Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement

Robert J. Hodrick
Northwestern University

Alternative ways of conducting inference and measurement for long-horizon forecasting are explored with an application to dividend yields as predictors of stock returns. Monte Carlo analysis indicates that the Hansen and Hodrick (1980) procedure is biased at long horizons, but the alternatives perform better. These include an estimator derived under the null hypothesis as in Richardson and Smith (1991), a reformulation of the regression as in Jegadeesh (1990), and a vector autoregression (VAR) as in Campbell and Shiller (1988), Kandel and Stambaugh (1988), and Campbell (1991). The statistical properties of long-horizon statistics generated from the VAR indicate interesting patterns in expected stock returns.

In this article, I examine the statistical properties of three alternative methods for conducting inference and making measurements in long-horizon forecasting experiments with an application to dividend yields as predictors of stock returns. Recent evidence on the

I have benefited greatly from my conversations with Lars Hansen, who deserves more than the usual thank you to a colleague. Any errors are, of course, my own responsibility. I also want to thank Laurie Bagwell, Tim Bollerslev, John Cochrane, Wayne Ferson, Ken French, Michael Gibbons, Narayana Kocherlakota, Bob Konjczyk. Ravi Jagannathan, Matt Richardson. Steve LeRoy, Mark Watson, and the referee (James Stock) for useful comments and discussion. I am very grateful to Geert Bekaert for excellent research assistance. This research was partially suppotted by a grant from the Lynde and Harry Bradley Foundation. Address correspondence to Robert J. Hodrick, Finance Department, Kellogg Graduate School of Management, Northwestern University, Evanston, IL 60208-2006.

The Review of Financial Studies 1992 Volume 5, number 3, pp. 357-386

predictability of stock returns at horizons of one year or longer has generated considerable controversy, and this analysis helps to resolve some of the outstanding disagreement.[1]

There are two aspects to this debate. First, some researchers question whether there is solid evidence that expected returns vary. The poor small sample properties of some statistical methods and the low power of tests contribute to this problem. By examining the statistical properties of the three methodologies in Monte Carlo experiments, I provide evidence on the bias of the various approaches and on their power to reject the null of no expected return variability. Second, there is substantial debate about why expected returns might vary. While there are sound rational economic theories that predict movements in expected returns, some economists argue that such movements reflect irrational transitory components in stock prices. Although I cannot resolve the debate on the efficiency of the market, I examine the link between short-run and long-run predictability of returns and demonstrate that a relatively large amount of long-run predictability is consistent with only a small amount of short-run predictability.

Many forecasting studies use ordinary least squares (OLS) regression. Consequently, it is my first methodology, and I use the Fama and French (1988b) regression as my canonical example. They investigate the ability of dividend yields to predict compound returns on the value-weighted and equal-weighted NYSE portfolios for intervals between one month and four years. In the first section of the article, I reexamine the asymptotic distribution theory of the OLS estimator in long-horizon forecasting situations. I demonstrate how to formulate an alternative estimator of the standard errors that imposes the null hypothesis of no serial correlation in returns but does not impose an assumption of conditional homoskedasticity. This approach builds on Richardson and Smith (1991).

The second methodology builds on Jegadeesh (1990), who advocates a reformulation of the regression in the first methodology in order to assess statistical significance of the forecasts. If the slope coefficient in an OLS regression is different from zero, the covariance of the regressand and the regressor must be nonzero. In the first methodology, the regressand is the compound return between time $t + 1$ and time $t + k$, while the regressor is the dividend yield at time t. An alternative way to examine the statistical significance of the

Many authors have used dividend yields and other variables to examine the predictability of returns. Campbell (1991) and Cochrane (1992) attribute a large fraction of the variance of the price-dividend ratio to variation in expected returns. Nevertheless, the controversy in this literature is typified by the arguments of Jegadeesh (1990); Kim, Nelson, and Startz (1989); Mankiw, Romer, and Shapiro (1989); Nelson and Kim (1990); Richardson: (1990); and Richardson and Stock (1989). These authors argue that the case for predictability of stock returns is weak when one corrects for small sample biases in test statistics.

long-horizon forecasts is to estimate the same numerator covariance in the slope coefficient but by measuring the regressand as the return at time $t + 1$, while summing the regressor into the past from time t to time $t - k + 1$.

The third methodology recognizes that long-horizon linear predictions can be generated by iterating one-step-ahead linear predictions from a vector autoregression (VAR) as in Campbell and Shiller (1988), Kandel and Stambaugh (1988), and Campbell (1991). The VAR completely characterizes the autocovariances of the time series, and I explore how it can be used to generate implicit long-horizon statistics without actually measuring data over a long horizon.

Much of the literature on the predictability of stock returns only addresses questions of inference. Researchers ask whether a test rejects the null hypothesis of constant expected returns using the .05 or .01 critical values derived from a asymptotic distribution theory. There are two problems with this approach. First, proper inference requires knowledge of the small sample distribution of the test statistic when the null hypothesis is true. If the small sample distribution of the test statistic coincides with that predicted by asymptotic distribution theory, the asymptotic critical value provides the correct type I error. If the statistic is poorly behaved in small samples, correct inference requires that an alternative critical value be determined, possibly from a Monte Carlo experiment. There is now ample evidence that some of the techniques used in the literature do not have good small sample properties, especially if the forecast horizon is large relative to the sample size. The second point is that examination of the null hypothesis using the .05 or .01 critical values from the asymptotic distribution ignores the trade-off between type I and type II errors. Unfortunately, without a well-specified alternative hypothesis we cannot determine the power of various tests.

To address these issues, I examine each of the methodologies in Monte Carlo experiments using artificial data generated from a first-order VAR of returns, dividend yields, and the Treasury-bill rate relative to its recent average, as in Campbell (1991). Much of the article is devoted to the issue of inference under the restrictive null hypothesis that expected returns are a constant. However, since the statistical properties of the VAR approach are found to be quite good, I also perform a number of different measurements of long-horizon statistics as well. The validity of the asymptotic distributions of these test statistics is also evaluated with Monte Carlo experiments.

In Section 1, I examine the OLS methodology with long-horizon returns as the regressand. In Section 2, I examine the second methodology that reorganizes the regression to have only one-step-ahead forecasts. The VAR alternative is discussed in Section 3, where implicit

long-horizon statistics are derived. I present the estimation of the three methodologies, as well as the small sample properties of the estimators, in Section 4. In Section 5, I present the estimates and the small sample properties of the implicit long-horizon statistics from the VAR. The VAR is extended to include term premiums and default premiums in Section 6. Addition of these variables does not change any of the inference from the three variable VARs. A conclusion is provided in Section 7.

1. Forecasting Long-Horizon Returns with OLS

In this section, I examine OLS as a forecasting methodology as in the analysis of Fama and French (1988b). After describing their methodology, I examine the asymptotic distribution of the OLS estimator and demonstrate an alternative way to estimate the standard errors, one that imposes the null hypothesis that stock returns have a constant conditional mean.

Fama and French (1988b) use CRSP monthly data, which begin in January 1926. Because dividends are highly seasonal, they construct annualized dividend yields by summing the previous 12 months of dividends. Consequently, their sample begins in January 1927 and ends in December 1986 for 720 observations. Define the one-period real return as $R_{t+1} \equiv (P_{t+1} + d_{t+1})/P_t$, where P_t is end-of-month real stock price and d_t is real dividends paid during month t. Define the annualized dividend yield as D_t/P_t.

A typical OLS specification of Fama and French (1988b) is the following:

$$\ln(R_{t+k,k}) = \alpha_{k,1} + \beta_{k,1}(D_t/P_t) + u_{t+k,k}, \tag{1}$$

where $\ln(R_{t+k,k}) \equiv \ln(R_{t+1}) + \cdots + \ln(R_{t+k})$ is the continuously compounded k-period rate of return. The error term $u_{t+k,k}$ is an element of the time $t + k$ information set, and if the data are sampled more finely than the compound return interval, it is serially correlated, even under the null hypothesis, as discussed in Hansen and Hodrick (1980). If all of the monthly data are employed, $u_{t+k,k}$ is correlated with $k - 1$ previous error terms. Under alternative hypotheses in which returns have a variable conditional mean, $u_{t+k,k}$ can be arbitrarily serially correlated if dividend yields do not capture all of the variation in the conditional mean.

Since the regressor is only predetermined and not strictly exogenous, asymptotic distribution theory must be used to generate standard errors. Traditional OLS standard errors are appropriate asymptotically if there is no serial correlation of the error term and if it is

conditionally homoskedastic. The error term is serially uncorrelated when forecasting one month ahead under the null hypothesis, but the variability of conditional variances of returns, documented, for example, by French, Schwert, and Stambaugh (1987), makes an assumption of conditional homoskedasticity inappropriate. To avoid inducing serial correlation when forecasting quarterly and annual returns, Fama and French sample the data and use the traditional OLS standard errors. This gives 240 quarterly and 60 annual nonoverlapping observations. For the longer horizons of two, three, and four years, they use annual observations with overlapping data and modify the standard errors.

