NBER WORKING PAPER SERIES

PREDICTABLE STOCK RETURNS: REALITY OR STATISTICAL ILLUSION?

Charles R. Nelson

Myung J. Kim

Working Paper No. 3297

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 1990

We are grateful to John Campbell for providing us with the data set assembled by him and Robert Shiller. Helpful comments from Andrew Lo, Pierre Perron, Richard Startz and participants in seminars at the Federal Reserve Board, Princeton University, and the University of Washington are acknowledged with thanks, but responsibility for errors is entirely that of the authors. This paper is part of NBER's research program in Financial Markets and Monetary Economics. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #3297
March 1990

PREDICTABLE STOCK RETURNS: REALITY OR STATISTICAL ILLUSION?

ABSTRACT

Recent research suggests that stock returns are predictable from fundamentals such as dividend yield, and that the degree of predictability rises with the length of the horizon over which return is measured. This paper investigates the magnitude of two sources of small sample bias in these results.

First, it is a standard result in econometrics that regression on the lagged value of the dependent variable is biased in finite samples. Since a fundamental such as the price/dividend ratio is a statistical proxy for lagged price, predictive regressions are potentially subject to a corresponding small sample bias. This may create the illusion that one can buy low and sell high in the sample even if the relationship is useless for forecasting. Second, multiperiod returns are positively autocorrelated by construction, raising the possibility of spurious regression. Standard errors which are computed from the asymptotic formula may not be large enough in small samples.

A set of Monte Carlo experiments are presented in which data are generated by a version of the present value model in which the discount rate is constant so returns are not in fact predictable. We show that a number of the characteristics of the historical results can be replicated simply by the combined effects of the two small sample biases.

Charles R. Nelson
Dept. of Economics, DK-30
University of Washington
Seattle, WA 98195

Myung Jig Kim
Dept. of Economics, Finance, &
Legal Studies
College of Commerce & Bus. Admin.
University of Alabama
Tuscaloosa, AL 35487-0224

Introduction

The proposition that stock returns are not predictable was until very recently regarded as one of the most (some would say the only) firmly established empirical results in economics. In his classic 1965 paper "The Behavior of Stock-Market Prices" Eugene Fama concluded that it had "presented strong and voluminous evidence in favor of the random walk hypothesis." Subsequent research over the next two decades only reinforced the evidence that neither past returns nor publicly available information were of any value in prediction. This large literature was widely interpreted as providing strong evidence that the capital markets are efficient in the sense of incorporating all available information in current prices. The extent to which the non-predictability result has been overturned in just the last few years in reflected in the opening statement in a recent paper by Fama and French (1988): "There is much evidence that stock returns are predictable." They cite estimates made by themselves and others that 25 to 40% of the variance in returns over periods of three to five years is predictable from past returns.

Two sources of predictability have been identified in the recent literature: past returns themselves and "fundamentals" such as dividend yield and price-earnings ratios. Poterba and Summers (1988) and Fama and French (1988b) report negative autocorrelation in returns over long horizons. Apparently, moves in prices tend to be reversed over several years, a tendency referred to as *mean reversion*. Lo and MacKinlay (1988) have reported evidence of positive autocorrelation at lags measured in weeks, suggesting persistence in returns over shorter periods. In an earlier paper (see Kim, Nelson and Startz (1989)) we have questioned the strength of the evidence for mean reversion over long horizons by demonstrating its dependence on pre-1947 data and by showing that randomization methods suggest that estimated standard errors may have been too small

The hypothesis that fundamentals should be useful in predicting stock returns follows from the seminal paper of Shiller (1981) which concluded that stock prices move too much to be justified by subsequent movement in dividends. If stock prices contain transient components unrelated to fundamentals but are anchored to fundamentals over the long term, then the fundamentals should contain information that is useful in predicting the future direction of prices. Indeed, Keim and Stambaugh (1986), Campbell and Shiller (1988), Fama and French (1988) and Cutler, Poterba, and Summers (1989) report that lagged ratios of dividends or earnings to price may explain more than 25% of the variation in stock returns measured over intervals of several years.

A finding that stock returns are to some extent predictable would not of itself contradict the efficient markets hypothesis since expected returns may well vary over time in a predictable way. Whether the observed degree of predictability constitutes evidence against efficient markets is, however, a subject of continuing debate in the literature. Cecchetti, Lam and Mark (1990) show that mean reversion can be a property of equilibrium returns in an economy where investors are risk averse.

The purpose of this paper is to consider the extent to which inferences about the predictability of stock returns might be influenced by small sample biases. One source of concern that we have comes from the use of multi-period overlapping stock returns data which introduces positive serial correlation by construction. For example, if annual one-year returns are serially random then overlapping annual observations on ten-year returns will have a MA(9) structure with first order autocorrelation equal to 0.9 by construction. Granger and Newbold (1974) cautioned against interpreting a high R^2 as evidence in itself of a relationship when the data are positively autocorrelated. Hansen and Hodrick (1980) and subsequent authors have recognized the need to correct regression standard errors in the case of overlapping observations. We are interested in investigating the adequacy of the correction in relevant sample sizes.

Another potential small sample problem is closely related to the bias which occurs in regressions involving a lagged dependent variable. To motivate the possibility of such a bias, consider the regression of the log of price, denoted p, on its lagged value,

$$p_t = a + b\, p_{t-1}.$$

If price is a random walk with drift then the true coefficient of p_{t-1} is one. It is a standard result in econometrics that least squares is biased in regression on a lagged dependent variable for finite samples, and in this case we know the sample slope coefficient b to be biased towards zero; see Fuller (1976) and Evans and Savin (1984). Subtracting lagged price from both sides of the equation we have

$$p_t - p_{t-1} = a + b' p_{t-1}$$

where b' = (b-1). In the case that price is a random walk, the true slope is zero but the expected value of the OLS coefficient b' is negative. Evans and Savin show that the bias is a decreasing function of the true value of the intercept "a" and of the sample size n. Thus, it will appear that the change in price is predictable from the price level.

