of no correlation between the instruments and residuals. The successful results in column (5) are perhaps especially noteworthy, since failures of rational expectations models to pass this test are quite common [Hansen and Singleton, 1982; West, 1986b].

Most important, the discount rate b is estimated plausibly and precisely in all regressions. See column (3) in Table IA. The implied annual real expected returns are a reasonable 6 to 7 percent, and are quite close to the arithmetic means for ex post returns: 8.1 percent for the Standard and Poor's (S and P) index (1872–1981) and 7.4 percent for the Dow Jones index (1929–1979). Moreover, the entries in column (6) give little evidence that the rate was different in the two halves of either sample. The only specification for which the null hypothesis of equality can be rejected at the 5 percent level is Standard and Poor's, undifferenced, $q = 2$. In addition, no evidence

TABLE IB
REGRESSION RESULTS: EQUATIONS (12a) AND (12b)

Data set	(1) Differenced	(2) q	(3) μ	(4) ϕ_1	(5) ϕ_2	(6) ϕ_3	(7) ϕ_4	(8) ρ	(9) Q/sig	(10) Stability/ sig
S and P										
1873–1980	no	2[a]	0.168	1.0196	−0.238			0.045	36.87/0.181	12.93/0.005
			(0.084)	(0.114)	(0.103)			(0.025)		
1874–1980	yes	2[a]	0.034	0.262	−0.214			0.002	22.79/0.824	2.71/0.438
			(0.029)	(0.118)	(0.071)			(0.023)		
1875–1980	no	4	0.150	1.247	−0.480	0.227	−0.029	0.001	21.39/0.875	33.49/0.000
			(0.080)	(0.116)	(0.093)	(0.113)	(0.066)	(0.010)		
1876–1980	yes	4	0.036	0.264	−0.230	0.026	−0.006	0.001	23.98/0.773	4.34/0.501
			(0.031)	(0.115)	(0.094)	(0.080)	(0.153)	(0.011)		
Modified Dow Jones										
1931–1978	no	3[a]	1.945	1.265	−0.664	0.333		0.002	4.05/1.000	7.53/0.111
			(1.037)	(0.112)	(0.108)	(0.098)		(0.054)		
1933–1978	yes	4[a]	0.275	0.302	−0.351	0.051	0.050	−0.024	9.77/0.939	8.06/0.153
			(0.405)	(0.119)	(0.133)	(0.093)	(0.176)	(0.067)		
1932–1978	no	4	1.925	1.263	−0.662	0.330	0.004	0.005	4.06/1.000	10.22/0.069
			(1.900)	(0.111)	(0.208)	(0.209)	(0.134)	(0.022)		

See notes to Table IC.

against the constancy of the discount rate may be found in a comparison of the two halves' mean ex post returns. For the S and P index, these were (in percent) 8.09 (1872–1926) versus 8.12 (1927–1981); for the Dow Jones the figures are 7.87 (1929–1954) versus 6.92 (1955–1979).

The specification of the arbitrage equation (1), then, appears acceptable. Let us now consider the estimates for the dividend process, reported in Table IB. The entries in columns (8) and (9) indicate little evidence of serial correlation in the disturbance to equations (11a) and (11b). Both test statistics in all regressions are far from significant, except for the estimate of the first-order serial correlation coefficient $\hat{\rho}$ for the S and P index, undifferenced, lag length $q = 2$. This regression's Q statistic in column (9) does, however, comfortably accept the null hypothesis of no serial correlation. Overall, then, no serial correlation to the residuals to (11a) and (11b) is apparent. Also, the estimates of most regression coefficients are statistically significant, at least when the lag length q was chosen by the Hannan and Quinn [1979] procedure. Finally, the null hypothesis that the parameters of the dividend process are the same in the two halves of each sample can be rejected at the 5 percent level only for the S and P index, undifferenced. See column (10). In general, then, the specification of the dividend process seems acceptable, with the possible exception of the S and P data set, undifferenced.

Estimates of the third and final equation, (12a) or (12b), are in Table IC. Parameter estimates are fairly precise for undifferenced specifications, less so for differenced specifications.

In contrast to the coefficients of the other two equations, however, the estimates of the coefficients of equations (12a) and (12b) are probably not sensible from the point of view of the simple efficient markets model that says $p_t = \Sigma_1^\infty b^i E d_{t+i} | I_t$. For the estimates of these coefficients are uniformly incompatible with the estimates of the coefficients of the other two equations. The test of whether these estimates are in fact compatible—that is, the test of the null hypothesis that bubbles are absent—may be found in Table II. Equation (14) is calculated in column (4). Every specification but those for the S and P, differenced, rejects the null at any conventional significance level. One of the S and P differenced specifications rejects the null at the 5 percent level, the other at the 10 percent level.

