Predicting returns with financial ratios

Jonathan Lewellen

MIT Sloan School of Management
50 Memorial Drive E52-436, Cambridge, MA 02142
(617) 258-8408
lewellen@mit.edu

First draft: March 2000
Current version: September 2001

I thank John Campbell, Ken French, Jay Shanken, Tuomo Vuolteenaho, two anonymous referees, and workshop participants at MIT, University of Chicago, and Grantham, Mayo, Van Otterloo, & Co. for helpful comments and suggestions. Financial support from the MIT / Merrill Lynch Partnership is gratefully acknowledged.

Predicting returns with financial ratios

Abstract

This paper reports new evidence on the predictive power of dividend yield, book-to-market, and the earnings-price ratio. I show that previous studies overstate the bias in predictive regressions, and consequently, understate the forecasting power of the three financial ratios. Dividend yield predicts stock returns from 1946 – 1997, as well as in various subperiods. Book-to-market and the earnings-price ratio predict returns during the shorter 1963 – 1997 sample. The evidence remains strong despite the ratios' poor forecasting ability in recent years.

Predicting returns with financial ratios

1. Introduction

Nearly fifty years ago, Kendall (1953) observed that stock prices seem to wander randomly over time. Kendall, and the early literature on market efficiency, tested whether price changes could be predicted using past returns. Empirical tests later expanded to other predictive variables, including interest rates (Fama and Schwert, 1977; Campbell, 1987), default spreads (Keim and Stambaugh, 1986; Fama and French, 1989), dividend yield and the earnings-price (Fama and French, 1988; Campbell and Shiller, 1988a), and more recently, the book-to-market ratio (Kothari and Shanken, 1997; Pontiff and Schall, 1998).

The three financial ratios – DY, B/M, and E/P – share several common features. First, each ratio measures price relative to 'fundamentals.' Because price is low when expected returns are high, and vice versa, the ratios should fluctuate positively with expected returns. According to the mispricing view, prices are low when investors are too pessimistic; the ratios predict high returns because stocks are underpriced. The efficient-markets view argues, instead, that prices are low when discount rates are high; the ratios predict returns because they track time-variation in the risk premium. The financial ratios also share similar time-series properties. At a monthly frequency, they have autocorrelations near one and most of their movement is caused by price changes in the denominator. These statistical properties have important effects on empirical tests.

This paper re-examines the predictive power of DY, B/M, and E/P. I focus primarily on DY because it has received the most attention in the literature. I also focus exclusively on short-horizon tests – monthly returns regressed on lagged DY – to avoid the complications arising from overlapping returns. Previous research has produced little evidence that DY forecasts monthly returns. Fama and French (1988) find that DY predicts monthly NYSE returns from 1941 – 1986, with t-statistics between 2.20 and 3.21 depending on the definition of returns (equal- vs. value-weighted; real vs. nominal). However, Stambaugh (1986, 1999) and Mankiw and Shapiro (1986) show that predictive regressions can be severely biased toward finding predictability. Nelson and Kim (1993) replicate the Fama and French tests, correcting for bias using bootstrap simulations, and estimate that the p-values are actually between 0.03 and 0.33. More recently, Stambaugh (1999) derives the exact small-sample distribution of the estimates (assuming that DY follows an AR(1) process). He reports a one-sided p-value of 0.15 when NYSE returns are regressed on lagged DY from 1952 – 1996.[1]

In this paper, I show that Stambaugh's (1999) analysis, as well as the Monte Carlo simulations common in the literature, can substantially overstate the bias in predictive regressions. Stambaugh's results are based on repeated sampling of both returns and DY, and they implicitly ignore information conveyed by the sample autocorrelation of DY. The slope estimate in a predictive regression and the sample autocorrelation of DY are strongly correlated, so any information conveyed by the autocorrelation helps produce a more powerful test of predictability. Incorporating this information into empirical tests has two effects: (1) the bias in the predictive slope coefficient is often smaller than Stambaugh's estimate; and (2) the standard error of the slope coefficient is much lower. In combination, the two effects substantially raise the power of empirical tests.