The asymptotic distribution of the OLS estimator of $\delta'_{k,1} = (\alpha_{k,1}, \beta_{k,1})$ can be derived from Hansen's (1982) generalized method of moments (GMM) when the data are sampled more finely than the forecasting interval and when one allows for conditional heteroskedasticity of unknown form. It can be demonstrated that $\sqrt{T}(\hat{\delta}_{k,1} - \delta_{k,1}) \sim N(0, \Omega)$,, where $\Omega = Z_0^{-1} S_0 Z_0^{-1}$, $Z_0 = E(x_t x_t')$ with $x_t' = (1, D_t/P_t)$, and S_0 is the spectral density evaluated at frequency zero of $w_{t+k} = u_{t+k,k} x_t$. Under the null hypothesis that returns are not predictable,

$$S_0 = \sum_{j=-k+1}^{k-1} E(w_{t+k} w'_{t+k-j}), \tag{2}$$

which may be estimated with

$$S_T^a = C_T(0) + \sum_{j=1}^{k-1} [C_T(j) + C_T(j)'], \tag{3}$$

where $C_T(j) = (1/T) \sum_{t=j+1}^{T} (w_{t+k} w'_{t+k-j})$ and the estimated residuals are used in w_t. The estimator of Z_0 is $Z_T = (1/T) \sum_{t=1}^{T} x_t x_t'$. The resulting standard errors for the coefficients in Equation (1) are the conditionally heteroskedastic counterparts to the standard errors of Hansen and Hodrick (1980) and are henceforth labeled standard errors (1A).

An alternative standard error for specification (1)

In this section, I develop an alternative estimator of S_0 that is valid only under the null hypothesis. The new estimator utilizes the fact that the values of unconditional expectations of stationary time series depend only on the intervals between the observations.[2]

Notice that under the null hypothesis, $u_{t+k,k} = (e_{t+1} + \cdots + e_{t+k})$, where e_{t+1} is the serially uncorrelated one-step-ahead forecast error. Estimates of e_{t+1} can be obtained from the residuals of a regression

[2] Lar Hansen suggested this estimator. which is a heteroskedastic counterpart to the covariance matrix in Richardson and Smith (1991).

of ln (R_{t+1}) on a constant. To derive the new estimator, examine a typical term in Equation (2), $E(w_{t+k}w'_{t+k-j})$, for $j > 0$, and substitute $(e_{t+1} + \cdots + e_{t+k})$ for $u_{t+k,k}$. The result is

$$E(u_{t+k,k}x_t u_{t+k-j,k}x'_{t-j}) = E\left[\left(\sum_{i=1}^{k} e_{t+i}\right)x_t\left(\sum_{b=1-j}^{k-j} e_{t+b}\right)x'_{t-j}\right]$$

$$= E\left[\left(\sum_{i=1}^{k-j} e^2_{t+i}\right)x_t x'_{t-j}\right]. \tag{4}$$

With stationary time series, the unconditional expectation on the right-hand side of Equation (4) is the sum of $k - j$ unconditional expectations, each depending only upon the distance between the terms. Hence, rather than summing e_{t+i} into the future, one can sum $x_t x'_{t-j}$ into the past:

$$E\left[\left(\sum_{i=1}^{k-j} e^2_{t+i}\right)x_t x'_{t-j}\right] = E\left[e^2_{t+1}\left(\sum_{i=0}^{k-j-1} x_{t-i}x'_{t-j-i}\right)\right]. \tag{5}$$

Applying the same logic to all of the terms in Equation (2) implies that

$$S_0 = E\left[e^2_{t+1}\left(\sum_{i=0}^{k-1} x_{t-i}\right)\left(\sum_{i=0}^{k-1} x_{t-i}\right)'\right]. \tag{6}$$

By estimating the residual series e_{t+1} and forming

$$wk_t = e_{t+1}\left(\sum_{i=0}^{k-1} x_{t-i}\right), \tag{7}$$

the alternative estimator of S_0 from Equation (6) is

$$S^b_T = \frac{1}{T}\sum_{t=k}^{T} wk_t wk'_t. \tag{8}$$

The new standard errors for the coefficients in Equation (1) are henceforth labeled standard errors (1B).

Two aspects of the estimator S^b_T are important, and both are induced by the fact that it avoids the summation of autocovariance matrices as in Equation (3). First, the estimator is positive definite since it estimates the variance of wk_t, in contrast to S^a_T which is not guaranteed to be positive definite. Second, if summation of autocovariance matrices in finite samples causes poor small sample properties of test statistics, the small sample properties of test statistics constructed with S^b_T ought to be better.

2. A Reorganization of the Long-Horizon Regression

In this section, I demonstrate how inference about the statistical significance of dividend yields as predictors of long-horizon returns can be conducted by considering the regression of one-period returns on the sum of the dividend yields.[3] This specification also avoids the summation of autocovariance matrices and may have better small sample properties under the null hypothesis than specification (1A).

Notice that because the compound k-period return is the sum of k one-period returns, the numerator of the regression coefficient $\beta_{k,1}$ in specification (1) is an estimate of

$$\text{cov}[\ln(R_{t+1}) + \cdots + \ln(R_{t+k}); (D_t/P_t)]. \tag{9}$$

This covariance is the sum of k covariances of returns and dividend yields separated by between one and k periods. With stationary time series the covariance (9) is identical to

$$\text{cov}[\ln(R_{t+1}); (D_t/P_t) + \cdots + (D_{t-k+1}/P_{t-k+1})] \tag{10}$$

which is the numerator of the slope coefficient in the following regression:

$$\ln(R_{t+1}) = \alpha_{1,k} + \beta_{1,k}[(D_t/P_t) + \cdots + (D_{t-k+1}/P_{t-k+1})] + u_{t+1}. \tag{11}$$

The first subscript on the coefficients of specifications (1) and (11) indicates how many periods ahead is the realization of the dependent variable, and the second subscript indicates how many terms are include&in the summation on the right-hand side. Under the null hypothesis, there is no serial correlation of the error term in Equation (11). Therefore, the asymptotic distribution of $\delta_{1,k} = (\alpha_{1,k}, \beta_{1,k})$ can be derived as in Section 1, but only one term ($j = 0$) is not zero in Equation (2). While both specification (1) with standard errors (1A) or (1B) and specification (11) are correct asymptotically, specifications (1B) and (11) should have better size in small samples if the null hypothesis is true because both avoid the summing of autocorrelations necessary under specification (1A).

If the null hypothesis is false, the power of the different specifications becomes important. Unfortunately, without specifying a precise alternative hypothesis, little can be said about type II errors. Fama and French (1988a, 3988b) and Poterba and Summers (1988)

[3] Jegadeesh (1990) uses similar logic and the explicit alternative hypothesis that stock prices have a first-order autocorrelated transitory component, as proposed by Fama and French (1988a) and Poterba and Summers (1988), to derive the test with the best asymptotic slope for investigating long-horizon predictability of returns using only lagged returns. He demonstrates that using the one-period return as the dependent variable and the sum of k lagged returns as the regressor is a superior way to conduct inference. The choice of k depends on the share of the variance of returns thought to be due to the transitory components in prices.

argue that an interesting alternative hypothesis is that stock prices have highly serially correlated temporary components that induce negative serial correlation in returns. Fama and French note that forecasting increasingly longer compound returns as in specification (1) allows these temporary components to manifest themselves because the variance of the compound returns grows less rapidly than if returns were serially uncorrelated. This makes forecasting ability easier to detect at long horizons if the small sample distributions of the statistics are well behaved. While estimation of specification (1) thus allows certain aspects of alternative hypotheses to arise naturally, which could improve power, standard errors (1A) and (1B) are only correct under the restrictive null hypothesis of this article because they do not allow for possible additional residual serial correlation that would be present under alternative hypotheses in which the dividend yield does not capture all of the predictable variation in stock returns.

Since returns have predictable components under plausible alternative hypotheses, there is also no reason to expect that the error term in Equation (11) will be serially uncorrelated. Because the order of the serial correlation under the alternative is unknown, I also construct an alternative covariance matrix for specification (11) by summing 12 autocovariance matrices with declining weights as in Newey and West (1987). Hence, tests of the statistical significance of dividend yields as predictors of returns can be conducted under the weaker null hypothesis that returns have predictable components.

3. A Vector-Autoregressive Alternative

A third way to conduct inference about the ability of dividend yields to predict returns at various horizons, and to measure these effects in the presence of alternative hypotheses that allow expected returns to vary, is to examine the incremental power of dividend yields with lagged returns and possibly other information present in the forecasting equation. This vector-autoregressive approach has been used by Campbell and Shiller (1988), Kandel and Stambaugh (1988), and Campbell (1991)) for example.

Inference about one-step-ahead predictability of returns is conducted by testing the forecasting-ability of the variables in the lagged information set. Measurement of long-horizon statistics relies on the fact that these statistics, such as the slope coefficient of the long-horizon return regression (1) or the variance ratios of Poterba and Summers (1988), are functions of the unconditional covariances of the data. These covariances are characterized by the parameter estimates of the VAR.

In this section, I demonstrate how to measure long-horizon statistics by estimating the parameters of the VAR and constructing the appropriate statistic that is a nonlinear function of these parameters. Standard errors are derived from the asymptotic distribution of the coefficients of the VAR.

Consider a first-order VAR in three variables: the continuously compounded real return on the CRSP value-weighted portfolio, the dividend yield, and the one-month Treasury-bill return relative to its previous 12-month moving average, which is denoted rb_t. This specification is used by Campbell (1991). Let

$$[\ln(R_t) - E(\ln(R_t)),\ D_t/P_t - E(D_t/P_t),\ rb_t - E(rb_t)]'.$$

Since Z_t follows a first-order VAR,

$$Z_{t+1} = AZ_t + u_{t+1}, \tag{12}$$

where A is a 3 × 3 matrix.[4] Although long-horizon statistics require long-horizon forecasts, the forecasting problem is simple since the error process u_{t+1} is unpredictable. If Φ_t is the information set consisting of current and past observations on Z_t, forecasts at horizon i are $E(Z_{t+i} \mid \Phi_t) = A^iZ_t$.

