NBER WORKING PAPER SERIES

STOCK MARKET FORECASTABILITY AND VOLATILITY: A STATISTICAL APPRAISAL

N. Gregory Mankiw

David Romer

Matthew D. Shapiro

Working Paper No. 3154

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 1989

A preliminary version of this paper was presented at IMSSS. We are grateful to Robert Shiller, Ian Tonks, and participants in the LSE conference for comments, and to William Nelson for research assistance. This paper is part of NBER's research program in Financial Markets and Monetary Economics. Any opinions expressed are those of the authors not those of the National Bureau of Economic Research.

NBER Working Paper #3154
October 1989

STOCK MARKET FORECASTABILITY AND VOLATILITY: A STATISTICAL APPRAISAL

ABSTRACT

This paper presents and implements statistical tests of stock market forecastability and volatility that are immune from the severe statistical problems of earlier tests. Although the null hypothesis of strict market efficiency is rejected, the evidence against the hypothesis is not overwhelming. That is, the data do not provide evidence of gross violations of the conventional valuation model.

N. Gregory Mankiw
Department of Economics
Harvard University
Cambridge, MA 02138
(617)495-4301

David Romer
Department of Economics
University of California
Berkeley, CA 94720
(415)642-1785

Matthew D. Shapiro
Department of Economics
University of Michigan
Ann Arbor, MI 48109
(313)764-5419

1. INTRODUCTION

To the casual observer, fluctuations in stock prices appear dramatically too large to be explained by changes in underlying economic factors. The value of the stock market often changes by a few percent over brief periods with little apparent relevant news, for example. In his pathbreaking paper, Shiller (1981) appeared to provide striking formal evidence in favor of this view. Stock prices, Shiller argued, are five times too volatile to be accounted for by changes in fundamentals. Subsequent work has established, however, that the test employed by Shiller suffers from severe statistical difficulties (Flavin, 1983; Marsh and Merton, 1986; and Kleidon, 1986).

The purpose of this paper is to present a test of the hypothesis of stock market efficiency in the spirit of Shiller's work that does not suffer from these difficulties. The test builds on our earlier paper (Mankiw, Romer, and Shapiro, 1985). In that paper, we found that a "naive forecast" outperforms the market price as a predictor of the perfect foresight price (the present discounted value of dividends plus the discounted terminal price). This apparent departure from the hypothesis that stock prices are the expected present discounted value of dividends, although striking on its face, should be interpreted with caution. We noted that the departure might not be statistically significant, or that it might be due to variation in the required rate of return.

In this paper, we develop procedures for conducting statistical inference using our test statistics that are appropriate in finite sample and we extend the tests to the case of a variable required rate of return. Variation in the required rate of return does not account for our finding of violations of the predictions of the efficient markets hypothesis. In fact, the required rate of return appears to vary somewhat in the opposite direction of what would be needed to reconcile our earlier findings with efficient markets. At the same time, we find that the rejections of the model (under either a constant or variable required rate of return) are only moderately statistically significant.

The reason that the tests fail to provide overwhelming evidence against efficient markets is simple. In informal discussions of volatility tests, it is often claimed that dividends are extremely smooth and thus that the efficient markets hypothesis implies that stock prices--which should equal the expected present value of future dividends--should also be very smooth. The large volatility of stock prices, the argument concludes, thus blatantly contradicts efficient markets.[1] But the assertion that dividends are dramatically less variable than stock prices is false: for the period 1872 through 1987, the standard deviations of the annual percentage changes in real dividends and in real stock prices are 12.4 and 17.6 percent respectively. In the simple case where dividends are a logarithmic random

---

[1]Keynes's discussion of the stock market is usually taken in this light. More recently, see Shiller (1981) and Summers (1986). Ackley, in his AEA presidential address states that "Shiller's [results] appear to demolish the possibility that movements of U.S. stock prices can be explained by the rational expectations of share holders." (Ackley, 1983, p. 13).

walk with drift (which is a good first approximation in the data) and the required rate of return is a constant, stock prices will be proportional to dividends. Consequently, the variances of the percentage changes in stock prices and in dividends will be equal. In this sense, at first glance stock prices appear to be only moderately too variable. The results of our formal tests confirm this initial impression: we find that although the data do not appear fully consistent with market efficiency, they do not grossly contradict it.

Other recent tests of stock market efficiency focus on forecastability rather than excess volatility (see, for example, Fama and French, 1988a, b and Poterba and Summers, 1988). Below we discuss how forecastability and volatility tests are closely related. They are derived from similar orthogonality conditions; they suffer from similar statistical difficulties. Conventional statistical tests are biased toward finding excessive forecastability in finite samples.

A frequent criticism of Shiller's volatility test is that a conjunction of smooth dividends and volatile prices would not be evidence against market efficiency because dividends may be slow to reflect new information about profitability. In this case, within-sample information about future profitability would be reflected in the end-of-sample price rather than in within-sample dividends (Kleidon, 1986; Marsh and Merton, 1986). Merton (1987) extends this class of criticisms to our previous work: he argues that if the end-of-sample price is sufficiently important, our volatility tests will be invalid statistically. Specifically, he shows that in the case in which stocks pay no dividends--so that only the end-of-sample price is relevant to assessing the rationality of pricing within sample--the standard errors of our estimates explode as the sample size becomes large. He also suggests that this problem will arise even if dividends are paid in the sample.

Whether the point raised by Merton invalidates our tests in practice depends on whether the end-of-sample price does dominate our estimates. We establish analytically that Merton's argument is not correct if stocks pay dividends in realistic quantities. The intuition behind this condition is simple. The crux of the criticism of volatility tests by Kleidon (1986), Marsh and Merton (1986), and Merton (1987) is that out-of-sample events might be the dominate determinant of fluctuations in stock prices. If so, in-sample dividends will not provide an adequate lever for tests of market rationality. Yet, if the <u>level</u> of dividends is high enough in the sample, the importance of out-of-sample events is circumscribed. Specifically, if dividends are sufficiently high so that much of the value of the stock is paid out in the sample, the possible contribution of out-of-sample events to in-sample stock prices is limited by the present value relation.

As an empirical matter, there are sufficient dividends paid during the sample for the tests of efficiency and volatility to be statistically valid. These results should bear on a wide range of tests that seek to test the conventional valuation model based on comparing the sample path of prices and dividends.[2]

---

[2] Shiller (1988), in his reply to Kleidon, shows that Kleidon's striking reversals of the inequalities arise from a specification with an unrealistically small quantity of dividends.

Our tests will also be statistically valid if we consider a variant that assumes fixed holding periods. With fixed holding periods, many end-of-holding period prices are observed in a sample. Therefore, a single observation on the end-of-sample price need not do the impossible job of statistically accounting for the in-sample revisions of expectations about out-of-sample events.

The remainder of the paper consists of four sections and an Appendix. Section 2 reviews our earlier tests and shows how statistical inference can be conducted for the tests. In addition, this section presents a variation on our test that focuses on holding returns over fixed horizons (rather than to the end of the sample period, as in previous volatility tests). Section 3 presents results for the constant required rate of return case. Section 4 extends the tests and results to the variable required return case. Section 5 compares our procedures to regression tests of the present value model and of the forecastability of stock returns. Section 6 summarizes the paper. Finally, the Appendix presents our discussion of Merton's argument.

2. TESTING STRATEGY

Models with rational expectations often imply that some observable variable is a rational forecast of another variable. Most obviously, the efficient markets hypothesis implies that stock prices are the expected present value of future dividends. Other theories that have implications of this form include the expectations theory of the term structure, uncovered interest rate parity, the permanent income hypothesis, and the hypothesis that preliminary announcements of economic variables are rational forecasts.

In this section we outline a general procedure for testing predictions of this type. Our strategy is to begin by constructing statistics that are valid under the null hypothesis under quite general conditions; that is, we consider sample statistics whose expected values under the null hypothesis can be determined employing minimal auxiliary assumptions. We then construct standard errors that are valid asymptotically under quite weak assumptions. Thus the only potential difficulties in employing the tests involve constructing appropriate finite sample standard errors--a problem in virtually all statistical inference. But because the tests begin with sample statistics whose expectations under the null are known, they attenuate the severe biases in tests of rational expectations models that can be caused by small samples (Flavin, 1983; Mankiw and Shapiro, 1986) and by non-stationarity (Marsh and Merton, 1986; Kleidon, 1986). The particular Monte Carlo method that we use for the finite sample inference is described below. Analytic results concerning the asymptotic validity of the tests are presented in the Appendix. We present the test in the context of the efficient markets hypothesis; our discussion can easily be modified to apply to other rational expectations theories.