It appears that the reason for the rejection is that the coefficients on dividends in the present value equations (12a) and (12b)

TABLE IC
REGRESSION RESULTS: EQUATIONS (13a) AND (13b)

Data set	(1) Differenced	(2) q	(3) $\hat{m}$	(4) $\hat{\delta}_1$	(5) $\hat{\delta}_2$	(6) $\hat{\delta}_3$	(7) $\hat{\delta}_4$
S and P							
1873–1980	no	2[a]	−24.28	31.152	−2.053		
			(11.42)	(6.331)	(4.118)		
1874–1980	yes	2[a]	0.79	−4.792	3.632		
			(1.51)	(2.842)	(2.205)		
1875–1980	no	4	−25.52	33.671	−14.257	11.029	−0.910
			(11.78)	(6.967)	(5.684)	(3.971)	(3.423)
1876–1980	yes	4	0.76	−5.099	4.436	−1.850	2.219
			(1.61)	(3.353)	(1.998)	(4.367)	(3.481)
Modified Dow Jones							
1931–1978	no	3[a]	−286.78	35.706	−22.869	19.866	
			(99.11)	(8.432)	(8.549)	(3.880)	
1933–1978	yes	4[a]	3.93	−10.438	9.213	−1.855	4.846
			(19.06)	(3.531)	(3.345)	(6.241)	(5.280)
1932–1978	no	4	−272.15	35.607	−21.639	20.015	−1.588
			(100.43)	(9.682)	(10.540)	(8.504)	(7.324)

Notes. A superscript a means lag length q chosen by Hannan and Quinn [1979] procedure. Asymptotic standard errors in parentheses. Symbols q, b, μ, ϕ_i, m, δ_i are defined in equations (1), (11a), (11b), (12a) and (12b). ρ = first order serial correlation coefficient of disturbance; H = Hansen's [1982] test of instrument-residual orthogonality, $H \sim \chi^2(q)$; "stability" is test for stability of coefficients, as described in text, distributed as $\chi^2(1)$ in Table IA and $\chi^2(q+1)$ in Table IB; Q is Box-Pierce Q statistic, $Q \sim \chi^2(30)$ for S and P, $Q \sim \chi^2(18)$ for Dow Jones. For the "H," "stability," and "Q" columns, "sig" refers to the probability of seeing the statistic under the null hypothesis.

TABLE II
TEST STATISTICS

Data set	(1) Differenced	(2) q	(3) Degrees of freedom	(4) Equation (14)/sig[b]
S and P				
1873–1980	no	2[a]	3	32.59/0.000
1874–1980	yes	2[a]	3	7.41/0.060
1875–1980	no	4	5	23.92/0.000
1876–1980	yes	4	5	11.40/0.044
Modified Dow Jones				
1931–1978	no	3[a]	4	45.12/0.000
1933–1978	yes	4[a]	5	19.93/0.001
1932–1978	no	4	5	40.22/0.000

a. Lag length q chosen by Hannan and Quinn [1979] procedure.
b. "Sig" refers to the probability of seeing the statistic under the null hypothesis.

are biased upwards. In six of the seven specifications, the sum of the biases in the $\hat{\delta}_i$ (not reported in any table) are positive. (The only exception is the S and P, differenced, $q = 2$.) Now, for undifferenced specifications, if there is a bubble, the bias in the estimate of the vector $(m,\delta_1, \ldots, \delta_q)$ is the probability limit of the vector of estimates of the parameters of a regression of the bubble c_{t+1} on a constant and $d_{t+1}, \ldots, d_{t-q+2}$. (See equation (10).) If bubbles reflect at least in part a tendency of the market to overreact to dividends or to news about future dividends [Shiller, 1984], this upward bias is precisely what would be expected. For differenced specifications the asymptotic bias in the estimate of the vector $(m,\delta_1, \ldots, \delta_q)$ is the probability limit of estimates of the parameters in a regression of the bubble on a constant and $\Delta d_{t+1}, \ldots, \Delta d_{t-q+1}$. If changes in bubbles tend to be associated with changes in lags of dividends, the $\hat{\delta}_i$ will also tend to be biased upward for differenced specifications.[6]

C. Dividends Follow a Lognormal Random Walk

The diagnostic tests discussed in the previous section found little fault with the specifications of the d_t process. Much research in finance, however, assumes that logarithmic and not arithmetic differences are necessary to induce stationarity in dividends [Kleidon, 1985]. As noted in Section II, it is possible to obtain a closed-form solution for $\Sigma_1^\infty b^i E d_{t+i} | H_t$ when $\Delta(\log d_t)$ is an iid

6. As noted in West [1985], the limiting distribution of the regression of c_t on dividends may not in general be a single vector. The statements in this paragraph therefore should be interpreted with caution.

normal random variable. This section applies the specification test, when d_t follows this lognormal random walk.