To gain some intuition, consider the model of returns analyzed by Stambaugh (1986, 1999) and Mankiw and Shapiro (1986):

$$r_t = \alpha + \beta\, x_{t-1} + \varepsilon_t, \tag{1a}$$

$$x_t = \phi + \rho\, x_{t-1} + \mu_t, \tag{1b}$$

where r_t is the return during month t and x_{t-1} is the dividend yield at the beginning of the month. The first equation is the predictive regression and the second equation specifies an AR(1) process for DY. The residuals, ε_t and μ_t, are correlated because positive returns lead to a decrease in DY. As a consequence, estimation errors in the two equations are closely connected:

$$\hat{\beta} - \beta = \gamma(\hat{\rho} - \rho) + \eta\,, \tag{2}$$

where η is a random error with mean zero and γ is a negative constant. This equation shows that $\hat{\beta}$

[1] DY predicts long-horizon returns more strongly, but the statistical significance is sensitive to changes in the definition of returns and the time period considered. See, for example, Hodrick (1992), Goetzmann and Jorion (1993), and Nelson and Kim (1993).

inherits many properties of sample autocorrelations. For example, taking expectations, the downward bias in $\hat{\rho}$ (Kendall, 1954) induces an upward bias in $\hat{\beta}$. That observation underlies Stambaugh's analysis and the simulations in other studies. Notice, however, that taking expectations throws out the information contained in the observed $\hat{\rho}$. The sample autocorrelation of DY is roughly 0.99 at the monthly frequency. Assuming DY is stationary, so that ρ is at most one, we know immediately that $\hat{\rho} - \rho > -0.01$. In turn, eq. (2) implies that the 'bias' in $\hat{\beta}$ is at most -0.01γ. This upper bound is often less than the bias derived by Stambaugh. In effect, the repeated-sampling inherent in Stambaugh's analysis includes many samples in which $\hat{\rho}$ is far below ρ, and we can eliminate these samples given the observed autocorrelation.

Empirically, exploiting the information in $\hat{\rho}$ strengthens the case for predictability. When NYSE value-weighted returns are regressed on log DY from 1946 – 1997, the OLS slope estimate is 1.21 with a standard error of 0.57. Based on Stambaugh's (1999) small-sample distribution, the bias-adjusted estimate is 0.51 with a one-sided p-value of 0.210. However, conditioning on DY's sample autocorrelation, the bias-adjusted estimate becomes 0.72 with a t-statistic of 4.09. This t-statistic has the standard distribution, so DY is strongly significant. We can also reject the null in shorter subperiods. For the first half of the sample, 1946 – 1971, the bias-adjusted estimate is 0.81 with a p-value less than 0.002. For the second half of the sample, 1972 – 1997, the bias-adjusted estimate is 0.72 with a p-value again less than 0.001. In short, by recognizing the upper bound on ρ, we obtain much more powerful tests of predictability.

As an aside, I also consider how the last few years of data affect the empirical results. In May 1995, DY reached a new low for the sample, predicting that returns going forward should be far below average. Contrary to the model, the NYSE index nearly doubled over the subsequent three years. When returns for 1995 – 1997 are added to the regression, the OLS slope coefficient drops almost in half, from 2.19 to 1.21, and the statistical significance declines from 0.061 to 0.210 using Stambaugh's small-sample distribution. Interestingly, the tests here are not sensitive to the recent data. The bias-adjusted slope drops from 0.95 to 0.72 and the p-value remains below 0.001. The reason is simple: the last few years have also lead to a sharp rise in the sample autocorrelation of DY, from 0.987 to 0.994. This rise means that the maximum bias in the predictive slope declines from 1.24 to 0.49, offsetting most of the decline in the OLS estimate. Regressions with the equal-weighted index are even more remarkable, finding stronger evidence of predictability after observing the recent data.

I also find that B/M and E/P have significant predictive power. The tests with B/M and E/P are restricted to 1963 – 1997 because of data requirements. Using the value-weighted NYSE index, B/M is positively related to expected returns in both the full sample (p-value of 0.062) and the truncated sample ending in 1994 (p-value of 0.012). Similarly, E/P predicts returns with p-values of 0.053 and 0.032 for the two periods. Again, these results are much stronger than previous studies. Kothari and Shanken (1997) and Pontiff and Schall (1998) find that B/M has little predictive power during this period, and Lamont (1999) finds no evidence that E/P (by itself) predicts quarterly returns from 1947 – 1994. I should mention, however, that the results for B/M and E/P are sensitive to the definition of returns; B/M and E/P forecast nominal returns on NYSE equal- and value-weighted indices, but they have little power to predict excess returns.

2. Predictive regressions

Predictive regressions are ubiquitous in the finance literature. They have been used to test whether past prices, financial ratios, interest rates, and a variety of other macroeconomic variables can forecast stock and bond returns. This section reviews the properties of predictive regressions, borrowing liberally from Stambaugh (1986, 1999).