Since the series are covariance stationary, Equation (12) implies

$$Z_{t+1} = (I - AL)^{-1}u_{t+1} = \sum_{j=0}^{\infty} A^j u_{t+1-j}, \tag{13}$$

where the three-dimensional identity matrix is I and L is the lag operator. The unconditional variance of the Z_t process is therefore

$$C(0) = \sum_{j=0}^{\infty} A^j V A^{j\prime}, \tag{14}$$

where $V \equiv E(u_{t+1}u'_{t+1})$.[5]

To allow for compounding of returns over k periods, consider the sum of k consecutive Z_t's. The variance of this sum is

$$V_k = kC(0) + \sum_{j=1}^{k-1} (k-j)[C(j) + C(j)'], \tag{15}$$

where $C(j)$ is the j th-order autocovariance of Z_t, and $C(j) = A^jC(0)$. Then, the total variance of the sum of k returns is $e1'V_ke1$, where $e1$ is the indictor vector, $e1' = (1, 0, 0)$.

The slope coefficient in Equation (1) is the covariance of the sum

[4] Higher-order systems can be handled in exactly the same way by stacking the VAR into first-order companion form as in Campbell and Shiller (1988, 1989).

[5] In actual calculations, I truncate the infinite sum in $C(0)$ at 127.

of returns from $t + 1$ to $t + k$ and the dividend yield at t divided by the variance of the dividend yield. The alternative estimator of this slope coefficient implied by the VAR is

$$\beta(k) = \frac{e1'[C(1) + \cdots + C(k)]e2}{e2'C(0)e2}, \tag{16}$$

where $e2$ is the indicator vector, $e2' = (0, 1, 0)$.

The R^2 from this implied regression is the ratio of the explained variance of the dependent variable to its total variance. Hence,

$$R_1^2(k) = \beta(k)^2\left(\frac{e2'C(0)e2}{e1'V_k e1}\right), \tag{17}$$

where the subscript 1 is used to distinguish this from the R^2 implied by the VAR. The explanatory power of the VAR at long horizons can be assessed by examining the ratio of the explained variance of the sum of k returns to the total variance of the sum of k returns. These long-horizon R^2 coefficients can be calculated as one minus the ratio of the innovation variance in the sum of k returns to the total variance of the sum of k returns.

The innovation variance of the sum of k returns is $e1'W_k e1$, where

$$W_k = \sum_{j=1}^{k} (I - A)^{-1}(I - A^j)V(I - A^j)'(I - A)^{-1'}.$$

Hence, the implied long-horizon R^2 from the VAR is

$$R_2^2(k) = 1 - \frac{e1'W_k e1}{e1'V_k e1}. \tag{19}$$

Implied variance ratio statistics, which are parametric counterparts to the estimates in Cochrane (1988), Lo and MacKinlay (19881, and Poterba and Summers (1988), can be calculated as

$$VR(k) = \frac{e1'V_k e1}{ke1'C(0)e1}. \tag{20}$$

If returns were independently and identically distributed, the variance of the sum of k returns would be equal to k times the variance of one return, and the variance ratio in Equation (20) would be 1. If the variance ratio falls below 1, this is evidence of negative serial correlation in returns, since the variance of the sum is growing less rapidly than an i.i.d. variable.

Asymptotic distributions for the statistics

Each of the long-horizon statistics derived above is a function of the slope coefficients A and the innovation covariance matrix V of the

VAR. Let n_0 be the vector of these parameters, and let $H(n_0)$ represent the true value of one of these functions. If n_T is an estimate of the parameters from a sample of size T, the asymptotic distribution theory of GMM implies that $(\eta_T - \eta_0) \sim N(0, \Theta)$. Numerical derivatives can be used to calculate the gradient of H evaluated at n_T, which is denoted ∇H, and by a Taylor's series approximation, the asymptotic distribution of the function is

$$\sqrt{T}[H(\eta_T) - H(\eta_0)] \sim N(0, \nabla H \Theta \nabla H'). \tag{21}$$

I estimate the nine slope coefficients of the VAR with OLS and the six parameters of V with the corresponding sample moments of the OLS residuals. The asymptotic distribution of n_T is derived by recognizing that these estimates coincide with GMM estimation of a just-identified system of orthogonality conditions. The first nine orthogonality conditions are the usual OLS conditions that the residuals are orthogonal to the right-hand-side variables, $E(u_{t+1} \otimes Z_t) = 0$. The last six orthogonality conditions are given by stacking the distinct elements of $E(u_{t+1}u_{t+1}' - V) = 0$ into a vector. In constructing the GMM weighting matrix, I impose the restriction on the first nine orthogonality conditions that u_{t+1} is serially uncorrelated, but I allow a Newey-West (1987) lag of 6 for the orthogonality conditions associated with the parameters of V, since the deviations of the squared residuals from the elements of V can be arbitrarily serially correlated.

4. Inference and Measurement with the Alternative Specifications

In the previous sections, I developed three different approaches to conducting inference about the predictability of stock returns. The statistics provide measurements that characterize the serial correlation properties of the returns and their cross-correlations with other variables. In this section, I analyze the properties of these alternative estimators. In each case, I report estimates of the statistics and their asymptotic standard errors, and I provide evidence on the small sample distributions of the statistics from Monte Carlo experiments.

A data appendix (Appendix A) provides a discussion of the construction of the data series, the NYSE value-weighted real market returns, the corresponding annualized dividend yields, and the nominal Treasury-bill returns. All data are from CRSP.

4.1. Estimation results for the VAR

I first present the results of the VAR for two reasons. The test statistics allow assessment of the short-run predictability of returns, and the point estimates are used to generate artificial data for the Monte Carlo simulations. Table 1 shows results for three different sample periods:

Table 1
First-order vector autoregression of returns, dividends yields, and relative Treasury-bill rates

Dependent variable	Coefficients on regressors: Constant (SE)	$\ln(R_t)$ (SE)	D_t/P_t (SE)	rb_t (SE)	$\chi^2(3)$ conf.	R^2
A: 1927:2 to 1987:11, 730 observations						
$\ln(R_{t+1})$	−12.432	0.110	3.941	−4.738	9.867	
	(14.098)	(0.063)	(3.276)	(2.360)	.980	.020
D_{t+1}/P_{t+1}	0.164	−0.0006	0.964	0.023	1480.352	
	(0.113)	(0.0004)	(0.027)	(0.011)	.999	.938
rb_{t+1}	0.188	0.0004	−0.041	0.673	156.227	
	(0.082)	(0.0004)	(0.017)	(0.063)	.999	.456
B: 1952:1 to 1987:11, 431 observations						
$\ln(R_{t+1})$	−14.654	0.056	5.243	−8.546	22.765	
	(10.410)	(0.061)	(2.590)	(2.356)	.999	.057
D_{t+1}/P_{t+1}	0.073	−0.0002	0.981	0.039	8065.333	
	(0.041)	(0.0002)	(0.011)	(0.010)	.999	.960
rb_{t+1}	0.421	0.0007	−0.103	0.748	236.820	
	(0.201)	(0.0010)	(0.058)	(0.054)	.999	.565
C: 1927:2 to 1951:12, 299 observations						
$\ln(R_{t+1})$	−23.753	0.139	5.515	11.967	5.314	
	(27.647)	(0.091)	(5.466)	(7.906)	.850	.019
D_{t+1}/P_{t+1}	0.337	−0.001	0.937	−0.025	632.463	
	(0.217)	(0.0006)	(0.044)	(0.040)	.999	.904
rb_{t+1}	0.079	0.0003	−0.020	0.295	4.948	
	(0.154)	(0.0005)	(0.025)	(0.158)	.824	.083

The variables are the continuously compounded real stock return, 1n (R_t) the corresponding annualized dividend yield, D_t/P_t, and the one-month Treasury-bill return relative to its previous 12-month moving average. Coefficient estimates are OLS, and standard errors are heteroskedasticity consistent. The $x^2(3)$ tests the joint hypothesis that all three coefficients are zero.

sample A from February 1927 to November 1987 has 730 observations; sample B from January 1952 to November 1987 has 431 observations; and sample C from February 1927 to December 1951 has 299 observations. Sample A includes all of the available data. Sample B coincides approximately with one of Campbell's (1991) samples. In recognition that the forecasting power of the Treasury-bill rate may depend on the monetary policy regime, it allows for the change in policy induced by the Treasury-Federal Reserve Accord of 1951. Sample B also leaves out the depression years. Kim, Nelson, and Startz (1989) argue that the time-series properties of returns before World War II are quite different from those of the postwar years. Sample C contains the years prior to 1952.

If returns are not predictable, each of the three coefficients on the lagged variables in the return equation must be zero. The test statistic of this joint hypothesis has a x^2-distribution with three degrees of freedom. For Sample B its value is 22.765 with a confidence level larger than .999. The evidence for return predictability is less strong when the full sample is employed, since the confidence level falls to

.980. This decrease occurs because the confidence level for the test statistic from sample C is only .850.

There is very strong evidence that the Treasury-bill return has predictive power in sample B, since the confidence level of its test statistic is larger than .999. The evidence that dividend yields predict returns is slightly less strong since the confidence level of its test statistic is 980. These confidence levels are from asymptotic distributions, and reliable inference requires that the small sample properties of the estimators coincide with the asymptotic distributions. In the next section, I examine this issue.

4.2 Small sample considerations

Monte Carlo experiments require a data-generating process that provides artificial stock returns, dividend yields, and Treasury-bill returns whose time-series properties are consistent with those of the actual data. I follow Campbell and Shiller (1989) and generate artificial data from simulations of the VAR.