Suppose that $\Delta(\log d_t) \sim N(\mu,\sigma^2)$. Let $H_t = \{d_{t-i} | i \geqq 0\}$. Then $\Sigma_1^\infty b^i E d_{t+i} | H_t = \delta_1 d_t$, $\delta_1 = \exp(\mu + \sigma^2/2)/[b^{-1} - \exp(\mu + \sigma^2/2)]$ [Kleidon, 1985, p. 21]. Our aim is to compare an estimate of δ_1 obtained by regressing p_t on d_t with that obtained from estimates of μ, σ^2, and b. For each of the two data sets, μ and σ^2 were obtained as (a) the sample mean and variance of $\Delta(\log d_t)$, and (b) $\mu = 0$, $\sigma^2 = T^{-1}\Sigma(\Delta \log d_t)^2$. ($T$ = sample size.) Case (b), which imposes $\mu = 0$ and calculates the variance conditional on this, was tried because the point estimate of μ in each data set was insignificantly different from zero. b^{-1} was set equal to the mean ex post return. A convenient way to test the null hypothesis is to note that the formula for δ_1 implies that

$$\sigma^2 = 2\log\{(1/b)[\delta_1/(1+\delta_1)]\} - 2\mu. \tag{15}$$

Since $\Delta(\log d_t) \sim N(\mu,\sigma^2)$, $\hat{\sigma}^2 \sim \chi^2(T)$ when $\mu = 0$ is imposed, $\hat{\sigma}^2 \sim \chi^2(T-1)$ when μ is estimated. It is straightforward to construct a 99 percent confidence interval around $\hat{\sigma}^2$, as described in Mood et al. [1974, p. 382]. We can then check whether the point estimates of b^{-1}, δ_1, and μ are such that the right-hand side of (15) falls in this confidence interval. Note that such a procedure ignores sampling uncertainty in the estimates of b^{-1}, δ_1, and μ. One reason I am nonetheless applying this procedure is that the usual asymptotic theory does not apply to the regression that produces $\hat{\delta}_1$.

The empirical results are in Table III. The first line for each

TABLE III
EMPIRICAL RESULTS, LOGNORMAL RANDOM WALK

Data set	(2) $\hat{\sigma}^2$	(2) 99% CI	(3) RHS of (15)	(4) $(1/\hat{b})$	(5) $\hat{\delta}_1$	(6) $\hat{\mu}$	(7) $\hat{\rho}$
S and P 1872–1980	0.016	(0.012, 0.026)	0.045	1.081	23.19	0.013	0.176
						(0.012)	(0.095)
			0.071			0.0	
Dow Jones 1929–1978	0.024	(0.015, 0.043)	0.043	1.074	23.44	0.008	0.236
						(0.021)	(0.137)
			0.059				
						0.0	

Notes. Symbols σ^2, $1/b$, δ_1 defined above equation (15); π defined below equation (15).
The "99% CI" column gives the lower and upper bounds of a 99 percent confidence interval around the entry in column (1). These are calculated for a χ^2 (100) random variable for the S and P (sample size = 109), for a χ^2 (50) random variable for the Dow Jones (sample size = 50), as described in Mood et al. [1974, p. 382].
In column (3), "RHS of (15)" means "right-hand side of equation (15)."
The numbers in parentheses in column 6 are standard errors; in column 7 are asymptotic standard errors.

data set uses the mean of $\Delta(\log d_t)$ for μ, the second imposes $\mu = 0$. Only one point estimate of $\hat{\sigma}^2$ is reported for each data set, since $\hat{\sigma}^2$ was the same to three decimal places whether or not $\mu = 0$ was imposed. The lower and upper bounds for the 99 percent confidence interval are reported in column (2). The mean ex post return for each data set is in column four. The OLS estimate of δ_1 that results from regressing p_t on d_t is in column five. (It may help as a point of reference to state that the mean p_t/d_t ratio for the S and P is 21.05, for the Dow Jones is 22.24.) Column six has the sample mean of $\Delta(\log d_t)$, or zero. Note that for both data sets, the sample mean is insignificantly different from zero, at any conventional significance level. Column (7) has the point estimate of ρ, the first-order serial correlation coefficient of the residual. For both data sets, the estimate is insignificantly different from zero at the 10 percent level, but not at the 5 percent level. Column (3) has the right-hand side of equation (15), calculated from the figures in columns (4) to (6). The numbers in this column are all on or above the upper end of the 99 percent confidence interval for $\hat{\sigma}^2$, reported in column (2).