2.1. The return-generating process

The paper focuses on the regression

$$r_t = \alpha + \beta\, x_{t-1} + \varepsilon_t, \tag{3a}$$

where r_t is the return during month t and x_{t-1} is a predictive variable known at the beginning of the month. It is easy to show that β must be zero if expected returns are constant. In many cases, including all of the ones discussed here, the alternative hypothesis is that $\beta > 0$, so we will be concerned only with one-sided tests. To complete the model, assume that x_t follows a stationary AR(1) process:

$$x_t = \phi + \rho\, x_{t-1} + \mu_t, \tag{3b}$$

where $-1 < \rho < 1$. An increase in stock prices leads to a decrease in DY, so the residuals in (3a) and (3b) are negatively correlated. Therefore, ε_t and x_t are correlated in the predictive regression, violating one of the assumptions of OLS (which requires independence at all leads and lags). For simplicity, I assume that all variables are normally distributed.

Before continuing, I should briefly discuss the stationarity assumption. The empirical tests depend on the assumption that ρ cannot be greater than one. Statistically, the tests actually remain valid if $\rho = 1$ (we just need an upper bound on ρ). I often assume that ρ is strictly less than one to be consistent with prior studies. It also makes little sense to predict returns with a nonstationary variable. Economically, x_t should be stationary unless there is an explosive bubble in stock prices. Suppose, for example, that x_t equals log DY. x_t will be stationary if log dividends and prices are cointegrated, implying that deviations from 'fundamental' value are not expected to grow forever. That assumption seems reasonable. There is a vast literature that argues against explosive bubbles and much direct evidence that DY is stationary using data over long sample periods.[2]

2.2. Properties of OLS

Denote the sample matrix of regressors as X, the coefficient vectors as $b = (\alpha, \beta)$ and $p = (\phi, \rho)$, and the residual vectors as ε and μ. The OLS estimates of eqs. (3a) and (3b) are then

$$\hat{b} = b + (X'X)^{-1}X'\varepsilon\,, \tag{4a}$$

$$\hat{p} = p + (X'X)^{-1}X'\mu\,. \tag{4b}$$

In the usual OLS setting, the estimation errors are expected to be zero. That is not true here. It is well-

[2] See Blanchard and Watson (1982), Tirole (1982, 1985), and Loewenstein and Willard (1999) for theoretical arguments and Hamilton and Whiteman (1985), Flood and Hodrick (1986, 1990), Diba and Grossman (1988), and West (1988) for empirical evidence.

known that autocorrelations are biased downward in finite samples, and this bias feeds into the predictive regression through the correlation between ε_t and μ_t. Specifically, note that we can write $\varepsilon_t = \gamma \mu_t + \nu_t$, where $\gamma = \sigma_{\varepsilon\mu} / \sigma_\mu^2$. Given our earlier assumptions, ν_t is independent of both μ_t and X. Substituting into (4a) yields

$$\hat{b} - b = \gamma(\hat{p} - p) + \eta, \tag{5}$$

where $\eta \equiv (X'X)^{-1}X'\nu$ has mean zero and variance $\sigma_\nu^2 (X'X)^{-1}$. This equation provides a convenient way to think about predictive regressions.

Consider, first, the distribution of $\hat{\beta}$ based on repeated sampling of both $\hat{\rho}$ and η. This is the distribution studied by Stambaugh. Eq. (5) shows that $\hat{\beta}$ inherits many of the statistical properties of autocorrelations. For example, taking expectations yields

$$E[\hat{\beta} - \beta] = \gamma\, E[\hat{\rho} - \rho]. \tag{6}$$

The downward bias in $\hat{\rho}$, approximately equal to $-(1+3\rho)/T$, induces an upward bias in the predictive slope. Further, autocorrelations are negatively skewed and more variable than suggested by OLS. These properties imply that $\hat{\beta}$ is positively skewed and also more variable than suggested by OLS. Stambaugh discusses these properties in detail.

Eq. (5) tells us more directly about the 'conditional' distribution of $\hat{\beta}$, where conditional means the distribution given $\hat{\rho}$ (and X). The correlation between $\hat{\beta}$ and $\hat{\rho}$ is strong, so information about $\hat{\rho} - \rho$ provides a lot of information about the predictive slope. The conditional expectation of $\hat{\beta}$ is

$$E_{\hat{\rho}}[\hat{\beta} - \beta] = \gamma(\hat{\rho} - \rho). \tag{7}$$

I refer to $\gamma(\hat{\rho} - \rho)$ as the realized bias in $\hat{\beta}$. The strong correlation between $\hat{\beta}$ and $\hat{\rho}$ implies that the conditional variance of $\hat{\beta}$ is substantially less than the unconditional variance. Also, $\hat{\beta}$ is normally distributed conditional on $\hat{\rho}$ and X. The irregularities in the sampling distribution of $\hat{\beta}$ are caused by its correlation with $\hat{\rho}$, not by the independent component.