The effect of the small sample bias is, therefore, to create the statistical illusion that it is possible to buy low and sell high when in fact future price changes are unpredictable in real time. To illustrate this we make use of the approximation due to Kendall (1954) which implies an expected value for b' of -(4/n) where n is the number of nonoverlapping observations. While this approximation was derived for stationary processes, it works well in the sample and parameter range relevant in this paper. Combining the bias in b' with the standard formula for the intercept in OLS one obtains the following expression for the predicted price change:

$$\widehat{(p_t - p_{t-1})} = \overline{(p - p_{-1})} - (4/n)\ (p_{t-1} - \bar{p}).$$

The regression says that the predicted price change is the average of past changes minus a fraction of the amount that the most recently observed price exceeds the average level of past prices. In other words, during the sample period it would have paid to buy when the price was below the average level in the sample and sell when it was above. The catch of course is that this rule for trading uses information that was only available after the sample period was over. The expected future price change given information available today is still just the drift parameter of the random walk.

Regressions of return on the log of the dividend-price ratio differ from this regression by the addition of the dividend return to the dependent variable and by the subtraction of the dividend from the explanatory variable, namely

$$p_t - p_{t-1} + rd_t = a + b''\ (p_{t-1} - d_{t-1})$$

where rd is the dividend return. In effect, the explanatory variable is now only one component of the log of price, the other component being the log of the dividend. Intuition would suggest that the size of the bias in b" would depend on the degree to which the price/dividend ratio is a proxy for price. If the dividend series is smooth then variation in (p - d) will be dominated by variation in p and the two will be strongly correlated over a finite sample. This correlation is apparent in the annual Standard and Poor's Index data from Campbell and Shiller (1988) plotted in Figure 1 for the period 1871-1986.

I. Lagged Price as a Predictor of Stock Returns

If lagged dependent variable bias is an important factor in the apparent predictive power of ratios involving price, then price by itself should also predict stock returns. The coefficients for the log of price and for the log of the price/dividend ratio should both be negative, reflecting the negative bias in regression on a lagged dependent variable. In Table 1 are reported the regression coefficients, t-ratios, and R^2 for a set of regressions based on those reported by Campbell and Shiller (1988) using their annual data set for the Standard and Poor's Index 1871-1987. The dependent variable is total return on the Index adjusted for inflation using the PPI. Explanatory variables are logs of the ratio of price to lagged dividends, versions of these using 10 and 30 year moving averages of dividends (denoted by superscripts), and the price alone. Note that price is measured at the beginning of year t and dividends are paid during year t. Return is measured alternatively over one year and overlapping three and ten year intervals. For each combination of return and price/dividend ratio there is a sample-matched regression on price. Following Hansen and Hodrick (1980), standard errors for t-ratios take into account the serial correlation induced by overlapping observations; see Appendix 1. The resulting ratio of coefficient to asymptotic standard error is denoted HH t(b).

Table I

OLS Coefficient, HH t-ratio, and R^2 for Predictive Regressions
Real Total Return for the Standard and Poor's Index 1871-1987

	b	HH t(b)	R^2
Predictor:	One Year Return		
(p - d)	-.12	-2.33	.039
p	-.06	-2.41	.036
$(p - d^{10})$	-.08	-1.85	.025
p	-.06	-2.25	.035
$(p - d^{30})$	-.11	-3.03	.061
p	-.07	-2.16	.037
	Three Year Return		
(p - d)	-.36	-3.00	.110
p	-.17	-2.96	.120
$(p - d^{10})$	-.24	-2.26	.077
p	-.18	-2.70	.110
$(p - d^{30})$	-.33	-3.47	.186
p	-.21	-2.60	.121
	Ten Year Return		
(p - d)	-.97	-3.42	.267
p	-.57	-5.16	.397
$(p - d^{10})$	-.94	-4.54	.372
p	-.59	-4.44	.359
$(p - d^{30})$	-1.02	-6.66	.559
p	-.68	-5.16	.393

As predicted by lagged dependent variable bias, all the slope coefficients are negative. Further, price by itself has about as much or more explanatory power and statistical significance as do the ratios of price to the lagged dividends and to the 10 year moving average of dividends. The coefficient on the price/dividend ratio is larger that on price in every case, reflecting in part the smaller sample variance of the ratio. While these features of the historical regressions are consistent in direction with the bias explanation of predictability, they are also consistent with the hypothesis that price itself is a source of information on future returns. We know of no economic motivation, however, for such a hypothesis. Further, it would imply that expected return is nonstationary with a positive drift through time since that is true of the price level.

Note that the longer the horizon over which return is measured, the greater is the fraction of variation that is predicted and the stronger is the statistical significance. The increase in R^2 with return horizon has been emphasized in this literature as a strong and important feature of the empirical evidence. However, the positive serial correlation induced by overlapping observations would in itself tend to increase R^2. Granger and Newbold (1974) showed that the expected value of R^2 rises with the degree of autocorrelation regardless of any relation between the variables, creating a spurious correlation. They conclude (*op cit* p. 114): "Thus, a high value of R^2 should not, on the grounds of traditional tests, be regarded as evidence of a significant relationship between autocorrelated series." The t-ratios are therefore of greater relevance in judging whether predictability increases with return horizon. However, since price by itself shares this property with the ratios, the horizon effect may also reflect the fact that the bias is a decreasing function of (non-overlapping) sample size which is reduced by the calculation of returns over longer horizons. The relation of significance to horizon may also involve the small sample properties of the HH standard errors used in the case of overlapping returns.

When the 30 year moving average of dividends used in calculating the price/dividend ratio it results in greater the explanatory power and statistical significance than price alone. One consequence of a longer moving average is shortened sample size and therefore larger bias, but price by itself in the sample-matched regressions shows no correspondingly large effect. Similarly, increased small sample bias in the HH standard errors would also have shown up in regressions on price alone. Whether the 30 year results are too large to be attributed to sampling error depends on the unknown small sample distribution.