Apparently, the point estimates of the right-hand side of (15) are too big, or those of the left-hand side of (15) too small, for the data to have been generated by a constant discount rate, lognormal random walk model, without bubbles. This is consistent with the subsection III.B results: one interpretation is that $\hat{\delta}_1$, the coefficient that results when p_t is projected onto d_t, is too big for $p_t = p_t^*$ to be correct. Another interpretation, consistent not only with the earlier results in this paper but of those in a companion paper as well [West, 1986c], is that σ^2, the variance of the innovation in the univariate dividend process, is too small.

It does not, however, seem wise to push either of these arguments too far. One reason is that the simple lognormal random walk specification may not adequately capture the dynamics of the d_t process. The figures in column (7) of Table III suggest some residual serial correlation. A second reason is that the figures in Table III do not really indicate a rejection of the model at the 99 percent level, since sampling uncertainty in the estimates of b^{-1}, δ_1, and μ is ignored. One way to emphasize that this is a practical and not just pedantic point is to consider the effects of column (3) of different values of b^{-1}. Suppose that $b^{-1} = 1.05$, a value within two standard deviations of the point estimates in Table IA. Then all four column (3) estimates would not only fall below the upper end of the 99 percent confidence interval in column (2), but would all be below the point estimate of $\hat{\sigma}^2$ in column (1).

In sum, the, the lognormal specification provides mild evidence against the null that $p_t = p_t^*$, versus $p_t = p_t^* + c_t$.

D. Time Varying Discount Rates

Time variation in discount rates can be allowed under the null, if, as in Shiller [1981a], the model is still linear. Let r_{t+j} be the one-period return expected by the market at period $t + j - 1$. Let $p_t^* = E\{\Sigma_{i=1}^{\infty} [\Pi_{j=1}^{i} (1 + r_{t+j})^{-1}] d_{t+1}\} | I_t$. Under the null hypothesis of no bubbles, $p_t = p_t^*$. Let us linearize p_t^* around $\bar{r}$ and $\bar{d}$; selection of $\bar{r}$ and $\bar{d}$ is discussed below. Define $\bar{b} = (1 + \bar{r})^{-1}$, $\bar{a} = -\bar{d}/\bar{r}$. Then [Shiller, 1981a]

$$p_t^* \approx E\{\Sigma_{i=1}^{\infty} \bar{b}^i [\bar{a}(r_{t+i} - \bar{r}) + d_{t+i}]\} | I_t = \text{(say) } E\{\Sigma_{i=1}^{\infty} \bar{b}^i y_{t+i}\} | I_t.$$

The arbitrage equation corresponding to the null hypothesis that $p_t \approx E\{\Sigma_{i=1}^{\infty} \bar{b}^i [\bar{a}(r_{t+i} - \bar{r}) + d_{t+i}]\} | I_t$ is

$$(16) \quad p_t \approx \bar{b}E(y_{t+1} + p_{t+1}) | I_t = \bar{b}E[\bar{a}(r_{t+1} - \bar{r}) + d_{t+1} + p_{t+1}] | I_t.$$

As before, solutions to (16) are of the form $p_t = E\{\Sigma_{i=1}^{\infty} \bar{b}^i y_{t+i}\} | I_t + c_t$ for any c_t that satisfies $Ec_t | I_{t-1} = \bar{b}^{-1} c_t$. The null hypothesis we wish to test is that $c_t = 0$.

This can be done by comparing two sets of estimates of expected present discounted values, with expectations conditional on the set of current and past dividends. Now, however, the variable being forecast is not just d_{t+i} but y_{t+i}. This will not involve an arbitrage equation; it will involve dividend and distributed lag equations, as before, and also a new equation, for forecasting expected returns using current and lagged dividends. A brief discussion follows. Algebraic details are available on request.