The VAR can be used to generate data that satisfy either the null hypothesis of no return predictability or an alternative hypothesis. When generating data under the null hypothesis, I simply set the coefficients on the lagged variables in the return equation equal to zero, and I set the constant in the return equation equal to the unconditional mean of returns implied by the original VAR. When generating data under the alternative, I set the coefficients at their point estimates from Sample B since it has the strongest evidence against the null.

Because the actual data are conditionally heteroskedastic, I estimate a generalized autoregressive conditionally heteroskedastic (GARCH) model of the conditional covariance matrix of the residuals of the VAR to generate realistic data.[6] Rather than consider several methods for different time periods, I work only with sample B to estimate the parameters. Estimation of the GARCH model and a complete description of the data-generating process are in Appendix B.

One additional aspect of the VAR that may affect inference is the assumed stationarity of the regressors. From Table 1 it is clear that dividend yields are highly persistent, which might negate the validity of the usual asymptotic distribution theory used to generate standard errors and test statistics. In order to determine the severity of this issue, I also estimated the VAR subject to the constraints that returns were not predictable, that dividend yields contain a unit root, and that dividend yields do not predict the Treasury-bill return relative to its past average. This latter constraint makes sense because the

[6] The GARCH estimation was done with FORTRAN programs written by Tim Bollerslev. who also helped me with the estimation. I am very grateful to him for his advice and assistance.

Table 2
Small sample properties of the VAR test statistics

A: Quantiles of the $\chi^2(3)$ test statistic under the null

Quantile	1%	2.5%	5%	50%	95%	97.5%	99%	Mean	SD
$\chi^2(3)$	0.115	0.216	0.352	2.366	7.815	9.348	11.34	3.000	2.449
Emp. 1	0.186	0.283	0.417	2.404	8.160	9.691	11.68	3.088	2.548
Emp. 2	0.241	0.416	0.626	3.646	10.43	12.17	15.18	4.385	3.215

B: Percent of observations greater than nominal critical values under the null hypothesis

	Test 1 Nominal size			Test 2 Nominal size			Test 3 Nominal size			Test 4 Nominal size		
	.100	.050	.010	.100	.050	.010	.100	.050	.010	.100	.050	.010
Emp. 1	.103	.050	.008	.100	.051	.011	.124	.067	.011	.108	.058	.012
Emp. 2	.097	.049	.005	.314	.200	.067	.119	.065	.012	.235	.134	.038

C: Simulated type II error rates for tests of 5% size and new critical values

Test 1			Test 2			Test 3			Test 4		
Err. rate	Crit. val. 1	Crit. val. 2	Err. rate	Crit. val. 1	Crit. val. 2	Err. rate	Crit. val. 1	Crit. val. 2	Err. rate	Crit. val. 1	Crit. val. 2
.800	3.838	3.820	.481	3.842	7.815	.126	4.427	4.336	.025	8.160	10.43

The results are for 2000 Monte Carlo experiments with 421 observations. Tests 1, 2, and 3 are Wald Tests of the respective null hypotheses that the lagged return, the dividend yield, or the Treasury bill variable does not forecast returns. Test 4 is the joint test that the three variables do not forecast returns. Panel A reports the quantiles of a $x^2(3)$ and those of two empirical distributions of Test 4 under the null hypothesis of no expected return variability. Emp. 1 is a stationary VAR, and Emp. 2 imposes a unit root in the dividend yield. Each entry in Panel B describes the fraction of the experiments under the null in which the value of the test statistic is larger than the critical value from a x^2-distribution corresponding to the nominal sizes of .10, .05, or .01 with either one degree of freedom for Tests 1.2, and 3 or three degrees of freedom for Test 4. Panel C shows the new .05 critical values of the empirical distributions and the fractions of the experiments that fail to exceed this value for Emp. 1. when the data are generated under the alternative hypothesis that returns are serially correlated as in the conditionally heteroskedastic. stationary VAR. The nominal critical value of a $x^2(1)$ is 3.841 and of a $x^2(3)$ is 7.815.

Treasury-bill variable is stationary and would inherit the assumed unit root in the dividend yield if predictability were allowed.

The small sample properties of the VAR tests of the null hypothesis of constant expected returns are presented in Table 2. Each experiment has 431 observations as in sample B, and 2000 experiments are conducted. With Tests 1,2, and 3, I examine the null hypotheses that returns are not predicted by lagged returns, lagged dividend yields, and lagged Treasury-bill returns, respectively. With Test 4, I examine the joint hypothesis that all three variables do not predict returns. The empirical distributions of test statistics under the stationary VAR are referred to as Emp. 1 and the empirical distributions of the test statistics for the nonstationary VAR are referred to as Emp. 2.

The small sample properties of the four tests for the stationary VAR are very good. The quantiles of the empirical distribution of Test 4 in Panel A are quite close to those of the $x^2(3)$. The empirical type

I error rates of the four tests are presented in Panel B. These are the percents of the 2000 experiments conducted under the null hypothesis in which the values of the test statistics are greater than the nominal .10, .05, and .01 critical values. The four tests seem quite reliable. For example, only 5.8 percent of the observations are greater than the .05 nominal critical value.[7] Inclusion of a unit root in the dividend yield does cause a noticeable deterioration in the performance of Tests 2 and 4. Most importantly, the new .05 critical value for Test 4 rises from 8.16 to 10.43, and the new .01 critical value for Test 4 rises from 11.68 to 15.18. Since these are still substantially below the sample statistic of 22.77, there is no reason to reassess the asymptotic inference discussed in connection with Table 1 above. Consequently, I only investigate the small sample properties of the statistics under the stationarity assumption.[8]

4.3 Results under the alternative hypothesis for VAR tests

In Panel C in Table 2, I examine the power of the VAR tests using the .05 critical values of the empirical distributions from the stationary VAR, when the alternative hypothesis is that returns are serially correlated as in the conditionally heteroskedastic VAR. Consistent with the findings of Poterba and Summers (1988), Test 1 has very low power since the type II error rate of a test with .05 size is 80 percent. This finding reflects the fact that returns generally have a large innovation variance, which makes it difficult to detect serial correlation in samples of this size. Tests 2 and 3 have better power since the corresponding type II error rates are 48.1 percent and 12.6 percent. Test 4, the joint test, is very powerful, with a type II error rate of 2.5 percent. Before examining the empirical distributions of the long-horizon statistics implied by the VAR, I consider the properties of the other approaches to inference and measurement developed above.

4.4 A comparison of specifications (1) and (11)

The results from estimation of specifications (1) and (11) for five horizons between one month and four years are presented in Table 3. In both cases, two standard errors are reported. Specification (1) has standard errors (1A) and (1B), and specification (11) has standard errors constructed either with no Newey-West lags or 12 lags. The

[7] If p is the nominal size and N is the number of experiments, the large sample standard error for the respective significance levels is $[p(1-p)/N]^{.5}$ which is .0067 for the .10 level, .0049 for the .05 level, and .0022 for the .01 level with 2000 experiments. Hence, although the percentages of the distributions of the test statistics that are greater than the nominal critical values are quite close to the nominal sizes, some of the estimated percentages are slightly more than two standard deviations from their nominal levels.

[8] The deterioration of some of the test statistics under the unit root assumption causes me to think that the empirical distributions of other statistics reported in this article would probably also be affected. A full exploration of this issue is beyond the scope of this project.

Table 3
Comparison of overlapping regressors, specification (1). and summed regressors, specification (11)

	Specification (1)							
Lead	$\beta_{k,1}$	$\beta^{\alpha}_{k,1}$	$z(A)$	95%	$z(B)$	95%	R^2	95%
1	3.390	2.381	1.023	1.966	1.005	1.935	.005	.006
12	4.501	3.651	2.395	2.275	1.607	1.991	.085	.066
24	4.337	3.597	4.960	2.677	1.962	1.954	.168	.111
36	3.844	3.165	4.723	3.247	2.276	1.938	.235	.139
48	3.616	2.967	4.605	3.825	2.937	1.917	.354	.154
$\chi^2(5)$					26.696	18.015		

	Specification (11)							
Lag	$\beta_{1,k}$	$\beta^{\alpha}_{1,k}$	$z(0)$	95%	$z(12)$	95%	R^2	95%
1	3.390	2.381	1.023	1.966	1.221	2.076	.005	.006
12	5.429	4.365	1.697	2.003	2.034	2.097	.010	.006
24	6.246	5.085	2.069	1.959	2.507	2.104	.011	.006
36	6.355	5.037	2.335	1.981	2.611	2.074	.010	.006
48	6.755	5.235	2.754	1.963	2.707	2.093	.011	.006
$\chi^2(5)$			10.545	11.497	8.326	13.143		

The $b_{k,1}$ are OLS estimates of Equation (1) with the dependent variable multiplied by $(1/k)$; the $b_{1,k}$ are OLS estimates of Equation (11) with the regressor multiplied by $(1/k)$. The adjusted coefficients $\beta^{\alpha}_{1,k}$ and $\beta^{\alpha}_{k,1}$ subtract the means of the Monte Carlo distributions from the OLS estimates. *Z*- statistics are unadjusted estimates divided by estimated asymptotic standard errors. The columns labeled 95% provide the 95th percentile of the empirical distributions from the Monte Carlo experiments conducted under the null hypothesis for the respective asymptotic statistics in the adjacent left column. The sample period for specification (1) depends upon the lead of the compound return. The first sample is January 1929 to December 1987 for 708 observations. Each higher compound return loses one observation. For specification (11), the sample period is January 1929 to December 1987 for 708 observations for k equal to 1 through 24. Twelve and 24 observations are lost from the beginning for k equal to 36 and 48. The $x^2(5)$ statistics test the joint hypothesis that all five slope coefficients are zero. The value for specification (1A) could nor be computed.

ratio of an estimated coefficient to its asymptotic standard error is reported as a z -statistic, which is asymptotically distributed as a standard normal under the assumption that the specification of the model is correct.