The next section of the paper describes a Monte Carlo experiment designed to investigate the extent to which the main features of the historical results can arise in hypothetical data where price does in fact adjust fully to fundamentals and returns are not predictable.

II. Monte Carlo Estimates of Small Sample Bias

The strategy behind a Monte Carlo experiment is to generate artificial data under the null hypothesis and tabulate the empirical distribution of sample statistics. We take as our null hypothesis the present value model of stock prices which says that the price is the discounted present value of expected dividends or net cash flow. This is, however, an incomplete specification since we need to say how information about the future is generated and how the term structure of discount rates moves over time. The actual predictability of returns will depend on these details of the complete specification and, as we shall see, the small sample bias in predictive regressions will also. For this reason there cannot be a unique bias, rather we can only report the bias found under particular specifications which at least account for some observable features of the data.

To complete the null hypothesis we assume that expected future real dividends are discounted at rate r which is constant through time. Under this specification, expected real return is of course constant through time and equal to r. The actual predictability of real returns is zero, so any that we find in predictive regressions is spurious. To investigate the effect of dividend smoothing on the bias we assume a version of the Lintner (1956) model with values of the smoothing parameter taken over a range suggested by the historical data. We anticipate that a smoother dividend series will produce stronger spurious predictability as the price/dividend ratio becomes a better proxy for price.

The target level of dividends toward which actual dividends adjust has the character of a long horizon forecast. From the result of Beveridge and Nelson (1981) it will therefore be a random walk with drift. The generating process is then

$$d^*_t = d^*_{t-1} + \gamma + u_t$$

where d^*_t is the target level for the log of dividends that is established in period t, γ is the long run growth rate of dividends, and u is i.i.d. $N(0,\sigma^2)$. The actual dividend paid during t depends on d^*_t and on past history according to the partial adjustment equation

$$d_t = d_{t-1} + \lambda\,(d^*_{t-1} - d_{t-1}).$$

These two equations specify the process for actual dividends and therefore the forecasting equation used to calculate price through the present value formula. As in the C&S data, price is generated at the beginning of the period (year) conditional on observed dividends through t-1. Since d^* is the long horizon forecast of d, it is irrelevant whether we assume that agents know d^* explicitly or simply forecast the actual dividend.

Values for parameters gamma and σ^2 to used in the simulation were obtained from corresponding sample moments for p since the change in p will be dominated by the change in d^* for a constant discount rate. These are .015 and .03 respectively. The discount rate is the historical average real return, .066. We expect the lagged dependent variable bias to depend on the degree to which $(p-d_{-1})$ is a proxy for p, which will in turn increase as we consider smaller values of λ. The AR(1) model estimated for historical dividends suggests a value of λ of about .2 with a standard error of 0.1, but it is clear that λ was considerably smoother after WWII than before. To see how the bias varies with λ we examine results for λ equal to 0.5 and 0.1, denoted experiments A and B respectively, representing a range that includes the point estimate. The process is given 100 periods to stabilize before the sample data are taken. The number of replications is 1000 and the random disturbances driving the d^* process are normally distributed. Further details may be found in Appendix 2.

Tables II, III, and IV present Monte Carlo estimates of the distribution of the slope coefficients, R^2, and their t-ratios respectively for this model. Sample size is initially set at 116 for comparison with the historical Standard and Poor's results from C&S which are tabulated for comparison. The dependent variable is total return over, successively, one, three and ten periods. The predictive variables are the log of the ratio of price to dividends and to ten and thirty year moving averages of dividends as well as the log of price by itself. In addition, a random walk variable called z which is unrelated to anything else is also a used as a predictive variable to check the distributions of t-ratios computed using the HH correction in the case of multiperiod returns and to calibrate the distribution of R^2 (the coefficient of z having expectation zero).

Table II A

Empirical Distribution of the Slope Coefficient
Speed of Adjustment of Dividends: $\lambda = 0.5$
Sample Size is 116

Predictor	Historical	Monte Carlo: Mean	Fractiles .025	.975
		One Year Return		
(p - d)	-.12	-.03	-.38	.26
p	-.06	-.04	-.12	.01
$(p - d^{10})$	-.08	-.03	-.16	.07
$(p - d^{30})$	-.11	-.04	-.17	.03
		Three Year Return		
(p - d)	-.36	-.09	-.81	.66
p	-.17	-.10	-.35	.02
$(p - d^{10})$	-.24	-.07	-.39	.18
$(p - d^{30})$	-.33	-.11	-.43	.09
		Ten Year Return		
(p - d)	-.97	-.28	-1.58	1.44
p	-.57	-.31	-.88	.10
$(p - d^{10})$	-.94	-.18	-.83	.56
$(p - d^{30})$	-1.02	-.31	-1.06	.26

Table II B

Empirical Cumulative Distribution of the Slope Coefficient
Speed of Adjustment of Dividends: $\lambda = 0.1$
Sample Size is 116

Predictor	Historical	Monte Carlo: Mean	Fractiles .025	.975
		One Year Return		
(p - d)	-.12	-.04	-.17	.05
p	-.06	-.03	-.11	.01
$(p - d^{10})$	-.08	-.03	-.13	.03
$(p - d^{30})$	-.11	-.04	-.13	.02
		Three Year Return		
(p - d)	-.36	-.11	-.46	.14
p	-.17	-.09	-.31	.02
$(p - d^{10})$	-.24	-.09	-.35	.08
$(p - d^{30})$	-.33	-.11	-.39	.05
		Ten Year Return		
(p - d)	-.97	-.30	-1.07	.38
p	-.57	-.26	-.74	.06
$(p - d^{10})$	-.94	-.24	-.82	.23
$(p - d^{30})$	-1.02	-.29	-.90	.17

Slope coefficients

The slope coefficients are negatively biased in all of these regressions. When price level is the explanatory variable, Kendall's approximation, (-4/n) where sample size n is taken to be the number of non-overlapping observations, provides a good rule of thumb for the magnitude of the bias. For one year returns we have -0.035, for three year returns -0.104, and for ten year returns -0.348. Thus, reduction in effective sample size explains why the multiperiod return regressions are more biased than those for one period returns.