The linearization parameters, $\bar{r}$, $\bar{b}$, and $\bar{a}$ were chosen as certain simple, plausible functions of the data. For both differenced and undifferenced specifications, the point of linearization for expected returns was the mean ex post return, $\bar{r} = T^{-1}\Sigma [(p_{t+1} + d_{t+1})/p_t] - 1$. Then $\bar{b} = (1 + \bar{r})^{-1}$. When dividends were assumed stationary, the point of linearization for $\bar{d}$ was mean dividends: $\bar{d} = T^{-1}\Sigma d_t$. When dividends were assumed to require (arithmetic) differences to induce stationarity, the point was $\bar{d} = (1 - \bar{b})\Sigma_{t=1}^{\infty} \bar{b}^{t-1} E_0 d_t$, $E_0 d_t = E_0 d_0 + tE\Delta d_t$, d_0 a presample value of dividends. Thus, $\bar{d} = d_0 + E\Delta d_t/(1 - \bar{b})$. Note that if dividends are stationary ($E\Delta d_t = 0$) and $d_0 = Ed_t$, this reduces to linearizing around mean dividends. For both differenced and undifferenced specifications, $\bar{a}$ was calculated as $\bar{a} = -\bar{d}/\bar{r}$. See Table IV for the resulting values of $\bar{r}$, $\bar{b}$, $\bar{d}$, and $\bar{a}$.

TABLE IV
LINEARIZATION PARAMETERS

Sample Period	(1) Differenced	(2) $\bar{r}$	(3) $\bar{b}$	(4) $\bar{d}$	(5) $\bar{a}$
1901–1981	no	0.0792	0.9266	4.0054	−50.56
1902–1981	yes	0.0772	0.9283	3.1441	−40.72

The dividend equation is precisely that used in the constant discount rate case, in subsection III.B.

For undifferenced specifications the distributed lag equation was obtained by projecting $E\Sigma\, \bar{b}^i y_{t+i+1} | I_{t+1}$ onto the space of current and lagged dividends H_{t+1}, as in equation (12a). For differenced specifications a difference of $E\Sigma\, \bar{b}^i y_{t+i+1} | I_{t+1}$ was projected onto H_t, as in equation (12b).

The final relationship involved is a regression to forecast expected returns. Let $R_{t+j} = (p_{t+j} + d_{t+j})/\, p_{t+j-1}$ denote the ex post return. Note that since H_t is a subset of I_t, $R_{t+j} = r_{t+j} + \nu_{t+j}$, with ν_{t+j} orthogonal to H_t. So $ER_{t+j}|H_t = Er_{t+j}|H_t$: a regression to forecast ex post returns also forecasts expected returns. The regressions are

$$R_{t+1} = g + \gamma_0 d_t + \ldots \gamma_n d_{t-n} + \eta_t \tag{17a}$$

$$R_{t+1} = g + \gamma_0 \Delta d_t + \ldots \gamma_n \Delta d_{t-n} + \eta_t \tag{17b}$$

η_t serially correlated in general, $Ex_s\eta_t = 0$ for x_s an element of H_t.

One can use (17a) to solve for $E\bar{b}^i[\bar{a}(R_{t+i} - \bar{r})]|H_t$. As before, the dividend equation (11a) yields $E\bar{b}^i d_{t+i}|H_t$. Together these produce the $E\Sigma\, \bar{b}^i y_{t+i+1}|H_{t+1}$. Similarly, (17b) and (11b) yield the distributed lag equation in differenced specifications.

For computational simplicity, the specification test was performed conditional on $\bar{a}$, $\bar{r}$, $\bar{b}$, and the parameters of equations (17a) and (17b). It may be shown that the parameters of the distributed lag equation can be estimated from the regressions

$$\tilde{p}_{t+1} = m + \delta_1 d_{t+1} + \ldots + \delta_q d_{t-q+2} + \tilde{w}_{t+1} \tag{18a}$$

$$\Delta\tilde{p}_{t+1} = m + \delta_1 \Delta d_t + \ldots + \delta_q \Delta d_{t-q+1} + \tilde{w}_{t+1}. \tag{18b}$$

The left-hand side variables $\tilde{p}_{t+1}$ and $\Delta\tilde{p}_{t+1}$ are calculated from p_{t+1} and Δp_{t+1}, and lags of d_t and Δd_t, using $\bar{a}$, $\bar{b}$, and the estimates of the parameters of (17a) and (17b). The δ_1 are functions of $\bar{b}$ and the parameters of the dividend process, as written out in equations (13a) and (13b). If the γ_i in equations (17a) and (17b) are identically zero, then $\tilde{p}_{t+1} = p_{t+1}$, $\Delta\tilde{p}_{t+1} = \Delta p_{t+1}$: if the return that is expected conditional on past dividends is constant, the test in this section reduces to that in subsection III.B.