The basic sample period for forecasting one month ahead is from January 1929 to December 1987 for 708 observations since specifications (1A) and (11) with no lags are the same.[9] Then, for specification (1), one observation is lost for each higher-order compound return, with the result that the four-year compound return equation has 661 observations. For specification (11), the sample period is constant until 36 or 48 lags are required in the sum of the dividend yields. Then, 12 and 24 observations are lost from the beginning of the sample, which allows 696 or 684 observations with 36 or 48 lags in the regressor.

To facilitate interpretation of the slope coefficients, the compound return in specification (1) at horizon k is multiplied by $(1/k)$ and for

[9] The sample differs slightly from sample A to allow for lags of the predetermined variable.

specification (11) the sum of k dividend yields is multiplied by $(1/k)$. The slope coefficients in specification (1) consequently measure the response of an annualized expected return over a given horizon to a change in the current dividend yield, while those in specification (11) measure the change in the annualized one-month return with a change in the average dividend yield. A coefficient of 3.6, for example, indicates that a 100 basis point increase in the dividend yield implies a 360 basis point increase in the expected annualized return.

Because the regressors in specifications (1) and (11) are only predetermined and not exogenous, estimates of the slope coefficients have small sample biases, as Stambaugh (1986) and Mankiw and Shapiro (1986) demonstrate. Consequently, I report adjusted estimates, β^{α}'s which are obtained by subtracting the means of the slope coefficients of the Monte Carlo distributions from the OLS estimates. Table 3 also reports the R^2 for each equation.

For each statistic, I provide the critical value associated with the 95th percentile of the empirical distribution of the test statistic from the Monte Carlo experiments conducted under the null hypothesis. The 2000 simulations were conducted exactly as in the actual estimation. Since the construction of standard errors (1A) does not guarantee a positive definite covariance matrix, 23 experiments were discarded when this occurred. The problems arose primarily in summing 48 lags.

Comparing a z-statistic to the 95th percentile of its empirical distribution provides a one-sided test of the null hypothesis that the slope coefficient is zero versus the alternative hypothesis that the coefficient is positive. A one-sided test is appropriate because both rational and irrational theories of time varying expected returns predict that high dividend yields forecast high expected returns. To interpret the results, recall that the critical value of the 95th percentile of a standard normal is 1.645, and compare this to the critical values of the empirical distributions. In all cases, the z-statistics are positively biased. Notice, though, that the results in Table 3 support the conjecture that the sizes of the test statistics are closer to the desired nominal sizes for specifications (1B) and (11) than for specification (1A).

For example, for specification (1A), the .05 critical values of the empirical distributions increase from 1.966 at the one-month horizon to 3.825 at the 48-month horizon, while the new critical values for specification (1B) fall slightly from 1.935 to 1.917. The primary source of bias for specification (1A) is the summing of the covariance matrices.[10]

[10] Richardson and Stock (1989) explore an alternative asymptotic distribution theory in which the ratio of the forecasting interval k to the sample size T limits to a nonzero constant as T goes to

There are two sources of potential bias in specification (11): the use of lags in the Newey-West (1987) standard errors when there is no residual serial correlation and the additional serial correlation in the regressor induced by summing the lagged dependent variable. The test statistics are slightly more biased, using 12 lags in the Newey-Westtechnique; however, since the regressor is already highly serially correlated, summing the regressor does not cause a pronounced deterioration in the test statistics as it might if the regressor were not initially very serially correlated.

Do the results of Table 3 indicate that dividend yields predict stock returns? The overall picture suggests the answer is yes. Although the results at the one-month horizon do not provide strong evidence against the null hypothesis, the overall evidence appears strong. At the annual horizon, for specification (1A) the z -statistic of 2.395 is above the empirical critical value of 2.275. Similarly, the test statistic for specification (11) with 12 lags is 2.034, compared to the empirical critical value of 2.097. The values of the R^2 's for the annual and longer horizons are also greater than the 95th percentiles of the empirical distributions. At the two-, three-, and four-year horizons, the differences in inference are less pronounced across the different specifications. In all cases, the point estimates of the test statistics and the R^2 's are well above the respective critical values of the empirical distributions.

Richardson (1990) argues correctly that interpretation of the above analysis must take account of the correlation of the different test statistics, which requires simultaneous estimation of the five forecasting equations. The test statistic of the joint hypothesis that the five slope coefficients are simultaneously zero has a x^2-distribution with five degrees of freedom. For specification (1B), its value is 26.696, which substantially exceeds the .05 critical value of the empirical distribution of the test statistic, 18.015. Since the nominal .05 critical value for a $x^2(5)$ is 11.071, there is a substantial bias in the joint test statistic. Simultaneous estimation of the five equations for specification (1A) results in failure of the GMM weighting matrix to be positive definite. For specification (11), the values of the test statistics are 10.545 (0 lags), and 8.326 (12 lags), but there is less bias in these joint tests as the .05 empirical values are 11.497 and 13.143, respectively. Although these latter findings support Richardson's (1990) conclusion that evidence for long-horizon predictability of returns is

infinity. In the context of the Fama and French (1988a) analysis using only return data, Richardson and Stock demonstrate that the small sample distributions are closer to this alternative asymptotic theory than to the traditional one. When data other than returns are present, derivation of the alternative asymptotic distribution depends upon nuisance parameters that characterize the serial correlation properties of the other series. I thank Lars Hansen for this insight.

not as strong as the individual test statistics indicate, the overall picture still appears to be one of return predictability.

4.5 The power of specifications (1) and (11)

I next examine how the two specifications perform when the null hypothesis is false. The alternative hypothesis allows returns to be serially correlated as in the data from the conditionally heteroskedastic VAR for sample B, and I again examined 2000 experiments. The probability of a type II error is calculated as the percent of the observations in which the test statistics are not larger than the .05 critical values associated with the empirical distributions calibrated under the null hypothesis. After correcting for small sample biases, the type II error rates are remarkably similar. For specifications (1A) and (1B), the type II error rates are, respectively, 5.1 percent and 5.0 percent when k = 1, *2.6* percent and 1.4 percent when k + = 12, 3.9 percent and 3.1 percent when k = 24, 9.2 percent and 8.4 percent when k = 36, 14.7 percent and 13.9 percent when k = 48, and 2.2 percent for the $x^2(5)$ for specification (1B). For specification (11) with no lags or 12 lags, the corresponding type II error rates are 5.1 percent and 12.3 percent when k = 1, 1.3 percent and 2.1 percent when k = 12, 2.9 percent and 3.9 percent when k = 24, 7.7 percent and 10.2 percent when k = 36, 14.2 percent and 18.4 percent when k = 48, and 14.7 percent and 19.8 percent for the $x^2(5)$. Hence, although specification (1A) and specification (11) with additional Newey-West lags allow aspects of the alternative hypothesis to manifest themselves in the estimation, the approaches are not more powerful than the approaches that impose the null hypothesis because the statistics are more biased under the null.

In terms of methodology, the message from this section is clear. If conducting inference without inducing a serially correlated dependent variable is possible, such an alternative procedure is preferred since its small sample properties under the null hypothesis are closer to the standard asymptotic distribution. However, even in this case, the potential for bias appears strong, and Monte Carlo analysis is appropriate.

5. Implied Long-Horizon Statistics

I next examine estimates of the implied long-horizon statistics, derived in Section 3. These are highly nonlinear in the underlying parameters of the VAR. I compare their asymptotic distributions to the empirical distributions under the null and alternative hypotheses. Given the good small sample properties of the coefficient estimates and basic

Table 4
Implied long-horizon statistics and asymptotic standard errors from the first-order VAR of returns, dividend yields, and relative Treasury-bill rates

A: Implied slope coefficients, long-horizon returns on dividend yields

Sample	$\beta(1)$ (SE)	$\beta(12)$ (SE)	$\beta(24)$ (SE)	$\beta(36)$ (SE)	$\beta(48)$ (SE)	$\beta(\infty)$ (SE)
A	3.375	47.448	77.320	95.058	105.587	120.966
	(2.575)	(28.879)	(37.226)	(36.776)	(33.997)	(28.253)
B	6.021	85.763	148.215	187.649	212.375	253.375
	(2.398)	(28.117)	(38.284)	(37.358)	(32.942)	(30.393)
C	3.429	48.667	70.835	80.279	84.302	87.287
	(3.869)	(40.252)	(46.868)	(45.210)	(42.943)	(39.583)

B: Implied $R_1^2(k)$ coefficients, long-horizon returns on dividend yields

Sample	$R_1^2(1)$ (SE)	$R_1^2(12)$ (SE)	$R_1^2(24)$ (SE)	$R_1^2(36)$ (SE)	$R_1^2(48)$ (SE)	$R_1^2(60)$ (SE)
A	.005	.070	.107	.122	.127	.125
	(.005)	(.056)	(.076)	(.082)	(.085)	(.087)
B	.011	.160	.272	.337	.373	.389
	(.005)	(.054)	(.076)	(.081)	(.082)	(.085)
C	.004	.061	.077	.076	.070	.063
	(.007)	(.077)	(.088)	(.083)	(.075)	(.069)