Define P_t^{*h} as the perfect foresight price for the strategy of buying a stock at time t and holding it for h periods. Assume that the required rate of return, denoted r, is constant. We relax this assumption in Section 4. Thus,

$$(1) \qquad P_t^{*h} = \sum_{j=0}^{h-1} \left(\frac{1}{1+r}\right)^{j+1/2} D_{t+j} + \left(\frac{1}{1+r}\right)^{h} P_{t+h} ,$$

where D_{t+j} are dividends in period $t+j$ and P_{t+h} is the market price in period $t+h$. (P_t and P_t^{*h} are beginning of period prices. We assume that dividends are paid in the middle of the period.)

Under the null hypothesis of market efficiency, there cannot be expected profit opportunities from the strategy of buying the stock at time t, holding the stock and collecting dividends until time $t+h$, and then selling the stock at the market price. It follows that

$$(2) \qquad P_t = E_t P_t^{*h} .$$

Note that (2) holds for all values of h.[3]

In our previous paper, we assume that h, for each observation, equals the length of time to the end of the sample. With this assumption, P_t^{*h} is the perfect foresight price for the strategy of holding the stock until the end of the sample period and then selling the stock at the prevailing price. This procedure of setting $h_t = T-t$ corresponds to the conventional practice in volatility tests. A natural alternative is to let h be constant across observations--that is, to let P_t^{*h} be the perfect foresight price for the policy of holding the stock until period $t+h$ and

[3] As we discuss in our previous paper, because the definition of P_t^{*h} includes the selling price, equation (2) holds even in the presence of speculative bubbles. See Flood and Hodrick (1986) for an elaboration of this point.

then selling it at the market price. In what follows, we consider both the variable and fixed horizon cases.

Equation (2) implies that $P_t^{*h} - P_t$ is uncorrelated with any information available at time t. In particular, let P_t^0 denote some "naive forecast" of P_t^{*h} that is available at time t. If (2) holds, then

$$(3) \qquad E_t(P_t^{*h} - P_t)(P_t - P_t^0) = 0 .$$

This in turn implies

$$(4) \qquad E_t(P_t^{*h} - P_t^0)^2 = E_t(P_t^{*h} - P_t)^2 + E_t(P_t - P_t^0)^2 .$$

The same relation holds if we normalize the data by information available at time t. That is, letting W_t be any variable known at time t,

$$(5) \qquad E_t\left(\frac{P_t^{*h} - P_t^0}{W_t}\right)^2 = E_t\left(\frac{P_t^{*h} - P_t}{W_t}\right)^2 + E_t\left(\frac{P_t - P_t^0}{W_t}\right)^2 .$$

Now define

$$(6) \qquad q_t = \left(\frac{P_t^{*h} - P_t^0}{W_t}\right)^2 - \left[\left(\frac{P_t^{*h} - P_t}{W_t}\right)^2 + \left(\frac{P_t - P_t^0}{W_t}\right)^2\right] .$$

Equation (5) implies $E_t q_t = 0$, and therefore (by the law of iterated projections) $Eq_t = 0$, where E denotes the expectation conditional on information available at the beginning of the sample period. We can write this as

(7) $$q_t = \alpha + \varepsilon_t ,$$

$$H_0 : \alpha = 0 ,$$

where ε_t is a zero-mean disturbance. To test the null hypothesis, we simply need to test whether α , the mean of the q_t's , is zero. Because the expectation of each q_t is zero, the expectation of the mean is zero. Note that the test that the mean of the q_t's is zero is similar to a regression test of the present value model (see Scott, 1985). Under the conventional valuation model given in equation (2), a regression of $P_t^{*h} - P_t$ on $P_t - P_t^0$ should yield a zero coefficient. Thus, both this regression and the our statistic in equation (7) generate tests of the orthogonality condition (3). In Section 5, we investigate the relationship between these tests in more detail.

Testing whether α is zero requires constructing a standard error for α. We would like to construct a standard error that is valid under as general a set of conditions as possible. For the case of a fixed holding period (that is, h_t constant across observations), q_t and q_{t-j} are correlated for $j < h$ but uncorrelated for $j \geq h$: q_t is uncorrelated with any information known at t, and q_{t-j} becomes known at $(t-j)+h$. Moreover, we have no reason to rule out heteroskedasticity in the q_t's. Thus, we wish to compute the standard error for the sample mean of a process of unknown heteroskedasticity and h-1 order serial correlation of unknown form. In the case of a variable holding period (that is, $h_t = T-t$), the serial correlation is not truncated after a fixed number of lags. Thus, for this case we require the standard error for a process with unknown serial correlation and heteroskedasticity.

To compute the standard errors, we use the technique of Hansen (1982) and White and Domowitz (1984). For the Hansen-White-Domowitz standard errors to be valid, only weak regularity conditions need to be satisfied. These conditions are satisfied, loosely speaking, if the correlation of the elements in the sample die out rapidly enough.

In the case of a sample mean, it is very easy to compute the standard errors. If R_k is the k^{th} sample autocovariance of the data used to compute the mean, define $\bar{R}$ as

$$(8) \qquad \bar{R} = R_0 + 2 \sum_{k=1}^{J} \omega_k R_k$$

where ω_k are weights.[4] The standard error is given by the square root of $\bar{R}$ divided by the sample size.[5] To implement the Hansen-White-Domowitz standard errors, we must choose the number of autocorrelations J to be included in the above expression. When the serial correlation is known to be zero after a certain lag--as in the fixed holding return case--J is chosen to equal the lag (Hansen, 1982). When the serial correlation is of

[4] The estimates are consistent with weights equal to one, but $\bar{R}$ is not guaranteed to be positive. We use the weighting scheme of Newey and West (1987), which also yields consistent estimates and is guaranteed to be positive. Specifically, $\omega_k = (J+1-k)/(J+1)$.

[5] Note in the case with no lags, the robust estimate of the standard error for the mean is just the ordinary one.

unknown and possibly infinite order--as in the variable $h_t = T-t$ case--White and Domowitz establish that the above computed standard error remains consistent if J increases at a rate at least as slow as the square root of the sample size but goes to infinity as the sample size increases. We choose J equal to 10 in this case.[6]

The Hansen-White-Domowitz standard errors are only valid asymptotically. Below we report results of a Monte Carlo experiment in which we attempt to judge the validity of the standard errors in small samples. The economic assumptions underlying the Monte Carlo experiment are discussed in the next section of the paper.

In his analysis of our previous paper, Merton (1987, pp. 110-116) argues that our tests have extremely undesirable asymptotic properties. In particular, he shows that for a case in which no dividends are paid, properly constructed standard errors for the test based on the variable holding period diverge as the sample period becomes large. Intuitively, for this case P_t^{*h} for each observation depends solely on the terminal value of the stock price, P_T, and thus observations remain correlated as the time between them becomes large. As a result, the regularity conditions needed for valid inference fail. In the Appendix, we show that Merton's argument does not hold either if dividends are paid in sufficient quantity or if we consider a fixed holding period. In both cases, the role of the terminal price in P_t^{*h} for a given t becomes small as the sample size becomes large; the tests therefore have desirable asymptotic properties.

---

[6]We have annual data from 1871 to 1988, so 10 is roughly the square root of the sample size.

Finally, as we emphasize in our previous paper, the equality (4) implies the following two inequalities:

$$(9) \qquad E\,\frac{1}{T}\sum_t \left(\frac{P_t^{*h} - P_t^0}{W_t}\right)^2 \geq E\,\frac{1}{T}\sum_t \left(\frac{P_t^{*h} - P_t}{W_t}\right)^2$$

and

$$(10) \qquad E\,\frac{1}{T}\sum_t \left(\frac{P_t^{*h} - P_t^0}{W_t}\right)^2 \geq E\,\frac{1}{T}\sum_t \left(\frac{P_t - P_t^0}{W_t}\right)^2 .$$

In the context of formal statistical tests of the efficient markets hypothesis (2), the role of these inequalities is to provide a means of interpreting the results of the tests. Inequality (9) states that the weighted mean square error using the naive forecast P_t^0 to predict P_t^{*h} should exceed the error using the optimal forecast P_t. Inequality (10) states that the volatility of P_t^{*h} around P_t^0 should exceed the volatility of P_t around P_t^0. Thus, comparison of the sample counterparts of the expressions in (9) and (10) provides a way of gauging the direction and substantive magnitude of departures from the null hypothesis.

3. IMPLEMENTATION AND RESULTS

In this section, we discuss details of the implementation of the test, the specification of the Monte Carlo experiment used to study the finite-sample properties of the test, and the data used to carry out the tests. We then present the results for the constant required rate of return case.