The length of the distributed lag of ex post returns on dividends was set to 30, as in Shiller [1984, Table I]. This was done because OLS standard errors suggested insignificant γ_i for both the Dow Jones and the S and P for a lag length of ten years. Because of degrees of freedom limitations resulting from the thirty-year lag, the test in this section was applied only to the S and P. An unconstrained lag was used, since both γ_0 and the sum of the γ_i were estimated more precisely with this lag than with Shiller's [1984] polynomial distributed lag.

The regression of returns on dividends is reported in Table VA. In contrast to the results of the previous section, some predictability of returns is suggested. γ_0, the coefficient on d_t or Δd_t, was significantly different from zero at the 95 percent level. So, too, was the sum of the other distributed lag coefficients. See columns (3) and (4). The significance of the coefficients is, however, probably somewhat overstated, since, as explained above, some experimentation was done to obtain a specification with significant coefficients.

Table VB has estimates of the dividend equations. These look quite similar to those in Table IB. Table VC has estimates of the distributed lag equations (18a) and (18b). Note that the coefficients are woefully insignificant for differenced specifications.

Results of the test of the null hypothesis of no bubbles are reported in Table VI. The null is strongly rejected for undifferenced specifications, not at all for differenced specifications. The sum of the biases of the $\hat{\delta}_i$ was positive for all specifications except differenced, lag length = 4.

It is rather disturbing that allowing for time varying discount rates yields stronger evidence against the model for undifferenced specifications, weaker evidence for differenced specifications. One possible explanation is that the equation (17) forecasts of future

TABLE VA
REGRESSION RESULTS: EQUATIONS (17a) AND (17b)

Sample period	(2) Differenced	(2) $\hat{g}$	(3) $\hat{\gamma}_0$	(4) $\Sigma_{i=1}^{29} \hat{\gamma}_i$	(5) $d.w.$
1901–1980	no	0.0623	−0.1560	0.1893	1.93
		(0.0891)	(0.0585)	(0.0693)	
1902–1980	yes	0.1456	−0.1470	−1.2790	1.92
		(0.0349)	(0.0592)	(0.5938)	

TABLE VB
REGRESSION RESULTS: EQUATIONS (12a) AND (12b)

Sample period	(1) Differenced	(2) q	(3) $\hat{\mu}$	(4) $\hat{\phi}_1$	(5) $\hat{\phi}_2$	(6) $\hat{\phi}_3$	(7) $\hat{\phi}_4$
1901–1980	no	2	0.295	1.185	−0.253		
			(0.132)	(0.123)	(0.108)		
1902–1980	yes	2	0.033	0.265	−0.221		
			(0.039)	(0.129)	(0.073)		
1901–1980	no	4	0.266	1.241	−0.500	0.244	−0.046
			(0.124)	(0.124)	(0.099)	(0.116)	(0.070)
1902–1980	yes	4	0.034	0.270	−0.235	0.027	−0.041
			(0.040)	(0.153)	(0.099)	(0.086)	(0.168)

TABLE VC
REGRESSION RESULTS: EQUATIONS (18a) AND (18b)

Sample period	(1) Differenced	(2) q	(3) $\hat{m}$	(4) $\hat{\delta}_1$	(5) $\hat{\delta}_2$	(6) $\hat{\delta}_3$	(7) $\hat{\delta}_4$
1901–1980	no	2	6.39	11.969	4.427		
			(5.63)	(1.418)	(1.778)		
1902–1980	yes	2	0.59	−1.168	0.659		
			(0.58)	(0.844)	(0.713)		
1901–1980	no	4	4.03	13.110	−3.007	1.980	5.047
			(5.27)	(2.086)	(2.356)	(1.423)	(1.932)
1902–1980	yes	4	0.58	−1.129	−0.645	0.147	0.171
			(0.58)	(1.030)	(0.575)	(1.534)	(0.943)

Notes: Asymptotic standard errors in parentheses.
Symbols: q = lag length of dividend autoregression (12a) and (12b); $\bar{r}$, $\bar{b}$, $\bar{d}$, and $\bar{a}$ are defined above equation (16); μ, ϕ_i, g, γ_i, m, and δ_i are defined in equations (12a), (12b), (17a), (17b), (18a), and (18b); *d.w.* is the Durbin-Watson statistic.