The tests in this paper are based on the conditional distribution of $\hat{\beta}$. The idea is simple. Even though we do not know $\hat{\rho} - \rho$, we can put a lower bound on it by assuming $\rho \approx 1$. In turn, this assumption gives us an upper bound on the bias in $\hat{\beta}$. Define the bias-adjusted estimator:

$$\hat{\beta}_{adj} = \hat{\beta} - \gamma(\hat{\rho} - \rho) . \tag{8}$$

This estimator is normally distributed with mean β and variance $\sigma_v^2(X'X)^{-1}$. I will show later that the parameter γ can be estimated very precisely and, therefore, treated as known. Although we do not know the true autocorrelation, $\rho \approx 1$ is the most conservative assumption we can make for testing predictability: the bias in eq. (7) is maximized, and the estimator in eq. (8) is minimized, if we assume $\rho \approx 1$. If $\hat{\beta}$ is significantly different from zero given this assumption, then it must be even more significant given the true value of ρ.

Previous studies focus on the unconditional distribution of $\hat{\beta}$. Implicitly, their tests assume that we have no information about $\hat{\rho} - \rho$ other than its unconditional distribution. That assumption will be true in many cases: given the sample autocorrelation of DY, we have no way of knowing whether it is above or below ρ. However, when the sample autocorrelation is close to one, the unconditional distribution throws out useful information. Suppose, for example, that $\hat{\rho}$ = 0.99 and T = 300. In this case, the unconditional bias in $\hat{\rho}$ is approximately E[$\hat{\rho} - \rho$] = –0.016. However, given the observed sample autocorrelation, the *minimum* possible value of $\hat{\rho} - \rho$ is actually –0.010. It follows that the unconditional distribution overstates the bias in $\hat{\beta}$ by at least 60%.

Figure 1 illustrates these ideas. Panel A plots the unconditional distribution of $\hat{\beta}$ and Panel B plots the conditional distribution given two different values of $\hat{\rho}$, one above and one below the true ρ. For the simulations, β = 0, ρ = 0.99, the correlation between ε_t and μ_t is –0.92, and T = 300. (The correlation between ε_t and μ_t is similar to its empirical value.)

Panel A approximates Stambaugh's small-sample distribution. It clearly shows the strong bias and skewness in $\hat{\beta}$: the mean is 0.30, the median is 0.24, and the skewness is 1.29. In contrast, the conditional distributions in Panel B are symmetric and approximately normal.[3] Comparing the two panels, it is clear that high values of $\hat{\beta}$ must correspond to samples in which $\hat{\rho}$ is far below its true value. For example, in Panel A, $\hat{\beta}$ is greater than 0.40 roughly 30% of the time. However, if we eliminate samples with low $\hat{\rho}$ (more than 0.0075 below the true value), Panel B shows that $\hat{\beta}$ will rarely be that large.

The discussion here should not be interpreted as a criticism of Stambaugh's (1999) analysis. In general, his small-sample distribution will give the best estimate of β. It will only be misleading when the sample autocorrelation is close to one. Also, Stambaugh actually advocates Bayesian methods. In some of his Bayesian analysis, he imposes the constraint that $\rho < 1$ and recognizes that this can lead to stronger rejections of the null. My contribution is to show that the constraint can also improve inferences in a frequentist setting. In fact, the approach here is similar in many ways to Stambaugh's Bayesian analysis, in that both condition on observed data in deriving the distribution of β (or $\hat{\beta}$ here). The tests are essentially identical if the Bayesian approach starts with a perfectly informative prior that $\rho \approx 1$ (but no information about β). Any other prior that places zero weight on $\rho > 1$ would produce even stronger rejections of the null.[4]

3. Data and descriptive statistics

I use the methodology outlined above to test whether DY, B/M, and E/P forecast stock returns. Return, market value, and dividend data come from the Center for Research in Security Prices (CRSP).

[3] The distributions in Panel B condition only on the sample autocorrelation of DY, not on the full matrix X. The mathematical analysis conditioned on both, so the graphs are not exactly comparable to the discussion above. The variances of the two conditional distributions differ because the sample autocorrelation of DY is correlated with the sample variance.