We consider two specifications for the naive forecast. First, following our previous paper, we generate the naive forecast P_t^0 by letting it equal what the discounted value of the infinite stream of future dividends would be if real dividends never changed from their most recently observed value, D_{t-1}. Thus,

$$(11) \qquad P_t^0 = \frac{(1+r)^{1/2}}{r} D_{t-1} .$$

As we noted in the introduction, dividends themselves are highly variable. To produce a potentially smoother naive forecast, we consider a specification where a thirty-year moving average of dividends is capitalized. Here,

$$(11') \qquad P_t^0 = \frac{(1+r)^{1/2}}{r} \frac{1}{30} \sum_{i=1}^{30} D_{t-i} .$$

The specification of the naive forecasts (11') is meant to capture the spirit of Shiller's (1981) original test while remaining statistically valid in our framework. If there is important mean-reversion in dividends, the naive forecast defined by (11') will be much smoother than that defined by (11). Centering on a potentially smoother naive forecast might make it easier for Shiller's excess volatility to emerge.[7]

---

[7]Additionally, Campbell and Shiller (1988b) find that long moving averages of earnings help to forecast stock returns. We use the moving average of dividends to try to capture the same phenomenon.

Now consider the choice of weights. Because the variables are growing over time, setting $W_t = 1$ for all t is likely to lead to severe heteroskedasticity and thus to inefficient estimates. We therefore weight by the market price; that is, we set $W_t = P_t$. Under this rule, that $(P_t - P_t^0)/W_t$ is uncorrelated with $(P_t^{*h} - P_t)/W_t$ can be written as

$$(12) \qquad E\left[\left(1 - \frac{P_t^0}{P_t}\right)\left(\frac{P_t^{*h}}{P_t} - 1\right)\right] = 0 .$$

Since P_t^0 is proportional to dividends, P_t^0/P_t is proportional to the dividend-price ratio. P_t^{*h}/P_t is the h-period excess holding return (plus one). By weighting by P_t, we are thus testing the prediction of the theory that the dividend-price ratio cannot help predict the excess holding return.[8]

The properties of the test statistic we present in the previous section have been established only asymptotically. In judging the outcome of the test procedure, it necessary to take into account the possible inadequacy of the asymptotic distribution for making inferences in finite samples. Because Flavin (1983), Kleidon (1986), and Marsh and Merton (1986) have shown that previous volatility tests suffer from severe finite sample biases, attention to small sample properties is particularly important in the case of volatility tests. We do note, however, that the

[8] In our previous paper, we choose dividends as the weights. We prefer to weight by price because it yields the natural interpretation just discussed. Weighting by dividends rather than prices has little effect on the results.

major criticisms of the finite sample properties of the original volatility tests, which mainly concern the treatment of trends, do not apply to our test.

To generate a distribution of our statistics under the null hypothesis (7), we postulate a process followed by dividends, draw a random realization of this process, and calculate prices and our test statistics. In order to address some of the criticisms of the earlier tests directly, we choose a process for dividends along lines suggested by Marsh and Merton. Suppose that the permanent component of dividends, denoted as η_t, follows a logarithmic random walk with drift. That is,

$$(13) \qquad \eta_t = (1+g)\eta_{t-1}\exp(\nu_t)$$

where g is the growth rate and where $\exp(\nu_t)$ is a lognormal random variable with mean one. Following Lintner, Marsh and Merton suggest that corporations adjust dividends slowly to news. Marsh and Merton maintain that this smoothness of dividends results in Shiller's test having misleading statistical properties. To allow for this smoothing, we assume that dividends are given by

$$(14) \qquad D_t = (1+g)\theta D_{t-1} + (1-\theta)\eta_t \ .$$

Iterating (14) forward and taking conditional expectations yields

(15) $$E_t D_{t+i} = (1+g)^{i+1}\theta^{i+1}D_{t-1} + (1-\theta)(1+g)^{i+1}\left[\frac{1-\theta^{k+1}}{1-\theta}\right]\eta_{t-1} .$$

Substituting (15) into (1) and (2), letting the holding period go to infinity, and collecting term yields the value of price under this dividend process and the conventional valuation model. That expression for price is

(16) $$P_t = (1+g)\left(\frac{1}{1+r}\right)^{1/2}\left[\frac{\theta}{1-\theta\frac{1+g}{1+r}}(D_{t-1} - \eta_{t-1}) + \frac{1}{1-\frac{1+g}{1+r}}\eta_{t-1}\right] .$$

Note that we are able to derive an exact expression for price that is linear in variables despite the fact that the driving process (14) is log-linear.[9] We consider cases where θ takes values 0.1 and 0.5, corresponding to more or less rapid adjustment of dividends.[10]

We now turn to applying our test to the data. We use annual data on the aggregate stock market from 1871 to 1988. Until 1926, the price and dividend data are the Cowles (1939) All-Stock Index. Since then, they are the Standard and Poor's composite. The financial variables are deflated by

---

[9] Scott (1985) carries out a similar exercise for an alternative, regression test of Shiller's model. He estimates θ to be approximately 0.5. In his specification of dividends, expected dividends are a nonlinear function of the forcing variables, so no simple, closed-form solution is available.

[10] In the simulations, the dividends and earnings processes are started at identical, fixed values. (Their unconditional distributions do not exist.) The variance of the innovation, ν_t, is chosen so that the variance of simulated dividends is expected to equal its historical value. The growth rate is set to zero and the required rate of return is set to the value assumed in computing P_t^*. Five hundred simulations were carried out using RATS software on an IBM 3083 computer.

the Producer Price Index. The data are an updated version of those used by Shiller (1981) and by Mankiw, Romer, and Shapiro (1985).

Table 1 shows the results for various holding periods with constant required rates of return of 5, 6, and 7 percent.[11] These rates encompass estimates of the mean return over the sample. Columns (2), (3), and (4) show the sample means of $\left(\frac{P_t^{*h} - P_t^0}{P_t}\right)^2$, $\left(\frac{P_t^{*h} - P_t}{P_t}\right)^2$, and $\left(\frac{P_t - P_t^0}{P_t}\right)^2$. The theory implies that, in expectation, the column (2) should exceed each of columns (3) and (4) and should equal their sum. Column (5) gives the $\chi^2(1)$ statistic for the null hypothesis that the equality holds. It is computed by squaring the difference between the entry in column (2) and the sum of the entries in columns (3) and (4) and then dividing this quantity by the estimated variance of the difference; the variance of the difference is computed using the Hansen-White-Domowitz procedure described above. The statistic thus tests whether α equals zero (equation 7). The final columns give the marginal significance level for the test of the equality restriction. Column (6) gives the p-value from the asymptotic $\chi^2(1)$ distribution. The p-values in columns (7) and (8) are calculated based on the Monte Carlo distributions under the two assumptions about the dividend process.

---

[11]We have calculated the statistics for a wider range of required rates of return (from 3 to 10 percent) and for a finer grid of rates (by tenths of a percent). The results are generally similar to those reported in the tables. The only noteworthy exception is that for extreme values of the discount rate the <u>level</u> of P_t^* is uniformly too high or too low relative to P_t, because the discount rate is too low or too high. A required rate of return of between six and seven makes the average of P_t^*/P_t unity in the sample.

Under the null hypothesis, columns (3) and (4) should not exceed column (2). In Table 1, for holding periods of 5, 10, and T-t years with the required rate of return equal to 5 percent and for holding periods of 10 and T-t years with the required return equal to 6 percent, column (3) does in fact exceed the column (2). For these cases, the naive forecast is a better forecast of realized returns than is the stock price itself. In the other cases, the forecast error inequality is not rejected. The second inequality, corresponding to the variance bound test, is not violated. The $E(P - P^0)^2$ figures are always less than the $E(P^{*h} - P^0)^2$ figures. Hence, these calculations do not indicate that the price is excessively volatile about the naive forecast.

The probability values for the test statistic that the equality holds in the sample indicate that the null hypothesis tends to be rejected, but not overwhelmingly so. Moreover, whether one rejects or not depends on the holding period used for the test and the assumed required rate of return. For r equal to 5 percent, the conventional valuation model is always rejected using tests of standard size. For r equal to 6 percent, it is only rejected for the longer holding returns. For r equal to 7, the hypothesis is not rejected, though it is close to rejection for the case of the holding period equal to the time to the end of the sample.

Note that the Monte Carlo experiment indicates that the asymptotic distribution often leads to reliable inference. On the other hand, in some cases the tests of standard size are misleading. The asymptotic distribution becomes more misleading as the horizon over which the statistic is computed increases. This illustrates the importance of careful attention to finite sample properties in this setting.

Finally, to try to come closer to the spirit of Shiller's original tests, Table 2 reports results analogous to those in Table 1 with both the naive forecast (P_t^0) and the weight (W_t) now based on a thirty-year moving average of dividends (see equation 11'). This obviates the possibility that our test statistics are misleading because of excess volatility in the stock price series used for detrending in the results reported in Table 1. The results reported in Table 2 are similar to those in Table 1. Two differences are worth noting. First, for long holding periods the null hypothesis in now not clearly rejected for r equal to 5 and 6 percent and is now rejected of r equal to 7 percent. Second, the source of the rejections has changed. The naive forecast is often not a better forecast of the perfect foresight price than is the stock price, but prices are more volatile around the naive forecast than in Table 1. Although we do not always find excess volatility in the sense that column (4) exceeds column (2), the volatility of prices plays more of a role in the rejection of equation (4) than it does in the results in Table 1.