TABLE VI
TEST STATISTICS

Sample period	Differenced	q	Degrees of freedom	Equation (14) sig
1901–1980	no	2	3	33.72/0.000
1902–1980	yes	2	3	2.00/0.572
1901–1980	no	4	5	30.86/0.000
1902–1980	yes	4	5	2.67/0.750

See Notes to Table II.

expected returns are quite noisy, and very different from the market's actual expected returns. The fitted values from equation (17a) could then be spuriously leading to a rejection for undifferenced specification, or conversely, those from (17b) could be incorrectly suggesting little evidence against the model for differenced specifications. A priority for future research, then, is performing similar tests with a more tightly constrained parameterization of what determines expected returns. The evidence in this section does little to pin down the extent to which the rejections in the constant discount rate model are due to variations in discount rates.

IV. Discussion and Conclusions

This section contains some concluding comments on the previous section's results. The first comment to make is that any diagnostic tests of (1′) will clearly have arbitrarily small power against "near rational" bubbles that are arbitrarily close to being rational. This is, if $p_t = p_t^* + c_t$, and $Ec_t|I_{t-1} = k^{-1}c_{t-1}$ for some k that is very close to b, diagnostic tests on equation (1′) may well fail to reject equation (1′). Summers [1986] calculated the small sample power of tests similar (though not identical) to those performed in subsection III.B, and, unsurprisingly, found that such tests are unlikely to detect variations in expected returns caused by near rational bubbles.

The presence of near rational bubbles certainly means that equation (1) is, strictly speaking, invalid. This fact does not, however, seem to me to be of great importance for the interpretation or implications of the results of Section III. A near rational bubble that tends to generate nearly constant expected returns will tend to generate nearly the same time series pattern of prices as will a rational bubble. That the tests in subsection III.B have little power to distinguish between such a bubble and a strictly rational one is not, then, very important for the interpretation of the evidence presented in Section III, at least at the level of generality of this paper.

The second comment to make concerns what determines whether a constant expected return specification is a good approximation for the purposes of the specification test. The subsection III.D analysis of a linearized model with time variation in expected returns suggests that the key requirement is not near constancy of returns expected by the market, but near constancy of the return that is expected conditional on the past history of dividends. In the linearized model, $Ep_t^*|H_t = \bar{a}\Sigma_{i=1}^{\infty} \bar{b}^i E(r_{t+i} - \bar{r})|H_t + \Sigma_{i=1}^{\infty} \bar{b}^i Ed_{t+1}|H_t$. In that model, then, $Ep_t^*|H_t = \Sigma_{i=1}^{\infty} \bar{b}^i Ed_{t+i}|H_t$ not only when $r_{t+i} = \bar{r}$ but also when $Er_{t+i}|H_t = \bar{r}$ for $i > 0$, i.e., when past dividends do not help predict future expected returns. Intuitively, if only past dividends are used to forecast the expected present discounted value of future dividends, and variations in expected returns are independent of past dividends, it is reasonable to forecast expected returns to be at their unconditional mean and to discount future dividends at a constant rate. This statement holds in a strict mathematical sense in a linearized model, and therefore may hold approximately in the underlying nonlinear model.

An implication is that most of the mild evidence against the constant expected return specification in Shiller [1981a, 1984] and all of the somewhat stronger evidence in Flood, Hodrick, and Kaplan [1986] may well not be directly relevant to the interpretation of the results of this paper.[7] It is, of course, of interest to investigate further whether, for the purposes of the specification test, it is adequate as an approximation to consider only variations in expected returns that are predictable from the past history of dividends, and the extent to which such variations explain what seems to be anomalous behavior in the data. This is an important task for future research.

What is required is a reconciliation of what appear to be incompatible price and dividend data. The incompatibility is manifested in an upward bias in the estimates of the coefficients of the projection of prices onto lagged dividends. A reconciliation that involves a parametric model for bubbles, or fads, and which allows for variation in expected returns, is a challenging task for future research.

PRINCETON UNIVERSITY

REFERENCES

Ackley, Gardner, "Commodities and Capital: Prices and Quantities," *American Economic Review*, LXXIII (1983), 1–16.
Blanchard, Olivier, and Mark Watson, "Bubbles, Rational Expectations and Financial Markets," NBER Working Paper No. 945, 1982.

7. Using the same data as in this paper, Shiller [1981a, 1984] and Flood, Hodrick, and Kaplan [1986] obtain coefficients significant at the 5 percent level when the ex post return is regressed on certain sets of lagged prices, dividend-price ratios, and ex post returns. Shiller [1984] also obtains significant coefficients using lagged earnings.

Box, G. E. P., and G. M. Jenkins, *Time Series Analysis: Forecasting and Control* (San Francisco, CA: Holden Day, 1976).

Brealey, Richard, and Stewart Myers, *Principles of Corporate Finance* (New York, NY: McGraw Hill, 1981).