C: Implied $R_2^2(k)$ coefficients, long-horizon returns on all three variables

Sample	$R_2^2(1)$ (SE)	$R_2^2(12)$ (SE)	$R_2^2(24)$ (SE)	$R_2^2(36)$ (SE)	$R_2^2(48)$ (SE)	$R^2(60)$ (SE)
A	.024	.074	.108	.123	.127	.125
	(.016)	(.056)	(.075)	(.081)	(.084)	(.086)
B	.062	.187	.278	.339	.373	.389
	(.028)	(.056)	(.075)	(.079)	(.081)	(.085)
C	.029	.063	.078	.077	.070	.063
	(.029)	(.077)	(.088)	(.083)	(.075)	(.069)

D: Implied variance ratios

Sample	VR(12) (SE)	VR(24) (SE)	VR(36) (SE)	VR(48) (SE)	VR(60) (SE)
A	1.085	0.904	0.826	0.739	0.673
	(.146)	(.191)	(.213)	(.218)	(.213)
B	1.168	1.025	0.884	0.769	0.678
	(.193)	(.215)	(.216)	(.210)	(.200)
C	1.026	0.855	0.746	0.674	0.625
	(.181)	(.235)	(.258)	(.264)	(.263)

Sample A: 1927:2 to 1987:11, 730 observations; sample B: 1952:1 to 1987:11, 431 observations; sample C: 1927:2 to 1951:12, 299 observations. The implied slope coefficient $\beta\ (k)$ in Panel A is derived in Equation (16), the $R^2_1\ (k)$ in Panel B is derived in Equation (17), the $R^2_2\ (k)$ in Panel C is derived in Equation (19). and the variance ratio VR (k) in Panel D is derived in Equation (20).

test statistics of the VAR, the question is whether the nonlinearities in estimating the implied long-horizon statistics induce biases.

In Table 4, I report estimates of the implied long-horizon statistics (with their associated asymptotic standard errors in parentheses) for the three sample periods. Estimates of $\beta\ (k)$, the implied slope coefficient in the regression of the sum of k future returns on the current

dividend yield, derived in Equation (16), for k equal to 1, 12, 24, 36, and 48 months ahead, and for the infinite horizon forecast are contained in Panel A. The $R^2_1(k)$ in Equation (17) associated with this implied regression is reported and the 60-month result is substituted for the infinite horizon in Panel B. Estimates of the $R^2_2(k)$ derived in Equation (19) are contained in Panel C, and the implied variance ratios of Equation (20) are reported in Panel D. The empirical distributions of the four statistics are presented in the corresponding panels of Table 5 under the null hypothesis of no return predictability and in Table 6 under the alternative hypothesis.

First, consider the point estimates relative to their asymptotic standard errors. Since only in sample B is there strong evidence of predictability of returns in the linear VAR, it is not surprising that only for this sample are the standard errors of the long-horizon statistics small relative to their point estimates. To conserve space, I therefore focus only on the relation of the estimated results from sample B to the empirical distributions in Tables 5 and 6, which are generated from the sample B point estimates.

The $\beta(k)$ coefficients in Panel A are not divided by the horizon as in Table 1 to allow computation of the infinite horizon forecast. When they are divided by the horizon, the results are 7.147 (k = 12), 6.176 (k = 24), 5.212 (k = 36), and 4.424 (k = 48). These coefficients indicate that an increase in the dividend yield of 1 percent implies a 7 percent per annum increase in the expected return on stocks over the next year and a 4 percent per annum increase over the next four years.

The measures of the predictive power of dividend yields and the full VAR for long-horizon returns as estimated by the two R^2's reported in Panels B and C are essentially the same. Although only 6 percent of the return is predictable over the next month, the dynamics of the VAR imply that the ratio of the explained variance of the compound return to its total variance rises to 19 percent at 12 months, 28 percent at 24 months, 34 percent at 36 months, and 39 percent at 48 months. The variance ratios in Panel D first rise above 1 before falling below 1. This indicates that serial correlation in returns is initially positive, then negative.

Now examine the empirical distributions of the implied statistics. The results for the first three sets of statistics are quite similar. Compare the implied $\beta(k)$, the implied $R^2_1(k)$, and the implied $R^2_2(k)$ to the empirical distributions in the corresponding panels of Tables 5 and 6. In all three cases, the point estimates from the data are larger than the 99th percentile of the empirical distributions calculated under the null hypothesis in Table 5, except when k = 1, in which case they are larger than the 95th percentile. The estimates are also

Table 5
Quantiles of the empirical distributions of the implied long-horizon statistics. under the null hypothesis

A: Implied slope coefficients, long-horizon returns on dividend yields

Quantile	$\beta(1)$	$\beta(12)$	$\beta(24)$	$\beta(36)$	$\beta(48)$	$\beta(\infty)$
1%	−4.475	−50.004	−89.845	−121.137	−147.026	−250.658
5%	−3.293	−36.052	−63.082	−83.657	−98.373	−148.230
10%	−2.590	−29.005	−50.068	−65.159	−75.913	−104.830
50%	0.418	5.037	8.300	10.375	11.575	13.548
90%	4.928	45.659	67.779	78.036	83.343	89.871
95%	6.470	60.162	86.998	99.209	105.438	112.218
99%	9.695	85.410	117.241	132.616	138.253	144.089
Mean	0.859	7.281	9.101	8.658	7.529	0.589
SD	3.082	29.533	45.918	55.724	62.615	81.417

B: Implied $R_1^2(k)$ coefficients, long-horizon returns on dividend yields

Quantile	$R_1^2(1)$	$R_1^2(12)$	$R_1^2(24)$	$R_1^2(36)$	$R_1^2(48)$	$R_1^2(60)$
1%	.000001	.000004	.000004	.000006	.000007	.000007
5%	.00001	.0001	.0002	.0002	.0002	.00001
10%	.00004	.0004	.0006	.0006	.0006	.0005
50%	.001	.009	.013	.013	.012	.013
90%	.006	.053	.072	.070	.067	.061
95%	.009	.075	.097	.100	.093	.087
99%	.014	.119	.147	.157	.153	.145
Mean	.002	.020	.026	.026	.025	.023
SD	.003	.027	.034	.035	.033	.031

C: Implied $R_2^2(k)$ coefficients, long-horizon returns on all three variables

Quantile	$R_2^2(1)$	$R_2^2(12)$	$R_2^2(24)$	$R_2^2(36)$	$R_2^2(48)$	$R_2^2(60)$
1%	.0004	.0003	.0002	.0002	.0001	.000007
5%	.0010	.0011	.0009	.0007	.0006	.0005
10%	.0014	.0021	.0018	.0016	.0013	.0011
50%	.006	.014	.015	.014	.013	.012
90%	.015	.059	.075	.072	.068	.064
95%	.019	.081	.101	.103	.096	.088
99%	.026	.129	.152	.160	.155	.148
Mean	.007	.024	.028	.028	.026	.024
SD	.006	.028	.035	.036	.034	.031

D: Implied variance ratios

Quantile	VR(12)	VR(24)	VR(36)	VR(48)	VR(60)
1%	0.754	0.640	0.561	0.508	0.473
5%	0.810	0.709	0.638	0.591	0.564
10%	0.844	0.757	0.702	0.658	0.629
50%	0.994	0.982	0.972	0.961	0.959
90%	1.169	1.273	1.365	1.444	1.510
95%	1.224	1.375	1.501	1.615	1.719
99%	1.338	1.548	1.748	1.937	2.088
Mean	1.003	1.004	1.009	1.016	1.024
SD	0.127	0.202	0.265	0.317	0.359

Each experiment has 431 observations, and 2000 experiments were conducted. The row entries are the values of the test statistics associated with the quantiles of the empirical distributions. The last row reports the sample standard deviation of the empirical distribution. See also notes for Table 4.

Table 6
Quantiles of the empirical distributions of the implied long-horizon statistics, under the alternative hypothesis

A: Implied slope coefficients, long-horizon returns on dividend yields

Quantile	$\beta(1)$	$\beta(12)$	$\beta(24)$	$\beta(36)$	$\beta(48)$	$\beta(\infty)$
1%	2.429	38.978	74.340	99.056	118.357	177.628
5%	3.450	52.068	95.979	128.505	152.206	204.798
10%	4.019	60.571	108.679	143.354	168.384	215.057
50%	6.892	92.417	155.490	192.494	214.478	250.177
90%	11.132	132.267	201.892	234.928	251.323	288.660
95%	12.420	146.333	216.167	246.091	260.843	299.948
99%	15.212	169.995	240.729	265.188	281.410	330.405
Mean	7.290	94.653	155.311	190.399	211.404	251.053
SD	2.825	28.569	36.521	35.776	33.292	30.622

B: Implied $R_1^2(k)$ coefficients, long-horizon returns on dividend yields

Quantile	$R_1^2(1)$	$R_1^2(12)$	$R_1^2(24)$	$R_1^2(36)$	$R_1^2(48)$	$R_1^2(60)$
1%	.003	.053	.097	.123	.136	.145
5%	.005	.083	.150	.191	.215	.226
10%	.007	.102	.177	.225	.248	.258
50%	.013	.173	.283	.340	.367	.375
90%	.023	.266	.396	.451	.475	.485
95%	.027	.291	.425	.480	.510	.522
99%	.034	.345	.481	.534	.568	.588
Mean	.014	.179	.285	.339	.364	.373
SD	.007	.064	.085	.089	.091	.091

C: Implied $R_2^2(k)$ coefficients, long-horizon returns on all three variables

Quantile	$R_2^2(1)$	$R_2^2(12)$	$R_2^2(24)$	$R_2^2(36)$	$R_2^2(48)$	$R_2^2(60)$
1%	.016	.068	.101	.126	.138	.146
5%	.024	.104	.157	.193	.215	.227
10%	.029	.123	.183	.227	.249	.259
50%	.055	.196	.288	.341	.367	.375
90%	.100	.284	.398	.452	.475	.485
95%	.125	.312	.427	.481	.510	.522
99%	.210	.377	.483	.536	.568	.587
Mean	.063	.201	.290	.340	.367	.364
SD	.042	.065	.084	.089	.090	.091

D: Implied variance ratios

Quantile	VR(12)	VR(24)	VR(36)	VR(48)	VR(60)
1%	0.829	0.629	0.498	0.408	0.357
5%	0.897	0.715	0.575	0.480	0.416
10%	0.949	0.770	0.627	0.526	0.456
50%	1.100	0.949	0.820	0.710	0.625
90%	1.340	1.191	1.051	0.937	0.850
95%	1.441	1.308	1.156	1.029	0.931
99%	1.777	1.654	1.438	1.307	1.191
Mean	1.133	0.977	0.837	0.729	0.645
SD	0.203	0.212	0.198	0.184	0.171

Each experiment has 431 observations, and 2000 experiments were conducted. The row entries are the values of the test statistics associated with the quantiles of the empirical distribution. The last row reports the sample standard deviation of the empirical distribution. See also notes for Table 4.