The bias in the regression on (p - d) is larger when λ is 0.1 than when λ is 0.5. The more smoothing there is of dividends the more (p - d) reflects variation in price and is therefore a better proxy for price. The fractiles also show that the sampling distribution of the ratio coefficient becomes more like that of the price coefficient with smaller values of λ. The influence of λ on bias is smaller when the moving average of dividends is used to calculate the ratio, presumably since dividends are smoothed anyway by the moving average. Recall that the coefficients for the ratios were larger in magnitude than those for price in the historical regressions. This is also the case on average in the artificial data when λ is 0.1.

Calculation of the price/dividend ratio using a moving average of dividends evidently has offsetting effects. As we go from no smoothing to the ten year moving average the bias decreases in magnitude. The averaging of dividends has a smoothing effect, tending to make the ratio a better proxy for price, and a noise-introducing effect since variation in the moving average involves lagged dividend innovations which are irrelevant to current price. When we go to the thirty year moving average from ten years, the bias increases in magnitude. Observations are lost in calculating the moving average and the smaller sample size tends to increase the bias.

To what degree do the sampling distributions of the Monte Carlo experiments account for the historical results? The negative signs of the coefficients, the increase in magnitude with the horizon over which return is measured, the coefficient for the ratio being larger than the coefficient of price alone, and the U-shaped pattern with respect to the length of the moving average of dividends are common to both. The historical coefficients, however, are larger in magnitude than the Monte Carlo means. One is outside the 95% range for λ equal to 0.5 and two are outside for λ equal to 0.1; all these cases are for ten year returns with dividend averaging. Of course, these are not statistically independent.

Table III A

Empirical Cumulative Distribution of R^2
Speed of Adjustment of Dividends: $\lambda = 0.5$
Sample Size is 116

Predictor	Historical	Monte Carlo: Mean	Fractile .95
		One Year Return	
(p - d)	.039	.009	.034
p	.036	.023	.065
$(p - d^{10})$	.025	.011	.042
$(p - d^{30})$	.061	.018	.067
z	na	.009	.033
		Three Year Return	
(p - d)	.110	.016	.059
p	.120	.066	.186
$(p - d^{10})$	.077	.029	.109
$(p - d^{30})$	.186	.050	.179
z	na	.025	.094
		Ten Year Return	
(p - d)	.267	.028	.103
p	.397	.197	.481
$(p - d^{10})$	.372	.066	.233
$(p - d^{30})$	.559	.139	.445
z	na	.077	.274

Table III B

Empirical Cumulative Distribution of R^2
Speed of Adjustment of Dividends: $\lambda = 0.1$
Sample Size is 116

Predictor	Historical	Monte Carlo: Mean	Fractile .95
		One Year Return	
$(p - d)$	.039	.011	.042
p	.036	.024	.068
$(p - d^{10})$	.025	.013	.047
$(p - d^{30})$	.061	.021	.067
z	na	.008	.032
		Three Year Return	
$(p - d)$	.110	.031	.111
p	.120	.070	.184
$(p - d^{10})$	.077	.038	.133
$(p - d^{30})$	.186	.059	.186
z	na	.023	.088
		Ten Year Return	
$(p - d)$	.267	.082	.279
p	.397	.204	.490
$(p - d^{10})$	.372	.101	.341
$(p - d^{30})$	.559	.165	.479
z	na	.075	.278

R^2

Monte Carlo estimates of the mean of R^2 for the ratios increase with return horizon. This reflects the spurious regression phenomenon identified by Granger and Newbold (1974) arising from the induced positive autocorrelation of multiperiod returns, which is confirmed by similar results for z (the unrelated random walk variable).

The mean R^2 also increases with the length of the moving average for dividends. This second effect partially reflects decreased sample size since sample-matched regressions on price alone (not displayed) show a similar but less dramatic increase. It is also partially due to the ratio becoming a better proxy for price as the length of the moving average increases.

Note that price by itself has more explanatory power than do the ratios. The smaller value of λ makes actual dividends smoother, so again the ratio is a better proxy for price resulting in higher mean R^2.

The pattern of historical values is similar but not entirely consistent with the Monte Carlo sampling distribution. Historical R^2 does rise with horizon over which the return is calculated. Taking a ten year moving average of dividends does not increase R^2 in the one and three year regressions, but it does when we go to a thirty year moving average. Price by itself has a larger R^2 for the three and ten year returns regressions. Finally, the historical R^2 for the ratios are compared to the .95 fractile of the sampling distribution since it is only large values of the statistic that contradict the null hypothesis. Note that the value of λ has a large impact on the .95 fractile. Historical R^2 exceed the .95 fractile for all of the ten year, two of the three year, and one of the one year return regression when we assume λ is .5. When the dividend process is more smooth with λ equal to 0.1, all but two of the ten year return results are within the range.