4. VARIABLE REQUIRED RATE OF RETURN

The maintained assumption that the required rate of return on the stock market is a constant is highly restrictive. If one believes that changes in the required rate of return are an important cause of changes in stock prices, then the constant required rate of return model in equation

(2) is only of limited interest. In this section, we discuss a test in which the expected rate of return is allowed to vary.

The required rate of return on the stock market is unobserved. To formulate a test, we must use a model relating observed variables to the unobserved required rate of return. We assume that the required rate of return is equal to the riskless rate of interest plus a constant risk premium ϕ . This restriction is justified in a consumption-beta model if the relevant covariance matrix is constant through time (Hansen and Singleton, 1983).

Our formulation with a constant risk premium is much less restrictive than the standard assumption of fixed required rate of return. Yet it still restricts P_t^{*h} more than some theories suggest. One alternative would be to use the change in consumption to measure the required rate of return on the market. Theories using consumption to measure risk are very attractive a priori but have failed to account well for returns in the stock market (see Mankiw and Shapiro, 1986).

Because the one-period nominal interest rate is known in advance with certainty, it is convenient to carry out the analysis in terms of nominal instead of real stock prices. This procedure also avoids measurement error from use of the price indexes as deflators. Let R_t denote the nominal one-period riskless rate of interest. The one-period discount factor is then given by

$$(17) \qquad \gamma_t = 1/(1 + R_t + \phi) < 1 .$$

Let $\gamma_t^{(h)}$ denote the h-period nominal discount factor. That is,

$$(18) \qquad \gamma_t^{(h)} = \Pi_{i=1}^{h} \gamma_{t+i-1} .$$

The theory states that the expected holding return on the stock market equals the risk-adjusted interest rate. Thus

$$(19) \qquad P_t = E_t[\gamma_t P_{t+1} + \gamma_t^{1/2} D_t] .$$

Note that dividends are discounted by $\gamma_t^{1/2}$ because they are paid in the middle of the year $t+1$. In contrast to the previous section, here the price and dividend are nominal. Using the fact that γ_{t+1} is in agent's time $t+1$ information set and applying the law of iterated expectations, we can solve forward for P_t . We obtain

$$(20) \qquad P_t = E_t[\sum_{j=0}^{h-1} \gamma_{t+j}^{1/2} \gamma_t^{(j)} D_{t+j} + \gamma_t^{(h)} P_{t+h}]$$
$$\equiv E_t P_t^{*h} .$$

We adopt the naive forecast that nominal dividends are expected to grow at the risk-free rate. This leads to the naive forecast

$$(21) \qquad P_t^0 = \frac{(1+\phi)^{1/2}}{\phi} D_{t-1} .$$

The naive forecast reflects no information concerning variation in the interest rate. Hence, one would expect it to predict P_t^{*h} defined with a varying interest rate less well than it predicts P_t^{*h} defined with a fixed rate of return.

To perform the tests, we use the same horizons as in the constant required rate of return tests. We consider a range of values for the unobserved risk premium ϕ .

The data we use for the one-period interest rate are the annual commercial paper rate given in Friedman and Schwartz (1982) and updated from the Economic Report of the President. Ideally, we would use an asset with a guaranteed nominal return, such as Treasury Bills, but these are only available over a short time span. With these data, we compute the tests analogous to the ones with the constant required rate of return.

The results of the tests with the varying interest rates are presented in Table 3. Results for risk premia of 4, 5, and 6 percent are reported for the same holding periods as in Table 1. The probability values reported in the last two columns are computed under the same null hypothesis as in Table 1, that is, of constant required rate of return.

The results in Table 3 are not supportive of the hypothesis that stock prices are the expected present discounted value of future dividends. The results for the variance inequalities are roughly the same as with the fixed required rate of return case: the volatility inequality (column 2 versus column 4) is not rejected; the forecast error inequality (column 2 versus column 3) is rejected for long holding periods with the lower two risk premia. Yet, the hypothesis of stock market efficiency fails more dramatically when interest rate variation is taken into account. That is, the tests of the equalities indicate more serious rejections of the null than in Table 1.

Table 4 gives the test statistics for different values of the equity premium based on the moving average naive forecast and weight (equation 11'). As in the fixed required rate of return case (Tables 1 and 2), using the long moving average of dividends as weight and naive forecast leads to only marginal (five to ten percent) rejections of the null in all cases considered. Similarly, the volatility of the stock price (column 4) contributes more to the rejection of the null hypothesis when the naive forecast is based on the moving average of dividends. The choice of naive forecast affects the decomposition of variance between the two elements of the variance equality (4) and therefore the interpretation of the inequalities (9) and (10). Yet, we stress the robustness of the formal test of the equality to these decompositions.

5. COMPARISON WITH REGRESSION TESTS

Our tests of market efficiency are in the spirit of the "volatility test" literature begun by Shiller and LeRoy and Porter. A second strand of literature testing market efficiency focusses on regression tests. Regression tests of market efficiency over medium and long-term horizons have been the subject of considerable recent attention (Scott, 1985; Fama and French, 1988a, b; Campbell and Shiller, 1988a, b; and Flood, Hodrick, and Kaplan, 1986). In this section, we describe the relationship between our tests and regression tests and compare the empirical results that we obtain with those obtained using regression tests.

The implication of market efficiency that we test is that $(P_t^{*h} - P_t)/W_t$ be uncorrelated with $(P_t - P_t^0)/W_t$. The natural approach to testing this condition using a regression procedure would simply be to regress $(P_t^{*h} - P_t)/W_t$ on $(P_t - P_t^0)/W_t$ and a constant, that is

$$\frac{P_t^{*h} - P_t}{W_t} = \alpha + \beta \frac{P_t - P_t^0}{W_t} + e_t, \tag{22}$$

and then test the hypothesis that the coefficients are zero. Taking the weight W_t equal to P_t, the left-hand side variable of equation (22) is the return on holding the stock portfolio for h periods (with reinvested dividends earning the required rate r) and the right-hand side variable is one minus the dividend-price ratio capitalized at a constant rate.

To see the relationship between this regression test and our "pseudo volatility" test, consider estimating (22) without the constant term. Under the null hypothesis β is still zero. The estimate of β is given by

$$\hat{\beta} = \frac{\frac{1}{T}\sum_t \left(\frac{P_t - P_t^0}{W_t}\right)\left(\frac{P_t^{*h} - P_t}{W_t}\right)}{\frac{1}{T}\sum_t \left(\frac{P_t - P_t^0}{W_t}\right)^2} \tag{23}$$

Our pseudo volatility test focusses on the mean of the q_t's defined in equation (6). Their mean is given by

$$\bar{q} = \frac{2}{T}\sum_t \left[\frac{P_t - P_t^0}{W_t}\right]\left[\frac{P_t^{*h} - P_t}{W_t}\right]. \tag{24}$$

Comparison of (23) and (24) shows how our test is related to a regression test: our test focusses on the numerator of the regression coefficient in a regression estimated without a constant term. For the case of $P_t^0 = D_t/r$ and $W_t = P_t$ that we consider in our empirical work, $(P_t^{*h} - P_t)/W_t$ is simply $P_t^{*h}/P_t - 1$ and $(P_t - P_t^0)/P_t$ is $1 - (1/r)(D_t/P_t)$. Thus in this case our test is closely related to a regression of the excess holding return on the dividend-price ratio.

Compared with a regression test, our test has a disadvantage and an advantage. The disadvantage is that regression coefficients often lend themselves to natural interpretation. Fama and French (1988a), for example, discuss their results in terms of the fraction of variation in returns over various horizons that is predictable. In contrast, unless the inequalities (9) and (10) are violated, it is more difficult to interpret the results of our tests in terms of the economic magnitudes of the estimated departure from the null hypothesis. Yet, given the bias in the estimates of the regression coefficients and the understatement of their standard errors, direct appeal to ordinary least squares results is potentially quite misleading.

The advantage of our test is that it is likely to have better statistical properties. As we stress in the derivation of our test, both the expectation of the mean of the q_t's and its asymptotic standard error can be derived under minimal assumptions. Thus, the only difficult statistical issues in conducting inference involve the small sample properties of the standard errors. In the Monte Carlo experiments that we report above, we find that the asymptotic standard errors provide tolerably good guides to the finite sample properties of the tests.

In contrast, regression estimates of β from (22) provide extremely biased estimates of the true coefficient when P_t^{*h} is defined over long horizons. The difficulty is that although under the null hypothesis the right-hand side variable is uncorrelated with the contemporaneous error term, it is correlated with the lagged errors. Mankiw and Shapiro (1986) demonstrate that orthogonality conditions such as $\beta=0$ is equation (22) are rejected too often in finite sample if the innovations in the right-hand side variable and the left-hand side variable are highly correlated at lags and if the right-hand side variable is high serially correlated. These conditions obtain in the test of the present value model: the same process drives stock prices and dividends, so their innovations are correlated. The bias toward rejection stems both from the fact that the point estimates are biased negatively and from the fact that the standard errors of the regression coefficients are understated even if an asymptotically correct covariance matrix is calculated. As we demonstrate below, the biases involved are potentially severe.