Table 7
First-order vector autoregression of returns, dividend yields, relative Treasury-bill rates, term premiums, and default premiums

Dependent variable	Coefficients on regressors: $\ln(R_t)$ (SE)	D_t/P_t (SE)	rb_t (SE)	ytp_t (SE)	ydp_t (SE)	$\chi^2(5)$ conf.	R^2
A: 1927:2 to 1987:11, 730 observations							
$\ln(R_{t+1})$	0.109	3.880	−4.582	0.404	0.012	11.426	
	(0.062)	(2.434)	(2.729)	(2.403)	(6.526)	.956	.017
D_{t+1}/P_{t+1}	−0.001	0.972	0.022	0.007	−0.031	5183.405	
	(0.001)	(0.017)	(0.012)	(0.010)	(0.043)	.999	.938
rb_{t+1}	0.0003	−0.021	0.676	0.034	−0.083	220.871	
	(0.001)	(0.015)	(0.071)	(0.050)	(0.042)	.999	.456
ytp_{t+1}	−0.0001	0.008	−0.504	0.639	0.109	1447.686	
	(0.0003)	(0.013)	(0.049)	(0.041)	(0.031)	.999	.803
ydp_{t+1}	−0.001	0.008	−0.007	−0.005	0.969	4036.008	
	(0.0002)	(0.007)	(0.009)	(0.005)	(0.024)	.999	.960
B: 1952:1 to 1987:11, 431 observations							
$\ln(R_{t+1})$	0.058	6.180	−8.768	0.518	−4.022	24.786	
	(0.061)	(2.826)	(3.176)	(2.947)	(5.889)	.999	.054
D_{t+1}/P_{t+1}	−0.0002	0.972	0.043	0.009	0.006	8548.105	
	(0.0002)	(0.012)	(0.012)	(0.011)	(0.024)	.999	.960
rb_{t+1}	0.0007	−0.060	0.750	0.048	−0.174	311.224	
	(0.001)	(0.044)	(0.065)	(0.047)	(0.133)	.999	.568
ytp_{t+1}	−0.0005	−0.015	−0.503	0.607	−0.086	757.208	
	(0.001)	(0.028)	(0.038)	(0.037)	(0.086)	.999	.951
ydp_{t+1}	−0.0002	0.024	0.006	−0.006	0.963	4330.088	
	(0.0001)	(0.006)	(0.009)	(0.006)	(0.018)	.999	.951
C: 1927:2 to 1951:12, 299 observations							
$\ln(R_{t+1})$	0.138	4.697	12.431	−0.872	3.329	5.822	
	(0.088)	(4.355)	(7.981)	(5.153)	(8.826)	.676	.013
D_{t+1}/P_{t+1}	−0.001	0.950	−0.032	0.001	−0.001	1910.541	
	(0.001)	(0.032)	(0.038)	(0.023)	(0.058)	.999	.904
rb_{t+1}	0.0003	−0.002	0.285	0.052	−0.086	13.312	
	(0.001)	(0.025)	(0.163)	(0.097)	(0.061)	.979	.088
ytp_{t+1}	0.0002	−0.001	−0.724	0.685	0.110	1001.785	
	(0.0003)	(0.018)	(0.113)	(0.092)	(0.053)	.999	.783
ydp_{t+1}	−0.001	−0.006	−0.039	−0.002	0.967	4426.364	
	(0.0003)	(0.013)	(0.022)	(0.013)	(0.032)	.999	.958

Equation	Test	Exclusion tests: Sample A	Sample B	Sample C
$\ln(R_{t+1})$	$\chi^2(2)$	0.030	0.556	0.143
	conf.	.015	.243	.069
D_{t+1}/P_{t+1}	$\chi^2(2)$	0.744	0.606	1.008
	conf.	.311	.261	.396
rb_{t+1}	$\chi^2(2)$	4.324	2.301	3.521
	conf.	.885	.684	.828
ytp_{t+1}	$\chi^2(3)$	107.462	185.729	41.401
	conf.	.999	.999	.999
ydp_{t+1}	$\chi^2(3)$	22.117	21.256	28.392
	conf.	.999	.999	.999

See Table 1 for definitions of $\ln(R_t)$, D_t/P_t, and rb_t. The term premium, ytp_t, is the difference between the yield on long-term government bonds and the one-month Treasury bills. The default premium, ydp_t, is the yield differential between BAA and AAA corporate bonds. Exclusion tests examine the restrictions that ydp_t and ydp_t do not forecast $\ln(R_{t+1})$, D_{t+1}/P_{t+1}, and rb_{t+1}, and that the latter three variables at time t do not forecast the former two at time $t + 1$.

very close to the means of the empirical distributions calculated under the alternative hypothesis in Table 6. The sample standard deviations of the empirical distributions calculated under the alternative hypothesis are also very close to the asymptotic standard errors reported in Table 4. The conclusion is that the asymptotic distributions of these long-horizon statistics accord very well with the distributions calculated under the alternative hypothesis demonstrating that the nonlinearities do not induce bad small sample biases.

The point estimates of the implied variance ratios are not as far into the tails of the empirical distributions calculated under the null hypothesis, but they are relatively close to the means of the distributions calculated under the alternative hypothesis. These results suggest that implied variance ratios are not a powerful way of testing the null hypothesis.

6. Adding Term Premiums and Default Premiums to the VAR

Fama and French (1989) conduct additional forecasting analyses similar to their earlier paper for several portfolios of stock and bond returns. As in Keim and Stambaugh (1986), they use the slope of the yield curve and the default premium on low-grade bonds relative to high-grade bonds as well as dividend yields. In Table 7, I report the results of adding a term premium and a default premium to the VAR of this article. The term premium ytp_t is the difference between the yield on long-term government bonds and the one-month Treasury-bill rate, and the default premium ydp_t is the difference between the yield on BAA corporate bonds and AAA corporate bonds.[11]

The exclusion tests in Table 7 very strongly indicate that the two premiums provide no additional explanatory power for the market return, the dividend yield, and the relative Treasury-bill rate compared to forecasts made with lags of these variables. The tests also very strongly indicate that the term premium and the default premium can be predicted by the first three variables. Consequently, in regressions of stock returns on the two premiums, such as those reported in Table 3 of Fama and French (1989), the two variables do have explanatory power at long horizons if other variables are excluded. Given these findings, there is little reason to recompute the long-horizon statistics with an expanded VAR.

7. Conclusions

In this article, I explore three alternative techniques for conducting inference and measurement in long-horizon forecasting environ-

[11] I thank Rob Stambaugh for supplying me with data on yields. I updated his series using the *Federal Reserve Bulletin,* the original source.

ments. Procedures for constructing standard errors under the null hypothesis that do not involve summing large numbers of autocovariances have better size than ones constructed under the alternative. Monte Carlo experiments indicate that substantial bias can arise in test statistics in long-horizon forecasting. After correcting for such biases, inference across the different procedures is quite similar. Such procedures are also quite powerful. Since the vector autoregressive alternative has correct size and supplies long-horizon statistics that appear to be unbiased measurements, it emerges as the preferred technique. One caveat to this statement is that the order of the VAR is taken as known in the Monte Carlo analysis.

The application investigates the predictability of stock returns at five horizons, from one month to four years. The VAR tests provide strong evidence of the predictive power of one-month-ahead returns at least for the sample from 1952 to 1987. The VAR analysis provides an alternative way to calculate various long-horizon statistics, including implied slope coefficients, implied R^2 's, and variance ratios. These implied long-horizon statistics indicate very interesting dynamic patterns in the data. The estimates and Monte Carlo results support the conclusion that changes in dividend yields forecast significant persistent changes in expected stock returns.

Finding the economic explanation that is consistent with the rejection of the null hypothesis of no expected return variability and the long-horizon predictability of returns is a challenging area for future research. Kandel and Stambaugh (1990) and Cecchetti, Lam, and Mark (1990) explore whether representative agent rational expectations models can be calibrated to produce long-horizon variance ratios and R^2 's analogous to those reported in Fama and French (1988a) and Poterba and Summers (1988). While their success demonstrates that such models cannot be dismissed out of hand, the arguments of Shleifer and Summers (1990) suggest that models with differential information may be required to reconcile the patterns in the data that are reported here.