Diba, Behzad T., and Herschel I. Grossman, "Rational Bubbles in the Price of Gold," NBER Working Paper No. 1300, 1984.

Flood, Robert P., and Peter M. Garber, "Market Fundamentals versus Price Level Bubbles: The First Tests," *Journal of Political Economy,* LXXXVIII (1980), 745–70.

——, ——, and Louis O. Scott, "Multi-Country Tests for Price Level Bubbles," *Journal of Economic Dynamics and Control,* VIII (1984), 329–40.

Flood, Robert P., Robert Hodrick, and Paul Kaplan, "An Evaluation of Recent Evidence on Stock Market Bubbles," manuscript, 1986.

Granger, C. W. J., and Paul Newbold, *Forecasting Economic Time Series* (New York, NY: Academic Press, 1977).

Hamilton, James D., and Charles M. Whiteman, "The Observable Implications of Self-Fulfilling Prophecies," manuscript, 1984.

Hannan, E. J., and B. G. Quinn, "The Determination of the Order of an Autoregression, *Journal of the Royal Statistical Society Series B,* XLI (1979), 190–95.

Hansen, Lars Peter, "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica,* L (1982), 1029–54.

Hansen, Lars Peter, and Thomas J. Sargent, "Formulating and Estimating Dynamic Linear Rational Expectations Models," in Robert E. Lucas, Jr., and Thomas J. Sargent, eds., *Rational Expectations and Econometric Practice* (Minneapolis, MN: University of Minnesota Press, 1981, pp. 91–126.

Hansen, Lars Peter, and Kenneth Singleton, "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica,* L (1982), 1269–86.

Hausman, Jerry A., "Specification Tests in Econometrics," *Econometrica,* XLVI (1978), 1251–71.

Keynes, John Maynard, *The General Theory of Employment, Interest and Money* (New York NY: Harcourt, Brace and World, 1964).

Kleidon, Allan W., "Variance Bounds Tests and Stock Price Valuation Models," Stanford University Graduate School of Business Research Paper No. 806, 1985.

Leroy, Stephen, "Efficiency and Variability of Asset Prices," *American Economic Review,* LXXIV (1984), 183–87.

——, and Richard Porter, "The Present Value Relation: Tests Based on Implied Variance Bounds," *Econometrica,* LXIX (1981), 555–574.

Marsh, Terry A., and Robert C. Merton, "Dividend Variability and Variance Bounds Tests for the Rationality of Stock Market Prices," Sloan School of Management Working Paper No. 1584–84, 1984.

Mood, Alexander M., Franklin A. Graybill, and Duane C. Boes, *Introduction to the Theory of Statistics* (New York, NY: McGraw Hill, 1974).

Newey, Whitney K., and Kenneth D. West, "A Simple, Positive, Semidefinite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," forthcoming, *Econometrica,* 1986.

Pagan, A. R., and A. D. Hall, "Diagnostic Tests as Residual Analysis," *Econometric Reviews,* II (1983), 159–218.

Sargent, Thomas J., and Neil Wallace, "Identification and Estimation of a Model of Hyperinflation with a Continuum of 'Sunspot' Equilibria," manuscript, 1984.

Shiller, Robert J., "Rational Expectations and the Dynamic Structure of Macro Models," *Journal of Monetary Economics,* IV (1978), 1–44.

——, "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?" *American Economic Review,* LXXI (1981a), 421–36.

——, "The Use of Volatility Measures in Assessing Market Efficiency," *Journal of Finance,* XXXV (1981b), 291–304.

——, "Stock Prices and Social Dynamics," *Brookings Papers on Economic Activity* (1984), 457–98.

Summers, Lawrence H., "Does the Stock Market Rationally Reflect Fundamental Values," *Journal of Finance,* XXXX (1986), 591–601.

Taylor, John B., "On the Conditions for Uniqueness in the Solution of Rational Expectations Models," *Econometrica,* XLV (1977), 1377–85.

Tirole, Jean, "Asset Bubbles and Overlapping Generations," *Econometrica*, LIII (1985), 1071–1100.
West, Kenneth D., "A Specification Test for Speculative Bubbles," Princeton University Financial Research Memorandum No. 58, 1985.
——, "Asymptotic Normality, When Regressors Have a Unit Root," Princeton University Woodrow Wilson School Discussion Paper No. 110, 1986a.
——, "A Variance Bounds Test of the Linear Quadratic Inventory Model," *Journal of Political Economy*, XCIII (1986b), 374–401.
——, "Dividend Innovations and Stock Price Variability," NBER Working Paper No. 1833, 1986c.