Table IV A

Empirical Cumulative Distribution of the HH t-ratio
Speed of Adjustment of Dividends: $\lambda = 0.5$
Sample Size is 116

Predictor	Historical	Monte Carlo: Mean	Fractiles .025	.975
		One Year Return		
(p - d)	-2.33	-.17	-2.22	1.85
p	-2.41	-1.41	-3.39	.57
$(p - d^{10})$	-1.85	-.39	-2.61	1.63
$(p - d^{30})$	-3.03	-.79	-2.90	1.26
z	na	.02	-2.03	1.99
		Three Year Return		
(p - d)	-3.00	-0.35	-2.93	1.84
p	-2.96	-1.76	-4.37	.58
$(p - d^{10})$	-2.26	-.52	-3.09	1.87
$(p - d^{30})$	-3.47	-1.05	-3.86	1.36
z	na	0.01	-2.59	2.49
		Ten Year Return		
(p - d)	-3.42	-.81	-4.57	1.81
p	-5.16	-2.45	-6.89	1.03
$(p - d^{10})$	-4.54	-.95	-5.18	1.98
$(p - d^{30})$	-6.66	-1.71	-6.83	1.68
z	na	.03	-3.39	3.49

Table IV B

Empirical Cumulative Distribution of the HH t-ratio
Speed of Adjustment of Dividends: $\lambda = 0.1$
Sample Size is 116

Predictor	Historical	Monte Carlo: Mean	Fractiles .025	.975
		One Year Return		
$(p - d)$	-2.33	-.62	-2.61	1.39
p	-2.41	-1.49	-3.24	.40
$(p - d^{10})$	-1.85	-.74	-2.65	.1.17
$(p - d^{30})$	-3.03	-1.02	-2.83	.99
z	na	-.05	-2.08	1.88
		Three Year Return		
$(p - d)$	-3.00	-.82	-3.45	1.53
p	-2.96	-1.82	-4.35	.51
$(p - d^{10})$	-2.26	-.94	-3.59	1.45
$(p - d^{30})$	-3.47	-1.30	-3.98	1.18
z	na	-.06	-2.50	2.30
		Ten Year Return		
$(p - d)$	-3.42	-1.31	-5.34	1.54
p	-5.16	-2.48	-7.43	.64
$(p - d^{10})$	-4.54	-1.39	-5.46	1.48
$(p - d^{30})$	-6.66	-1.97	-6.72	1.38
z	na	-.06	-3.57	3.37

The t-ratio

There are two issues to be investigated with regard to the sampling distribution of the t-ratio, which is the ratio of the OLS slope coefficient to the HH standard error. One is the negative shift in the distribution due to the negative bias in the coefficient. The other is the appropriateness of the HH correction for autocorrelation induced by overlapping multiperiod returns. To isolate the latter effect we included z, the unrelated random walk series, as another predictor of returns.

The distribution of the t-ratio for the case of autoregression when the data are a driftless gaussian random walk has been investigated by Dickey (1976), Fuller (1976) and Dickey and Fuller (1979), and it is centered around roughly -1.5 for any sample size larger than 25. This is essentially what we see in Table IV for the regression of one year return on lagged price and estimated fractiles also compare closely with Table 8.5.2 of Fuller (1976). The rate of drift of the price series is not large enough to have an important effect on the distribution along the lines suggested by Evans and Savin (1984). We expect that the effect of a slower rate of adjustment of dividends will be to make the distribution of the t-ratio for the price/dividend ratio more like that for price by itself. This is what we see in Table IV as a slower speed of adjustment or a longer moving average is considered.

A multiperiod horizon for the return introduces positive autocorrelation in the residuals which the HH correction is designed to take into account. If we did not use the correction then standard econometric theory leads us to expect too large a dispersion for the t-ratio, but if the correction is successful the range between the .025 and .975 fractiles should be roughly 4. What we see in Table IV is that this range increases to about 5 for three year returns and to about 7 or 8 for ten year returns. Since the z variable shows the same pattern the spreading is clearly due to the HH adjustment to the standard error not being large enough to correct for autocorrelation. Of course, the HH standard errors would be correct asymptotically so this can be viewed as another small sample bias.

While the historical t-ratios are larger in magnitude than the sampling means, they are generally within the .025-.975 range. With λ at 0.5 there are three exceptions, and with λ at 0.1 the single exception is for the regression of the one year return on the ratio which uses the 30 year moving average of dividends. The difference between taking the historical t-ratios at face value and taking them as drawings from the sampling distributions of this experiment is the difference between a significance level around .05 and one that is infinitesimal. Equivalently, it is the difference between a test statistic that is about two standard deviations from the mean and one that is 3 to 6.

III. Larger Sample Size and Monthly Data

In this section we describe two additional experiments designed to show how small sample bias changes with sample size and with more frequent intervals of observation. In the first experiment we simply double the sample size to 232 while retaining the annual observations generating model. In the second experiment we consider what happens if prices and returns are observed monthly. This is different from just going from $n=116$ to $n=12x116$ because λ is the annual rate of adjustment of dividends. The results for the first experiment are given in Table V for a λ value of 0.1.

Table V

Empirical Means of the Slope, R^2, and HH t-ratio
Speed of Adjustment of Dividends: $\lambda = 0.1$
Sample Sizes 116 and 232 Years

	Slope		R^2		HH t-ratio	
Sample N:	116	232	116	232	116	232
			One Period Return			
(p - d)	-.04	-.02	.011	.005	-.62	-.46
p	-.03	-.01	.024	.009	-1.49	-1.15
$(p - d^{10})$	-.03	-.01	.013	.005	-.74	-.52
$(p - d^{30})$	-.04	-.01	.021	.007	-1.02	-.68
z			.008	.004	-.05	..05
			Three Period Return			
(p - d)	-.11	-.05	.031	.014	-.82	-.57
p	-.09	-.03	.070	.026	-1.82	-1.35
$(p - d^{10})$	-.09	-.04	.038	.016	-.94	-.63
$(p - d^{30})$	-.11	-.04	.059	.020	-1.30	-.82
z			.023	..013	-.06	..06
			Ten Period Return			
(p - d)	-.30	-.16	.082	.037	-1.31	-.79
p	-.26	-.09	.204	.085	-2.48	-1.61
$(p - d^{10})$	-.24	-.12	.101	.044	-1.39	-.81
$(p - d^{30})$	-.29	-.12	.165	.060	-1.97	-1.03
z			.075	..044	-.06	..06

The work of Dickey, Fuller and Evans and Savin give some guidance about what we should expect to see when the number of annual observations is doubled. For regression on the lagged level of a driftless random walk we know that the bias is proportional to 1/n, and the results of Evans and Savin (1984) suggest that the bias will decrease faster in the presence of a nonzero expected change. The bias for the coefficient of lagged price is indeed about a third as large when n is doubled, and there is a similar decrease for the price/dividend ratio. The t-ratio is essentially independent of sample size in the driftless random walk case, and this implies that R^2 will decline in proportion to 1/n. We see in Table V that the decline in mean R^2 is somewhat faster than this, relecting a decline in mean t-ratios with the doubling of n. These results suggest that the erroneous inference of predictability becomes less of a problem as sample size grows.