Recent work by Phillips (1988a, b) shows that the biases worsen as the lagged correlations between the regressors and the disturbances increase. In the long holding period cases (h large), these correlations are high because of the overlap of the observations. Therefore, one should expect to find that the finite sample performance of these tests worsens as the holding period increases. Phillips (1988a) also stresses that the R^2 is too large in finite sample in these settings. The focus on that statistic in the stock market forecastability literature is therefore problematic.

To investigate the differences between our tests and regression tests, we perform several regression tests using the same data, the same sample period, and the same Monte Carlo experiment to assess small sample properties that we use in our tests. We begin with the test, due to Scott (1985), most comparable to ours: the regression (22) of $(P_t^{*h}-P_t)/P_t$ on a constant and $(P_t-P_t^0)/P_t$, that is, the dividend-price normalized by a constant required rate of return (r). Table 5 reports the results. The table shows that finite sample biases are extremely important; the asymptotic distribution provides an extremely poor guide to the finite sample properties of the test statistics. For example, for the case of r equal to five percent and h = T - t, a chi-squared statistic of 58.4, which implies a rejection at virtually the zero percent level using the asymptotic distribution, corresponds to a rejection at only the 0.8 percent level in the Monte Carlo experiment.[12] The estimates do imply stronger rejections of the null hypothesis than those obtained using our pseudo volatility tests

[12] In other words, the biased-corrected chi-squared statistic is around five, or over ten times smaller than the nominal value of 58.4.

according to the Monte Carlo distributions; the final two columns of the table show that the null hypothesis is rejected at between the one and five percent level. But the fact that the required correction for bias is so large suggests that the precise confidence levels implied by our Monte Carlo experiment should be viewed with considerable caution; we have no reason to rule out the possibility that the bias correction needed would be somewhat larger, or smaller, under other reasonable specifications for the dividend process.

In Table 6, we consider the analogue of the test for the variable required rate of return case. Specifically, we perform a test that is the same as that in Table 5 except that P_t^{*h} is now computed using a constant equity premium rather than a constant required rate of return. As we stress above, tests of the constant equity premium model have the advantages that they focus on a null hypothesis with stronger theoretical foundations than the constant required rate of return model and that they do not require the use of a potentially imperfect price level series to deflate the stock prices and dividends. The null hypothesis is rejected less strongly in Table 6 than in Table 5; typical rejections are in the five to ten percent range. In contrast, our tests in the previous section yielded stronger rejections for the variable rather than fixed return case. And, as before, the asymptotic distributions provide extremely poor guides to the finite sample characteristics of the regression tests.

Finally, Table 7 reports the results of regression tests based on Fama and French (1988a). Here the h-period gross holding return is regressed on itself lagged h periods. These tests have the advantage of not requiring an independently specified required rate of return. There is little evidence against the null. Again, the asymptotic distribution is an extremely poor guide to statistical inference.[13]

Poterba and Summers and Fama and French emphasize rejections of the null hypothesis at horizons of three to four years. But in our longer sample, at these horizons the null is not rejected even with conventional statistical inference. Kim, Nelson, and Startz (1989) show convincingly that the outcome of these tests is strongly conditioned by the choice of sample interval. As the horizon is lengthened, in our sample, the spurious rejections do emerge strikingly.

All of these results are consistent with what other authors employing regression tests have found. Although the point estimates and asymptotic distributions often suggest considerable forecastability of returns and overwhelming rejections of the orthogonality conditions, when careful attention is paid to the finite sample properties of the test there proves to be only moderate evidence against the null. Thus regression tests suggest the same conclusion as implied by our tests of volatility and forecastability: there is some, but not overwhelming, evidence against the joint null hypothesis of market efficiency and either a constant required rate of return or a constant equity premium.[14]

---

[13] A natural variation on Scott's test, employed by Fama and French (1988b), is to regress returns over various horizons on the dividend-price ratio. Like the test in Fama and French (1988a), this specification has the advantage that the required rate of return need not be specified. The test yields results quite similar to those in Table 5.

[14] Investigations using "variance ratio" test (Poterba and Summers, 1988; Lo and MacKinlay, 1988) reach similar conclusions. West (1988), on the other hand, reaffirms Shiller's original conclusion by comparing the

6. CONCLUSION

This paper investigates whether aggregate stock prices exhibit excess volatility or predictable movements over horizons of a year or more. Building on work in Mankiw, Romer, and Shapiro (1985), we develop a testing procedure involving sample statistics whose expectations and asymptotic distributions under the null hypothesis of market efficiency can be found under minimal auxiliary assumptions. In addition, we employ a Monte Carlo procedure (the specifics of which are inspired by the arguments made by critics of previous variance bounds tests) to conduct finite sample inference. We find that, contrary to the predictions of the hypothesis of market efficiency, the difference between the current level of stock prices and a "naive forecast" based solely on current dividends is not orthogonal to future returns; in fact, over horizons of five years or more the naive forecast often outperforms the market price as a predictor of the perfect

---

variability of a forecast of dividend streams based on a restricted information set with the variability of prices. He concludes that stock prices are many times too volatile (see also West, 1987). The null hypothesis that West tests is considerably more restrictive than that considered in the other papers. First, speculative bubbles are not permitted under the null. Second, he parameterizes the dividend process in a particular way, including specifying an ARIMA process on levels rather than logarithms. His overwhelming rejections could be caused by these restrictive features of his tests. Moreover, as our introduction emphasizes, his finding of small dividend innovations relative to price innovations can only arise from estimates of the dividend process that are substantially mean reverting.

Campbell and Shiller (1988a) report qualitatively similar results to those of Fama and French concerning the forecastability of returns (an R-squared of roughly twenty percent for ten-year returns, for example). They report test statistics with very small p-values based on asymptotic distributions. In a Monte Carlo study, Campbell and Shiller (1988c) show, however, that their linear Wald test correspond to rejections of the conventional valuation model at the five to ten percent level. These rejections, which are similar to ours, are not overwhelming.

foresight price. These rejections of market efficiency are moderately, but not overwhelmingly, statistically significant. Finally, we find that the rejections of the model are in fact slightly stronger when the focus is on excess returns rather than on simple returns.

Our statistical appraisal encompasses other recent studies of stock market valuation. These tests in general find a considerable amount of forecastability to returns (and excess returns) over several year horizons. Our findings show, however, that the rejections are generally not overwhelming; the null hypothesis of efficient markets is typically roughly on the brink of rejection at conventional significance levels.

Although we present tests of the present value relationship that are robust to the main criticisms of earlier tests, the data are not decisive. Consideration of lack of power against interesting alternatives (Summers, 1986; Shiller and Perron, 1986) might suggest that one would expect only marginal rejections of the null. On the other hand, variations in the equity premium or revisions of expectations about rare events could account for marginal rejections even if the null is true. We find rejections of the model that are statistically significant using conventionally-sized tests, but are not overwhelming. Moreover, inspection of the variance equality and the component inequalities shows that the magnitude of the deviations from the null are not large. Therefore, contrary to the common claim that stock prices are grossly too volatile, the aggregate data are not glaringly inconsistent with market efficiency.

APPENDIX

Overview. Merton, in his analysis of our earlier paper, argues that the true standard errors of our statistics are so large that the point estimates we present do not constitute rejections of the model. Since our previous paper considers the inequalities (9) and (10) for the variable holding period case (that is, $h_t = T-t$), Merton considers these statistics. He explores their sampling properties for a case in which stocks do not pay dividends. In the absence of dividends, the statistics are governed entirely by the end-of-sample value of the stock price; Merton finds that as a result, the standard errors of the statistics diverge as the sample size grows.

In this Appendix we show that Merton's conclusions no longer hold if dividends are paid in sufficient quantity. Let S1 denote the difference between columns (2) and (3) of Table 1 (that is, the statistic for the forecast error inequality (9)); let S2 denote the difference between columns (2) and (4) of Table 1 (the volatility inequality (10)); and let S3 denote the mean of the equality (5). We show that if the dividend-price ratio exceeds one-half the instantaneous variance (σ^2) of the diffusion process for stock prices, the standard errors of S1 and S3 converge in probability limit to zero, so the tests based on them are consistent. Convergence of S2 requires that the dividend-price ratio exceed one and one-half times σ^2. Merton [p. 115] suggests that 0.04 for a value of σ^2 at annual rate. In our data, the variance of stock price growth is 0.031. The dividend-price ratio is, on average, 0.051, so sufficient dividends are paid for S1 and S3 to yield valid tests. On the other hand, these calculations suggest that the failure to reject the volatility inequality (S2) may be the result of lack of power.