Appendix A: Data

The data are from the Center for Research in Security Prices (CRSP) of the University of Chicago's Graduate School of Business. The four basic monthly series are the NYSE value-weighted with-dividend nominal return, RN_t, the value-weighted without-dividend nominal return, RX_t, the one-month Treasury-bill return, i_t, and the CPI inflation rate, π_t. The sample period is January 1926 to December 1987 for 744 observations.

Since $RX_t = (P_t - P_{t-1})/P_{t-1}$, a normalized nominal value-weighted price series is produced by setting the price in December 1925 equal to 1 and recursively setting $P_t = (1 + RX_t)P_{t-1}$. A normalized nominal divided series, d_t, is obtained by recognizing that

The annualized dividend for month t is

which sums the future values of the previous 11 months of dividends using the nominal interest rate factors obtained from the one-month Treasury-bill returns with the current dividend. The first observation is therefore December 1926, and the last observation is December 1987, for 733 observations.

A nominal goods price level, P_{gt}, is constructed from the monthly CPI inflation rates. Since $\pi_t = (P_{gt} - P_{gt-1})/P_{gt-1}$, a normalized nominal goods price level series is produced by setting the price in December 1925 equal to 1 and recursively setting $P_{gt} = (1 + \pi_t)P_{gt-1}$.

Real returns are constructed by dividing the nominal value-weighted price and divided for month t by the price level for that month and forming the return as real price plus real dividend divided by the previous month's real price.

Appendix B: The GARCH Model

In this appendix, I describe estimation of the GARCH model and its use as the data-generating process for the Monte Carlo simulations. Since there are six distinct elements in the conditional covariance matrix of the VAR, many GARCH models are possible. To avoid highly parameterized systems, the only model I explored is the constant conditional correlation model discussed in Bollerslev (1990). Let $H_t = E_t(u_{t+1}u'_{t+1})$ be the conditional covariance matrix of the VAR in Equation (11) with typical element $h_{ij,t}$. Model each of the three conditional variances as a first-order ARMA process:

$$h_{ij,t} = \omega_i + \beta_i h_{ii,t-1} + \alpha_i u_{it}^2, \qquad i = 1,2,3.$$

To model the covariances of H_t, estimate the nine parameters of Equation (B1) simultaneously with three constant correlation coefficients, ρ_{12}, ρ_{13}, and ρ_{23}, using maximum likelihood.

One problem with the estimation is induced by the large increase in the variance of the Treasury-bill return during the period from October 1979 to October 1982. Attempts to estimate a GARCH model that do not allow for an increase in the unconditional variance of the process during this period result in parameter estimates for which the conditional variance is an integrated process. I therefore first normalized the data for the Treasury-bill rate by dividing the error terms from this subperiod by their standard deviation estimated for

Table 8
Constant correlation GARCH model of the conditional variance matrix for the VAR of returns, dividend yields, and relative Treasury-bill rates

$$b_{ii,t} = \omega_i + \beta_i b_{ii,t-1} + \alpha_i u_{it-1}^2, \; i = 1,2,3$$

Conditional variance	ω_i (SE)	β_i (SE)	α_i (SE)	Type	Diagnostic tests Q(10)	Q(15)	Q(20)
$b_{11,t}$	117.714	0.908	0.046	A	9.513	12.563	17.155
	(54.563)	(0.031)	(0.016)	B	5.580	8.657	9.663
$b_{22,t}$	0.001	0.923	0.062	A	11.069	17.477	25.760
	(0.0002)	(0.013)	(0.011)	B	10.997	12.830	16.363
$b_{33,t}$	0.027	0.784	0.197	A	75.162	88.591	97.833
	(0.014)	(0.049)	(0.049)	B	7.969	12.420	15.084
	ρ_{12}	ρ_{13}	ρ_{23}	C	7.099	9.490	11.161
	(SE)	(SE)	(SE)	D	7.969	12.420	15.087
	−0.946	−0.072	0.052	E	14.888	16.815	19.399
	(0.005)	(0.048)	(0.047)				

Sample: 1952:1 to 1987:11, 431 observations. The conditional variance models are estimated simultaneously with the constant correlation coefficients. Conditional variances 1, 2, and 3 are innovation variances in the market return. the dividend yield, and the relative Treasury-bill return. Diagnostic test A refers to the level of the residual divided by the conditional standard deviation; B refers to the squared residual divided by the conditional variance; while C, D, and E refer to the product of two residuals divided by the product of their conditional standard deviations for (1, 2), (1,3), and (2,3), respectively. The .05 critical values of the x^2-statistics with 10, 15, and 20 degrees of freedom are 18.307, 24.996, and 31.410, respectively.

this period, by dividing the remaining data by their standard deviation estimated exclusive of this period, and by multiplying the entire series by the standard deviation for the whole sample period.

The parameter estimates are reported in Table 8. Most of the estimates are quite significantly different from zero. All three a, coefficients have confidence levels above .997, and the three β_i coefficients have confidence levels above .999. The contemporaneous correlation coefficient between the dividend yield and the equity return is highly significant as it should be, but the contemporaneous correlation between the relative Treasury-bill return and the equity return is not highly significant.

Also contained in Table 8 are diagnostic tests for serial correlation that examine either the residuals divided by their conditional standard deviations, the squared residuals divided by their conditional variances, or the cross-products of residuals divided by the cross-products of the respective standard deviations. In all cases except the normalized Treasury-bill rate, there is no evidence of additional serial correlation. The large values of the test statistics for the latter series indicate that the VAR may be misspecified.

To generate artificial data at each step in the Monte Carlo experiments, three standardized normal random variables, $\in_{t+1}$, are generated using the Gauss command RNDNS. The conditionally heteroskedastic innovations of the VAR are formed by taking the Cholesky decomposition of the conditional variance matrix, $C'_t C_t = Ht$, and

setting $u_{t+1} = C_t' \epsilon_{t+1}$. These residuals are fed into the VAR to generate the data, and they are used to update the conditional variance for the next step using Equation (B1). The initial residuals are generated from a normal distribution with an unconditional variance implied by the GARCH coefficients. In all simulations, the first 100 observations are discarded to reduce the influence of starting values.

References

Bollerslev, T., 1990, "Modeling the Coherence in Short Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model," *Review of Economics and Statistics,* 72. 498-505.

Campbell, J., 1991, "A Variance Decomposition for Stock Returns," *Economic Journal,* 101, 157-179.

Campbell, J. Y.. and R. J. Shiller, 1988, "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies,* 1, 195-228.

Campbell, J.. and R. Shiller, 1989, "The Dividend Ratio Model and Small Sample Bias: A Monte Carlo Study," *Economic Letters,* 29, 325-331.

Cecchetti, S., P. Lam. and N. Mark, 1990, "Mean Reversion in Equilibrium Asset Prices," *American Economic Review,* 80, 398-418.

Cochrane, J., 1988, "How Big Is the Random Walk Component of GNP?" *Journal of Political Economy,* 96, 893-920.

Cochrane, J. H., 1992, "Explaining the Variance of Price-Dividend Ratios," *Review of Financial Studies,* 5. 243-280.

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

Fama, E., and K. French, 1988b. "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics,* 22, 3-26.

Fama. E., and K. French, 1989. "Business Conditions and Expected Returns on Stocks and Bonds," *Journal of Financial Economics,* 25, 23-50.

French, K., W. Schwert, and R. Stambaugh, 1987, "Expected Stock Returns and Volatility," *Journal of Financial Economics,* 19, 3-30.

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

Hansen, L., and R. Hodrick, 1980. "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis," *Journal of Political Economy,* 88,829-853.

Jegadeesh, N., 1990, "Seasonality in Stock Price Mean Reversion: Evidence from the U.S. and the U.K.." working paper, University of California at Los Angeles.

Kandel, S., and R. Stambaugh, 1988, "Modeling Expected Stock Returns for Long and Short Horizons," working paper, University of Chicago and University of Pennsylvania.

Kandel, S., and R. F. Stambaugh, 1990, "Expectations and Volatility of Consumption and Asset Returns," *Review of Financial Studies,* 3, 207-232.

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

Kim, M., C. Nelson, and R. Startz, 1989, "Mean Reversion in Stock Prices? A Reappraisal of the Empirical Evidence," working paper, University of Washington.

Lo. A. W., and A. C. MacKinlay, 1988, "Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test," *Review of Financial Studies,* 1, 41-66.

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

Mankiw, N., D. Romer, and M. Shapiro, 1989, "Stock Market Forecastability and Volatility: A Statistical Appraisal," NBER Working Paper 3154.

Nelson, C., and M. Kim, 1990, "Predictable Stock Returns: Reality or Statistical Illusion?" working paper, University of Washington.

Newey. W.. and K. West, 1987, "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocor. relation Consistent Covariance Matrix," *Econometrica,* 55, 703-708.

Poterba, J., and L. Summers, 1988, "Mean Reversion in Stock Prices: Evidence and Implications," *Journal of Financial Economics,* 22, 27-59.

Richardson, M., 1990, "Temporary Components of Stock Prices: A Skeptic's View," working paper, University of Pennsylvania.

Richardson, M.. and T. Smith, 1991. "Test of Financial Models in the Presence of Overlapping Observations," *Review of Financial Studies,* 4, 227-254.

Richardson, M., and J. Stock, 1989, "Drawing Inferences from Statistics Based on Multi-Year Asset Returns," *Journal of Financial Economics,* 25, 323-348.

Shleifer. A., and L. Summers, 1990, "The Noise Trader Approach to Finance," *Journal of Economic Perspectives,* 4. 19-32.

Stambaugh, R., 1986, "Bias in Regressions with Lagged Stochastic Regressors," CRSP Working Paper 156, University of Chicago.