Recall that the sampling variance of the t-ratio was also too large for three and ten year returns because the HH correction for overlapping observations was not large enough. Fractiles of the t-ratio which are not shown confirm that the sampling distribution is also becoming less disperse with sample size, although it is still too large. For n=116 the distances between the .025 and .975 fractiles were approximately 4, 5, and 7 for one, three and ten year returns respectively when the independent variable was (p-d). For n=232 these distances become approximately 4, 4.8, and 5.5. A similar pattern was also found when the predictive variable was the unrelated random walk z. Thus the contribution of overlapping observations to the inference of predictability yields only slowing to increasing sample size.

In the second experiment we consider the whether monthly observations instead of annual mitigate small sample bias. The data generating mechanism has the same form as that for annual data except that the target dividend process is monthly with drift and variance one twelfth as large. Actual dividends adjust toward the target at one twelfth the annual rate (λ) and price is calculated from the discounted present value of expected future divdiends each month. We have 12x116=1392 monthly observations on total return, the log of price, and the price/dividend ratio. Results for one period return regressions are reported in Table VI.

Table VI

Empirical Distribution of the Slope, R^2, and HH t-ratio
Annual Speed of Adjustment of Dividends: $\lambda = 0.1$
Sample Sizes are 116 Years and 1392 (12x116) Months

		Slope Coefficient		
		Mean	.025 fractile	.975 fractile
(p-d)				
	Annual	-.04	-.17	.05
	Monthly	-.003	-.011	.003
p				
	Annual	-.03	-.11	.01
	Monthly	-.002	-.007	-.00

		R^2	
		Mean	.95 fractile
(p-d)			
	Annual	.011	.042
	Monthly	.001	.003
p			
	Annual	.024	.068
	Monthly	.002	.005

		HH t-ratio		
		Mean	.025 fractile	.975 fractile
(p-d)				
	Annual	-.62	-2.61	1.39
	Monthly	-.62	-2.22	1.11
p				
	Annual	-1.49	-3.24	.40
	Monthly	-1.45	-3.10	-.06

Note: Based on 200 replications of monthly data.

Note that the distribution of the t-ratio is essentially the same in monthly data as it is in annual data. Observing the process at monthly intervals instead of annual does not alter the apparent statistical evidence for predictability. This is consistent with the insensitivity of the distribution of the t-ratio to sample size reported by Dickey (1976) in autoregressions with unit roots. While the monthly observations multiply the number of independent observations by 12, that is irrelevant. The relevant difference is that we are observing the same nonstationary process 12 times more frequently, running what amounts to an autoregression as explained in the introduction, and that does not substantially alter our inferences about the process. The values of R^2 are proportionally smaller for monthly data in accordance with the algebraic relation between t and R^2. Of course the slope coefficient is about one twelfth as large for monthly data since the corresponding predicted return is at a montly rate.

Fama and French (1988) reported t-ratios in the range of 1.15 to 2.76 for regressions of the monthly returns from the CRSP tapes 1926-86 on the the raw dividend yield. We intend to run a further experiment that is more directly comparable to their regression, however based on the above experiment we would not expect that the reported t-ratios are as strongly significant as they had appeared.

IV. Summary and Conclusions

A number of recent studies have reported that stock returns can be predicted to a statistically significant degree by fundamentals and that predictability increases as one considers returns over multiperiod horizons. In this paper we have looked at whether these findings could to some degree be attributed to small sample biases. Regression on a lagged dependent variable is subject to small sample bias and this suggests that regression of stock returns on the price/dividend ratio or similar value measures involving price may be subject to a related bias. Further, multiperiod returns are by construction positively autocorrelated, suggesting the possibility of spurious regression as another source of apparent predictability. To investigate the size and economic significance of these biases we have run a Monte Carlo experiment in which prices and returns are generated by a version of the present value model in which expected returns are constant. While the generating model replicates some features of the distribution of historical annual real returns since 1871, it is not the only specification of the present value model nor the only one with constant expected returns. A different specification of the null hypothesis would presumably imply somewhat different biases.

Under the specification we have chosen, a number of the features of the historical regression results arise as the consequence of small sample biases: reliably negative coefficients on the price/dividend ratio which seems to suggest one can buy low and sell high, an increase in R^2 as returns are calculated over longer horizons, and the apparent increase in statistical significance as longer horizons and more dividend averaging are used. Judging the statistical significance of regressions with multiperiod returns requires an appropriate upward adjustment of standard errors due to the induced autocorrelation. We find that the correction based on asymptotic theory is not large enough in small samples, resulting in spuriously large t-ratios.

Generating monthly data of equivalent length in years, we find that statistical inferences are essentially the same as in annual data. Although it would seem that one would have twelve times the number of independent observations, the relevant analogy is to taking observations on a nonstationary series twelve times as frequently. The latter does not change greatly inferences about the process.

The magnitude of the historical results reported for annual returns by Campbell and Shiller (1988) and monthly returns by Fama and French (1988) is such that they cannot be dismissed entirely as the result of small sample bias under our generating model. Their statistical significance, however, would seem to be substantially weaker than if taken at face value.

REFERENCES

Beveridge, Stephen and Charles R. Nelson, 1981, A new approach to the decomposition of economic time series into permanent and transitory components with particular attention to measurement of the "business cycle," Journal of Monetary Economics 7, 151-174.