In addition, we show that S1, S2, and S3 for fixed holding periods have non-degenerate asymptotic normal distributions regardless of whether dividends are paid.

Following Merton, assume that $D(t) = \rho P(t)$, that $P(t)$ follows geometric Brownian motion with drift $r-\rho$ and instantaneous variance σ^2, and that $P^0(t) = 0$, where $D(t)$ is dividends, $P(t)$ price, and $P^0(t)$ the naive forecast. We first discuss the case in which $P^*(t)$ is the perfect foresight price for the strategy of holding the stock to the end of the sample period and then consider the case in which the stock is held for some fixed horizon h.

<u>Holding the Stock to the End of the Sample Period</u>. Consider first S3. We claim that if $\rho > \sigma^2/2$, then

(A-1) $$\lim_{T \to \infty} E[S3^2] = 0 ,$$

and thus that S3 converges in quadratic mean to its expectation of zero.

With the naive forecast assumed to always equal zero, the test statistic simplifies to

(A-2) $$S3 = \frac{1}{T} \int_{t=0}^{T} \left(\frac{P^*(t)}{P(t)} - 1\right) dt.$$

Thus

$$\text{(A-3)} \qquad E[S3^2] = \frac{1}{T^2}\int_{t=0}^{T}\int_{t'=0}^{T} E\left[\frac{P^*(t)}{P(t)}\,\frac{P^*(t')}{P(t')}\right] dt'dt$$

$$- 2\,\frac{1}{T}\int_{t=0}^{T} E\left[\frac{P^*(t)}{P(t)}\right] dt \quad + \quad 1 .$$

The expectation of $P^*(t)/P(t)$ is one. To find $E\left[\frac{P^*(t)}{P(t)}\,\frac{P^*(t')}{P(t')}\right]$, assume without loss of generality that $t' \geq t$ and define $\Delta t = t' - t$. $P^*(t)/P(t)$ can be written

$$\text{(A-4)} \qquad \frac{P^*(t)}{P(t)} = \int_{s=t}^{t+\Delta t} \rho \exp\left[(-\rho-\frac{\sigma^2}{2})(s-t) + \sigma(Z(s) - Z(t))\right]ds$$

$$+ \exp\left[(-\rho-\frac{\sigma^2}{2})\Delta t + \sigma(Z(t+\Delta t) - Z(t))\right] \frac{P^*(t+\Delta t)}{P(t+\Delta t)} ,$$

where Z is Wiener. The first term reflects innovations between times t and $t + \Delta t$, the second innovations between $t + \Delta t$ and T. Because $Z(s) - Z(t)$ $(s \leq t + \Delta t)$ and $Z(t+\Delta t) - Z(t)$ are known as of time $t + \Delta t$, they are independent of $P^*(t+\Delta t)/P(t+\Delta t)$. It follows that

$$\text{(A-5)} \qquad E\left[\frac{P^*(t)}{P(t)}\,\frac{P^*(t+\Delta t)}{P(t+\Delta t)}\right] =$$

$$E\left[\int_{s=t}^{t+\Delta t} \rho \exp[(-\rho-\frac{\sigma^2}{2})(s-t) + \sigma(Z(s) - Z(t))]ds\right] E\left[\frac{P^*(t+\Delta t)}{P(t+\Delta t)}\right]$$

$$+ E\left[\exp[(-\rho-\frac{\sigma^2}{2})\Delta t + \sigma(Z(t+\Delta t) - Z(t))]\right] E\left[\left(\frac{P^*(t+\Delta t)}{P(t+\Delta t)}\right)^2\right]$$

$$= \int_{s=t}^{t+\Delta t} \rho \exp[-\rho(s-t)] \, dt \; E\left[\frac{P^*(t+\Delta t)}{P(t+\Delta t)}\right] + e^{-\rho\Delta t} E\left[\left(\frac{P^*(t+\Delta t)}{P(t+\Delta t)}\right)^2\right]$$

$$= 1 - e^{-\rho\Delta t} + e^{-\rho\Delta t} E\left[\left(\frac{P^*(t+\Delta t)}{P(t+\Delta t)}\right)^2\right] ,$$

where the intermediate step uses the fact that $\exp[(-\rho-\frac{\sigma^2}{2})(t_2-t_1) + \sigma(Z(t_2) - Z(t_1))]$ is lognormal.

Now

$$\text{(A-6)} \qquad \frac{P^*(t)}{P(t)} = \int_{s=t}^{T} \frac{e^{-r(s-t)}\rho P(s)}{P(t)} \, ds + \frac{e^{-r(T-t)}P(T)}{P(t)}$$

$$= \int_{s=t}^{T} \rho \exp\left[\left(-\rho-\frac{\sigma^2}{2}\right)(s-t) + \sigma\left(Z(s) - Z(t)\right)\right] ds$$

$$+ \exp\left[\left(-\rho-\frac{\sigma^2}{2}\right)(T-t) + \sigma\left(Z(T)-Z(t)\right)\right] .$$

Thus

$$\text{(A-7)} \qquad E\left[\left(\frac{P^*(t)}{P(t)}\right)^2\right] = 2 \int_{s=t}^{T} \int_{s'=s}^{T} \rho^2 E\Big[\exp[(-\rho-\frac{\sigma^2}{2})(s-t) + \sigma(Z(s)-Z(t))]$$

$$\exp[(-\rho-\frac{\sigma^2}{2})(s'-t)+\sigma(Z(s')-Z(t))]\Big] ds' \, ds$$

$$+ 2 \int_{s=t}^{T} \rho E\Big[\exp[(-\rho-\frac{\sigma^2}{2})(s-t) + \sigma(Z(s)-Z(t))]$$

$$\exp\left[\left(-\rho-\frac{\sigma^2}{2}\right)(T-t) + \sigma(Z(T)-Z(t))\right]\Bigg]ds$$

$$+ E\left\{\left(-\rho-\frac{\sigma^2}{2}\right)(T-t) + \sigma\Big[Z(T)-Z(t)\Big]\right\}^2$$

$$= 2\int_{s=t}^{T}\int_{s'=s}^{T}\rho^2\exp\left[\left(-\rho-\frac{\sigma^2}{2}\right)(s-t) + \left(-\rho-\frac{\sigma^2}{2}\right)(s'-t)\right.$$

$$\left. + \frac{3}{2}\sigma^2(s-t) + \frac{1}{2}\sigma^2(s'-t)\right]ds'ds$$

$$+ 2\int_{s=t}^{T}\rho\exp\left[(-\rho-\frac{\sigma^2}{2})(s-t) + (-\rho-\frac{\sigma^2}{2})(T-t)\right.$$

$$\left. + \frac{3}{2}\sigma^2(s-t) + \frac{1}{2}\sigma^2(T-t)\right]ds$$

$$+ \exp\left[2(-\rho-\frac{\sigma^2}{2})(T-t) + 2\sigma^2(T-t)\right]$$

$$= \frac{2\rho - \sigma^2\exp[-(2\rho-\sigma^2)(T-t)]}{2\rho - \sigma^2},$$

where the second step again uses lognormality and where the final step involves tedious algebra.

Substituting this result into (A-5) yields

$$\text{(A-8)}\qquad E\left[\frac{P^*(t)}{P(t)}\,\frac{P^*(t+\Delta t)}{P(t+\Delta t)}\right] =$$

$$(1-e^{-\rho\Delta t}) + e^{-\rho\Delta t}\,\frac{2\rho - \sigma^2[\exp[-(2\rho-\sigma^2)(T-(t+\Delta t))]}{2\rho - \sigma^2} .$$

After substituting (A-8) into (A-3) and integrating, straightforward but tedious algebra shows that if $\rho > \sigma^2/2$ the limit as T approaches infinity of the expectation of $S3^2$ is zero.

The proof of the consistency of S1 parallels that for S3. When the naive forecast is zero, S1 can be written as simply

$$\text{(A-9)} \qquad S1 = \frac{1}{T}\int_{t=0}^{T} 2\frac{P^*(t)}{P(t)}dt - 1 .$$

Since $E[P^*(t)/P(t)] = 1$, the expectation of S1 is one. Analysis similar to that for S3 shows that $E[S1^2]$ converges to zero as T becomes large.

Finally, the remaining test statistic is

$$\text{(A-10)} \qquad S2 = \frac{1}{T}\int_{t=0}^{T} \left[\frac{P^*(t)}{P(t)}\right]^2 dt - 1 .$$

$S2^2$ thus involves terms in $(P^*(t)/P(t))^4$. One can show that the expectation of this fourth moment does not diverge as T becomes large as long as $\rho > \frac{3}{2}\sigma^2$. This condition proves to be necessary to be and sufficient for the consistency of S2.

Fixed Holding Period. Consider the test statistic based on a fixed holding time of h periods. Assume for simplicity that the statistics are constructed using only every h^{th} observation. The test statistic corresponding to S1, for example, would be

$$\text{(A-11)} \qquad S1(h) = \frac{1}{N}\sum_{i=0}^{N-1}\left[\frac{P^{*h}(ih) - P^{0}(ih)}{P(ih)}\right]^2 - \frac{1}{N}\sum_{i=0}^{N-1}\left[\frac{P^{*h}(ih) - P(ih)}{P(ih)}\right]^2,$$

where $N = T/h$ and $P^{*h}(ih)$ is the perfect foresight price for the strategy of buying the stock at time ih and holding it until time ih + h.

Because $P^{*h}(ih)$ is known as of time (i+1)h, the expectation of $P^{*h}(ih)[P^{*h}((i+1)h)-P((i+1)h)]$ is zero. Under the remaining assumptions employed by Merton, it follows that each of the test statistics S1(h), S2(h), and S3(h) is the sum of independent and identically distributed random variables with finite mean and variance. The statistics thus converge to their means (which are positive for S1(h) and S2(h), zero for S3(h)) and are asymptotically normal. Note that this argument requires no assumptions about the payment of dividends. Assuming that every observation rather than every h^{th} observation is used does not alter the conclusions. (We have not been able to establish any results concerning asymptotic normality for the case in which the holding period extends to the end of the sample period. Our conjecture is that the assumptions that imply consistency also imply asymptotic normality.)

Table 1

Volatility and Forecastability Tests

Naive Forecast based on Current Dividends

Constant Required Rate of Return (r)

(1)	(2)		(3)		(4)	(5)	(6)	(7)	(8)
							p-values		
h	$E\left[\frac{P_t^{*h} - P_t^0}{P_t}\right]^2$	=	$E\left[\frac{P_t^{*h} - P_t}{P_t}\right]^2$	+	$E\left[\frac{P_t - P_t^0}{P_t}\right]^2$	χ^2	asym.	θ=.1	θ=.5
			r = 5 percent						
1	.096		.030		.082	4.54	.003	.028	.036
2	.107		.066		.081	10.86	.001	.000	.000
5	.148		.179		.083	10.72	.001	.004	.002
10	.298		.406		.086	5.82	.016	.044	.044
T-t	.295		.446		.082	8.95	.003	.034	.050
			r = 6 percent						
1	.105		.029		.081	.42	.517	.516	.540
2	.124		.061		.080	1.78	.182	.188	.224
5	.175		.154		.080	5.81	.016	.046	.064
10	.300		.314		.083	4.78	.029	.064	.068
T-t	.194		.199		.082	7.59	.006	.060	.052
			r = 7 percent						
1	.141		.028		.116	.08	.777	.718	.780
2	.162		.058		.115	.43	.512	.564	.558
5	.205		.135		.114	2.01	.156	.236	.248
10	.290		.249		.117	2.39	.122	.188	.216
T-t	.141		.118		.117	5.16	.023	.096	.100

Note: Col. (1), horizon; cols. (2)-(4), sample analogues of terms in equation 4 (weighted by price); col. (5), $\chi^2(1)$ test statistic for test that col. (2) equals sum of cols. (3) and (4); col. (6), asymptotic p-value for test statistic; cols. (7) and (8), small sample p-values for test statistic (based on Monte Carlo experiment described in text [θ parameterizes adjustment of dividend]).

Table 2

Volatility and Forecastability Tests

Naive Forecast and Weight based on 30-Year Moving Average of Dividends

Constant Required Rate of Return (r)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
					p-values		
h	$E\left[\frac{P_t^{*h} - P_t^0}{P_t^0}\right]^2$	$= E\left[\frac{P_t^{*h} - P_t}{P_t^0}\right]^2$	$+ E\left[\frac{P_t - P_t^0}{P_t^0}\right]^2$	χ^2	asym.	θ=.1	θ=.5
			r = 5 percent				
1	.428	.059	.419	1.65	.198	.212	.194
2	.431	.122	.417	2.51	.113	.192	.192
5	.436	.251	.430	3.00	.083	.204	.206
10	.400	.449	.458	3.59	.058	.156	.162
T-t	.403	.339	.418	3.93	.047	.274	.284
			r = 6 percent				
1	.784	.083	.791	1.89	.170	.168	.162
2	.768	.170	.786	2.69	.101	.174	.170
5	.718	.341	.807	3.15	.076	.184	.192
10	.593	.593	.858	3.21	.073	.192	.196
T-t	.431	.352	.791	3.46	.063	.304	.300
			r = 7 percent				
1	1.267	.111	1.316	2.60	.107	.102	.096
2	1.208	.226	1.307	3.49	.062	.110	.110
5	1.050	.450	1.337	3.89	.049	.128	.128
10	.782	.784	1.416	3.51	.061	.174	.168
T-t	.462	.525	1.317	4.19	.041	.254	.248

See note to Table 1.

Table 3

Volatility and Forecastability Tests

Naive Forecast based on Current Dividends

Variable Required Rate of Return with Constant Equity Premium (ϕ)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
					p-values		
h	$E\left[\frac{P_t^{*h} - P_t^0}{P_t}\right]^2$ =	$E\left[\frac{P_t^{*h} - P_t}{P_t}\right]^2$ +	$E\left[\frac{P_t - P_t^0}{P_t}\right]^2$	χ^2	asym.	θ=.1	θ=.5
			ϕ = 4 percent				
1	.195	.030	.189	2.80	.094	.090	.128
2	.192	.063	.190	7.95	.005	.006	.000
5	.177	.134	.195	7.36	.007	.026	.002
10	.254	.258	.200	3.95	.047	.080	.022
T-t	.301	.460	.187	5.20	.023	.086	.050
			ϕ = 5 percent				
1	.100	.029	.083	2.28	.131	.098	.158
2	.107	.060	.083	9.44	.002	.010	.018
5	.112	.119	.084	11.12	.001	.044	.000
10	.186	.208	.088	9.55	.002	.092	.026
T-t	.195	.248	.083	8.57	.003	.088	.024
			ϕ = 6 percent				
1	.102	.029	.085	1.94	.164	.136	.174
2	.111	.057	.084	6.47	.011	.000	.044
5	.114	.109	.083	12.55	.000	.002	.000
10	.165	.177	.086	9.37	.002	.026	.022
T-t	.135	.172	.085	10.99	.001	.042	.030

See note on Table 1.

Table 4

Volatility and Forecastability Tests

Naive Forecast and Weight based on 30-Year Moving Average of Dividends

Variable Required Rate of Return with Constant Equity Premium (ϕ)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
					p-values		
h	$E\left[\frac{P_t^{*h} - P_t^0}{P_t^0}\right]^2$	$= E\left[\frac{P_t^{*h} - P_t}{P_t^0}\right]^2$	$+ E\left[\frac{P_t - P_t^0}{P_t^0}\right]^2$	χ^2	asym.	θ=.1	θ=.5
			ϕ = 4 percent				
1	.558	.061	.552	1.46	.226	.232	.232
2	.551	.120	.530	1.84	.175	.272	.270
5	.535	.231	.514	2.04	.153	.300	.308
10	.552	.435	.519	2.19	.139	.358	.342
T-t	.565	.453	.565	3.40	.065	.314	.306
			ϕ = 5 percent				
1	1.163	.094	1.184	1.83	.176	.168	.166
2	1.118	.182	1.141	2.27	.132	.204	.204
5	1.008	.341	1.105	2.41	.121	.268	.262
10	.920	.613	1.103	2.08	.150	.366	.366
T-t	.791	.513	1.210	2.87	.090	.352	.352
			ϕ = 6 percent				
1	2.008	.134	2.095	2.69	.101	.102	.088
2	1.887	.256	2.027	3.34	.068	.116	.110
5	1.592	.480	1.964	3.37	.066	.170	.158
10	1.306	.849	1.953	2.53	.112	.288	.290
T-t	1.014	.744	2.138	3.62	.057	.274	.286

See note to Table 1.

Table 5

Scott/Fama-French Regressions

$$(P_t^{*h} - P_t)/P_t = \alpha + \beta(P_t - P_t^0)/P_t + e_t$$

Constant Required Rate of Return (r)

(1)	(2)	(3)	(4)	(5)	(6)	(7)
					p-values	
h	α	β	χ^2	asym.	.1	.5
			r = 5 percent			
1	.032 (.016)	-.090 (.046)	3.85	.050	.072	.076
2	.064 (.027)	-.232 (.080)	8.46	.004	.012	.034
5	.135 (.060)	-.658 (.157)	17.52	.000	.000	.008
10	.233 (.099)	-1.083 (.287)	14.25	.000	.034	.040
T-t	.420 (.056)	-1.380 (.180)	58.49	.000	.008	.008
			r = 6 percent			
1	.040 (.018)	-.107 (.054)	3.86	.050	.066	.076
2	.089 (.032)	-.273 (.094)	8.50	.004	.012	.034
5	.210 (.072)	-.754 (.179)	17.65	.000	.000	.008
10	.336 (.125)	-1.193 (.313)	14.51	.000	.034	.040
T-t	.380 (.070)	-1.260 (.173)	52.88	.000	.008	.008
			r = 7 percent			
1	.047 (.023)	-.123 (.063)	3.87	.049	.066	.076
2	.113 (.041)	-.311 (.107)	8.54	.003	.012	.034
5	.277 (.090)	-.840 (.199)	17.78	.000	.000	.008
10	.416 (.154)	-1.281 (.333)	14.78	.000	.034	.040
T-t	.324 (.082)	-1.158 (.170)	46.55	.000	.008	.012

Table 5 (continued)

Note: Col. (1), horizon; cols. (2) and (3), estimated regression coefficients, col. (4), $\chi^2(1)$ test statistic for test that $\beta=0$; col. (5), asymptotic p-value for test statistic; cols. (6) and (7), small sample p-values for test statistic (based on Monte Carlo experiment described in text [θ parameterizes adjustment of dividend]).