Campbell, John Y., and Robert J. Shiller, 1988, Stock prices, earnings, and expected dividends, *The Journal of Finance* 43, 661-676.

Cecchetti, Stephen G., Pok-sang Lam, and Nelson C. Mark, 1990, Mean reversion in equilibrium asset prices, *American Economic Review* 80, ??.

Cutler, David M., James M. Poterba, and Lawrence H. Summers, 1989, Speculative Dynamics, paper presented at London School of Economics conference on econometrics of financial markets, June 1989.

Dickey, David A. and Wayne A. Fuller, 1979, Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association* 74, 427-431.

Dickey, David A., 1976, Estimation and hypothesis testing in nonstationary time series, Ph.d. dissertation, Iowa State University.

Evans, G. B. A. and N. E. Savin, 1984, Testing for Unit Roots: 2, *Econometrica* 52, 1241-1269.

Fama, Eugene F. and Kenneth R. French, 1988, Dividend yields and expected stock returns, *Journal of Financial Economics* 22, 3-25.

Fuller, Wayne A., 1976, *Introduction to Statistical Time Series* (John Wiley, New York).

Granger, C. W. J. and P. Newbold, 1974, Spurious regressions in econometrics, *Journal of Econometrics* 2, 111-120.

Hansen, Lars Peter, and Robert J. Hodrick, 1980, Forward Exchange rates as optimal predictors of future exchange rates: an econometric analysis, *Journal of Political Economy* 88, 829-853.

Kendall, M. G., 1954, Note on bias in the estimation of autocorrelation, *Biometrika* 41, 403-404.

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

Kim, Myung-Jig, Charles R. Nelson and Richard Startz, 1989, Mean reversion in stock prices: a reappraisal of the empirical evidence, NBER Working Paper No. 2795.

Lintner, John, 1956, Distribution of income of corporations among dividends, retained earnings and taxes, *American Economic Review* 46, 97-113.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskadasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703-708.

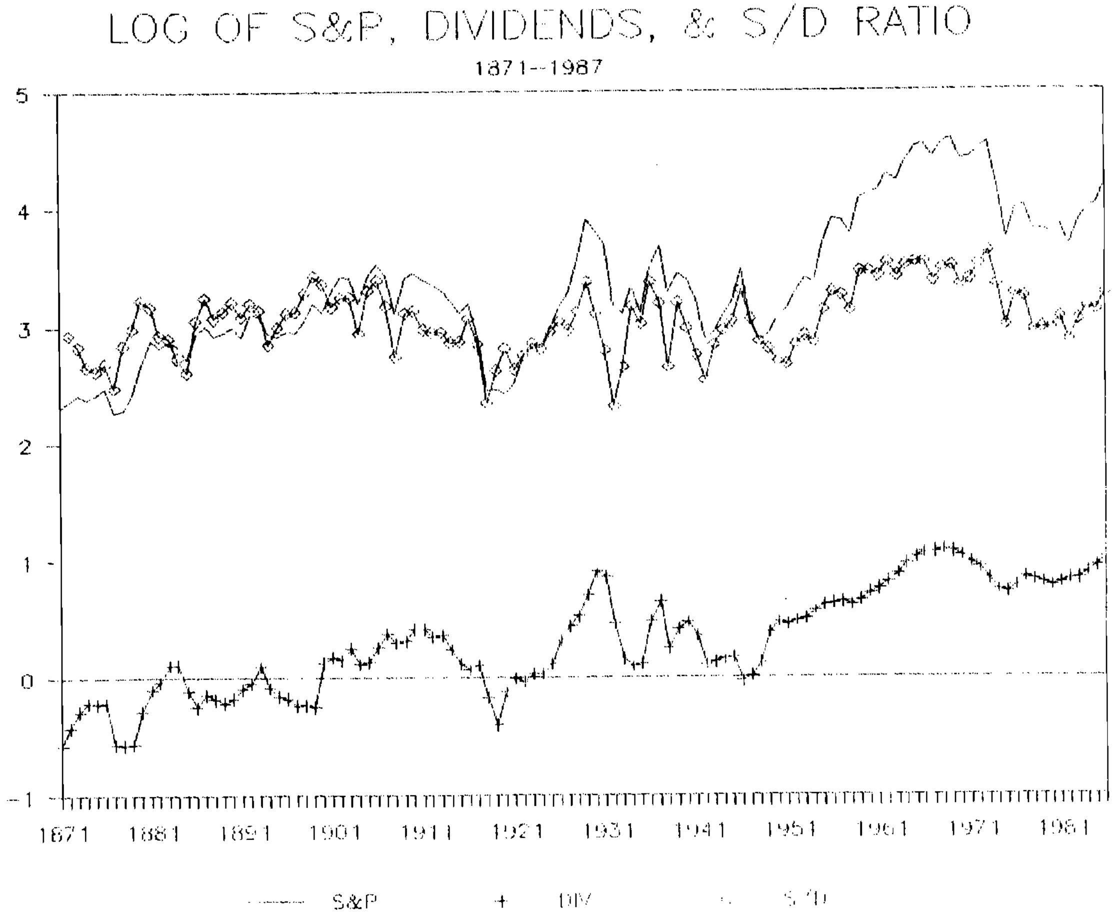
Phillips, P. C. B., 1987, Time series regression with a unit root, *Econometrica* 55, 277-301.

Poterba, James M. and Lawrence H. Summers, 1988, Mean reversion in stock prices, *Journal of Financial Economics* 22, 27-59.

Shiller, Robert J., 1981, Do stock prices move too much to be justified by subsequent changes in dividends, *American Economic Review* 75, 421-436.

White, Halbert, 1980, A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica* 48, 817-838.