Table 6

Scott/Fama-French Regressions

$$(P_t^{*h} - P_t)/P_t = \alpha + \beta(P_t - P_t^0)/P_t + e_t$$

Variable Required Rate of Return with Constant Equity Premium (ϕ)

(1)	(2)	(3)	(4)	(5)	(6)	(7)
					p-values	
h	α	β	χ^2	asym.	.1	.5
			ϕ = 4 percent			
1	.000 (.017)	-.061 (.042)	2.05	.153	.072	.184
2	-.010 (.028)	-.175 (.059)	8.80	.003	.006	.026
5	-.049 (.051)	-.453 (.128)	12.49	.000	.016	.030
10	-.026 (.091)	-.543 (.180)	9.14	.002	.056	.056
T-t	.044 (.099)	-.869 (.267)	10.58	.001	.118	.134
			ϕ = 5 percent			
1	.006 (.016)	-.075 (.052)	2.05	.152	.160	.182
2	.014 (.027)	-.214 (.072)	8.84	.003	.006	.026
5	.019 (.052)	-.541 (.152)	12.63	.000	.016	.030
10	.028 (.107)	-.627 (.205)	9.33	.002	.062	.056
T-t	.051 (.114)	-.823 (.265)	9.66	.002	.122	.140
			ϕ = 6 percent			
1	.012 (.020)	-.089 (.062)	2.05	.152	.160	.182
2	.038 (.032)	-.252 (.084)	8.88	.003	.006	.026
5	.080 (.066)	-.620 (.173)	12.76	.000	.016	.030
10	.069 (.127)	-.695 (.225)	9.53	.002	.062	.052
T-t	.027 (.128)	-.769 (.258)	8.88	.003	.128	.146

Note: See Table 5.

Table 7

Fama-French Regressions

$$R_t^h = \alpha + \beta R_{t-h}^h + e_t$$

(1)	(2)	(3)	(4)	(5)	(6)
				p-values	
h	α	β	χ^2	asym.	Monte Carlo
1	1.042 (.117)	.036 (.105)	.12	.729	.702
2	1.249 (.099)	-.065 (.089)	.52	.471	.530
5	1.511 (.241)	-.031 (.157)	.04	.845	.898
10	2.816 (.540)	-.340 (.160)	4.49	.034	.158
15	5.112 (.845)	-.698 (.166)	17.60	.000	.052

Note: $R_t^h = \Pi_{i=0}^{h-1} r_{t+i}$ where r_{t+i} is the one-period real stock return defined as $r_t = (P_{t+1}+\omega D_t)/P_t$. As described in the text, P_t is the January stock price, D_t is the calendar-year dividend, and $\omega = \sqrt{1.06}$ to correct for the fact that dividends are paid during the year rather than at the end of the year. The Monte Carlo p-value is invariant to the value of θ.

REFERENCES

Ackley, Gardner, "Commodities and Capital: Prices and Quantities," American Economic Review 73 (March 1983) 1-16.

Campbell, John Y., and Shiller, Robert J., "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," Review of Financial Studies 1 (1988) 195-228. (a)

Campbell, John Y., and Shiller, Robert J., "Stock Prices, Earnings, and Expected Dividends," Journal of Finance 43 (July 1988) 661-76. (b)

Campbell, John Y., and Shiller, Robert J., "The Dividend Ratio Model and Small Sample Bias: A Monte Carlo Study," NBER Technical Working Paper No. 67, 1988. (c)

Cowles, Alfred, and Associates, Common-Stock Indexes, 2nd. ed., Bloomington, Ind.: Principia Press, 1939.

Fama, Eugene F. and French, Kenneth R., "Permanent and Temporary Components of Stock Prices," Journal of Political Economy 96 (April 1988) 246-73. (a)

Fama, Eugene F. and French, Kenneth R., "The Dividend Yields and Expected Stock Returns," Journal of Financial Economics 22 (1988) 27-59. (b)

Flavin, Marjorie A., "Excess Volatility in Financial Markets: A Reassessment of the Empirical Evidence," Journal of Political Economy 91 (December 1983) 929-56.

Flood, Robert P. and Hodrick, Robert J., "Asset Price Volatility, Bubbles, and Process Switching," Journal of Finance 41 (September 1986) 831-42.

Flood, Robert P., Hodrick, Robert J., and Kaplan, Paul, "An Evaluation of Recent Evidence on Stock Market Bubbles," unpublished paper, Northwestern University, 1986.

Friedman, Milton and Schwartz, Anna J., Monetary Trends, Chicago: University of Chicago Press, 1982.

Hansen, Lars P., "Large Sample Properties of Generalized Methods of Moments Estimators," Econometrica 50 (1982) 1029-54.

Kleidon, Allan W., "Variances Bounds Tests and Stock Price Valuation Models," Journal of Political Economy 94 (October 1986) 953-1001.

LeRoy, Stephen F. and Porter, Richard D., "The Present Value Relation: Tests Based on Implied Variance Bounds," Econometrica 49 (May 1981) 555-574.

Lo, Andrew W. and MacKinlay, A. Craig, "Stock Market Prices of Not Follow Random Walks: Evidence from a New Specification Test," Review of Financial Studies 1 (1988) 41-66.

Mankiw, N. Gregory, Romer, David, and Shapiro, Matthew D., "An Unbiased Reexamination of Stock Market Volatility," Journal of Finance 40 (July 1985) 677-87.

Mankiw, N. Gregory and Shapiro, Matthew D., "Do We Reject Too Often? Small Sample Bias in Tests of Rational Expectations Models," Economics Letters 20 (1986) 139-145.

Marsh, Terry A. and Merton, Robert C., "Dividend Variability and Variance Bounds Tests for the Rationality of Stock Market Prices," American Economic Review 76 (June 1986) 483-503.

Merton, Robert C., "On the Current State of the Stock Market Rationality Hypothesis," in Macroeconomics and Finance: Essays in Honor of Franco Modigliani, Rudiger Dornbusch, Stanley Fischer, and John Bossons, eds. Cambridge: MIT Press, 1987.

Kim, Myung Jig, Nelson, Charles, R., and Startz, Richard, "Mean Reversion in Stock Prices: A Reappraisal of the Empirical Evidence," NBER Working Paper No. 2795 (December 1988).

Newey, Whitney K. and West, Kenneth D., "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," Econometrica 55 (May 1987) 703-8.

Phillips, Peter C.B., "Regression Theory for Near-integrated Times Series.," Econometrica 56 (September 1988) 1021-1043. (a)

Phillips, Peter C.B., "Optimal Inference in Cointegrated Systems," Cowles Foundation Discussion Paper No. 866, 1988. (b)

Poterba, James and Summers, Lawrence H., "Mean Reversion in Stock Prices: Evidence and Implications," Journal of Financial Economics 22 (1988).

Scott, Louis O., "The Present Value Model of Stock Prices: Regression Tests and Monte Carlo Results," Review of Economics and Statistics 67 (1985) 599-604.

Shiller, Robert J., "Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?" American Economic Review 71 (June 1981) 421-36.

Shiller, Robert J., "The Probability of Gross Violations of a Present Value Variance Inequality," Journal of Political Economy 96 (October 1988) 1089-92.

Shiller, Robert J., and Perron, Pierre, "Testing the Random Walk Hypothesis: Power versus Frequency of Observations," Economic Letters 18 (1985) 381-6.

Summers, Lawrence H., "Does the Stock Market Rationally Reflect Fundamental Values?" Journal of Finance 41 (July 1986) 591-601.

Warren, George F. and Pearson, Frank A., Gold and Prices. New York: Wiley, 1935.

West, Kenneth D., "A Specification Test for Speculative Bubbles," Quarterly Journal of Economics 102 (August 1987) 553-80.

West, Kenneth D., "Dividend Innovations and Stock Price Volatility," Econometrica 56 (January 1988) 37-61.

White, Halbert and Domowitz, Ian, "Nonlinear Regression with Dependent Observations," Econometrica 52 (1984) 123-142.