Figure 1



LOG OF S&P, DIVIDENDS, & S/D RATIO
1871--1987
5
4
3
2
1
0
-1
1871
1881
1891
1901
1911
1921
1931
1941
1951
1961
1971
1981
S&P
DIV
S/D

Appendix 1: The Correction for the Standard Errors in the Multiyear Returns Regression

Our procedure of calculating the standard errors is in the spirit of White's [1980] consistent estimators for the covariance matrix of estimators. That is, the asymptotic covariance matrix of β_K in the regression of K-multiyear returns ($r_{t,t+K}$) on the predictor X_t can be estimated from

$$\hat{\Sigma}_\beta = (X_T'X_T)^{-1}X_T'U_TU_T'X_T(X_T'X_T)^{-1}$$

where $X_T = (x_1, \cdots, x_T)', U_T = (\hat{u}_1, \cdots, \hat{u}_T)'$ and $\hat{u}_i$'s are the least square residuals. As is shown in H&H, the T by T matrix U_TU_T' will include not only the T-elements main diagonal, but also K-1 off-diagonal terms (pararell to the main diagonal) reflecting the (K-1)th order serial correlations induced by overlapping data, and zeros elsewhere. The difference between the original H&H correction and ours lies in that H&H imposed equality along the diagonals. In fact the standard errors estimated from both methods are very close. C&S reported SE's of .0570, .1443 and .2997 from the regression of 1, 3 and 10-year returns on the dividend-price ratio, while the White version estimates .0529, .1183 and .2843, respectively.

When there are large negative sample serial covariances the variances can take on negative values. To ensure the positive definiteness of the variances the nonzero off-diagonal terms may be modified by multiplying those by the arithematically declining weights proposed by Newey and West [1987] and Phillips [1987]. Thus the modification may be written as

$$X_T'U_TU_T'X_T = \sum_{t=1}^{T} u_t x_t' x_t u_t + 2 \cdot \sum_{l=1}^{K-1}(1 - \frac{l}{K+1}) \sum_{t=l+1}^{T} u_t x_t' x_{t-l} u_{t-l}$$

and the t-statistics based on the modified standard error are calculated in the Monte Carlo simulation and are reported as $HHt(b)$ in the paper.

Appendix 2: Data Generating Mechanism under the Present Value Model

This section briefly describes the data generating mechanism under the null hypothesis of the PV model. We assume that the actual log of dividend (d_t) follows a partial adjustment process towards the target level (d_t^*):

$$d_t = d_{t-1} + \lambda(d_{t-1}^* - d_{t-1}) \tag{1}$$

$$d_t^* = d_{t-1}^* + \gamma + u_t, \quad u_t \sim i.i.d.\ N(0, \sigma_u^2) \tag{2}$$

where γ is the long run growth rate of dividends. In the PV model, the price P_t is

$$P_t = E_t(\tilde{D}_t)/(1+r) + E_t(\tilde{D}_{t+1})/(1+r)^2 + \cdots \tag{3}$$

where " $\tilde{}$ " means a random variable and the capital letter for the level. For a simplifying approximation to the expected future dividend, we note

$$\begin{aligned} E_t(\tilde{D}_{t+i}) &= E_t\ exp\{\tilde{d}_{t+i}\} \\ &= exp\{E_t(\tilde{d}_{t+i}) + \frac{1}{2}Var_t(\tilde{d}_{t+i})\} \\ &\propto exp\{E_t(\tilde{d}_{t+i})\} \end{aligned} \tag{4}$$

since the density function for $\tilde{D}_{t+i}$ is log-normal and the variance $Var_t(\tilde{d}_{t+i})$ does not vary over time. Now let $\hat{d}_{t,t+i}$ denote the expected value of $\tilde{d}_{t+i}$ at the beginning of period t. Similarly, $\Delta\hat{d}_{t,t+i} = E_t\Delta\tilde{d}_{t+i}$. Using the ARI(1,1) model implied by the equation (1) and (2), we construct the forecasts recursively,

$$\begin{aligned} \Delta\hat{d}_{t,t} &= (1-\lambda)\Delta d_{t-1} + \lambda\gamma; \quad \hat{d}_{t,t} = d_{t-1} + \Delta\hat{d}_{t,t} \\ \Delta\hat{d}_{t,t+1} &= (1-\lambda)\Delta\hat{d}_{t,t} + \lambda\gamma; \quad \hat{d}_{t,t+1} = \hat{d}_{t,t} + \Delta\hat{d}_{t,t+1} \\ &\vdots \\ \Delta\hat{d}_{t,t+i} &= (1-\lambda)\Delta\hat{d}_{t,t+i-1} + \lambda\gamma; \quad \hat{d}_{t,t+i} = \hat{d}_{t,t+i-1} + \Delta\hat{d}_{t,t+i} \\ &\vdots \\ \Delta\hat{d}_{t,t+BIG} &\approx \gamma; \quad \hat{d}_{t,t+BIG} = \hat{d}_{t,t+BIG-1} + \gamma \end{aligned} \tag{5}$$

The price P_t is then

$$\begin{aligned} P_t &= exp\{\hat{d}_{t,t}\}/exp\{r\} + exp\{\hat{d}_{t,t+1}\}/exp\{2r\} + \cdots + exp\{\hat{d}_{t,t+i}\}/exp\{(i+1)r\} + \cdots \\ & \quad V_{BIG} + V_{BIG}exp\{\gamma - r\} + V_{BIG}exp\{2(\gamma - r)\} + \cdots \end{aligned} \tag{6}$$

where $V_{BIG} = exp\{\hat{d}_{t,t+BIG}\}/exp\{(BIG+1)r\}$ and the second line of equation (6) equals $V_{BIG}/[1 - exp\{\gamma - r\}$. From experimentation setting "$BIG = 50$" years takess $\Delta\hat{d}$ close enough to γ. Values for γ and σ_u^2 used in the simulation are .015 and .03, which are obtained from the historical data. The mean first order serial correlation of returns calculated from the generated data with 116 observations was -.0087 (-.0049 for 232 observations), which is close to the expected value $(-T^{-1})$ for random data.