[4] Several readers have questioned whether the tests in this paper are truly frequentist. They are. The tests are simply based on the repeated-sampling distribution of $\hat{\beta}_{adj} = \hat{\beta} - \hat{\gamma}(\hat{\rho} - 1)$. (We do not know the exact distribution of this statistic, because it depends on ρ, but we can put an upper bound on the p-value.) The test can also be interpreted as a joint test of the hypothesis that $\beta \leq 0$ and $\rho \leq 1$.

Earnings and book equity come from Compustat. The tests focus on NYSE equal- and value-weighted indices, primarily to be consistent with prior research. Also, the composition of the overall CRSP index changes considerably over time as AMEX and NASDAQ firms enter the database, which would complicate empirical tests.

DY is calculated monthly on the value-weighted NYSE index. It is defined as dividends paid on the index over the prior year divided by the current level of the index. Thus, DY is based on a rolling window of annual dividends. Similar to the approach of Fama and French (1988), DY is actually backed out of the 'with dividend' and 'without dividend' returns provided by CRSP (see their paper for details). I use value-weighted DY to predict returns on both the equal- and value-weighted indices, primarily because it is easier to compare slope coefficients across regressions when the independent variables are the same. The predictive regressions use the natural log of DY, rather than the raw series, for reasons discussed below.

The empirical tests with DY use returns from January 1946 – December 1997. I omit the Depression era because the properties of returns and DY are much different before and after 1945. Returns were extremely volatile in the 1930s and this volatility is reflected in both the variance and persistence of DY (see Fama and French, 1988). As a robustness check, I split the sample in half and look at the two subperiods, 1946 – 1971 and 1972 – 1997. Further, I investigate the influence of the last few years because recent stock returns have been so unusual.

The tests with B/M and E/P are restricted to 1963 – 1997 when Compustat data is available. B/M is calculated as the ratio of book equity in the previous fiscal year to market equity in the previous month. E/P is measured as the ratio of operating earnings before depreciation to market value. I use operating earnings because Shiller (1984) and Fama and French (1988) suggest that net income is a noisy measure of fundamentals; preliminary tests suggest that operating earnings are a better measure.[5] To ensure that

[5] Log E/P ratios are highly autocorrelated using either measure: 0.990 for operating earnings and 0.987 for net income. However, the residuals in an AR(1) regression are more variable when E/P is calculated from net income (standard deviation of 0.068 for net income compared with 0.049 for operating earnings). In addition, the residuals are less highly correlated with returns, –0.60 vs. –0.85. Net income seems to vary independent of price movements more than operating earnings, which might indicate additional noise in the process. In any case, the predictive

the tests are predictive, I do not update accounting numbers until four months after the fiscal year. Also, to reduce possible selection biases, a firm must have three years of accounting data before it is included in the sample (see Kothari, Shanken, and Sloan, 1995). The regressions use log B/M and log E/P, both measured on the value-weighted NYSE index.

Table 1 provides summary statistics for the data. The table shows that volatility is bit lower in the first half of the sample, but most of the time-series properties of returns seem to be stable across the two subperiods. The same observation applies to DY. DY averages 3.92% over the full sample and tends to be positively skewed, which is not terribly surprising since it is measured as a ratio. In principal, log DY should better approximate a normal distribution. The table confirms that log DY is more symmetric in the full sample; the skewness is somewhat positive in the first half of the sample and negative in the second half. The properties of B/M and E/P are similar to those of DY. The B/M ratio averages 0.56 and the E/P ratio averages 0.21. The raw series for both B/M and E/P are positively skewed, while the log series are nearly symmetric. (The average E/P ratio is relatively high because earnings are measured as operating earnings before depreciation. In comparison, the E/P ratio based on net income averages 0.07 over the period.)

Table 1 also shows that the financial ratios are extremely persistent. The first-order autocorrelations range from 0.986 to 0.996 for the various series. The autocorrelations diminish as the lag increases, consistent with a stationary process. (Nelson and Plosser, 1982, point out that the same pattern would be expected even with a unit root.) The table shows that log DY, B/M, and E/P tend to be a bit more highly autocorrelated than the raw series. This is important for the empirical tests because the maximum bias depends on $1 - \hat{\rho}$.

4. Empirical results

I estimate predictive regressions using DY, B/M, and E/P in the full sample and various subperiods.

power of the two series is similar and, for simplicity, I report only the tests using operating earnings (the results for net income are marginally weaker).

The influence of the last few years receives special attention because it highlights important properties of the conditional tests. The tests initially focus on DY since it has received the most attention in the literature.

4.1. Predicting with dividend yield

Table 2 explores the predictive power of DY in the full sample, 1946 – 1997 (624 months). The table reports estimates of the model analyzed in Section 2:

$$r_t = \alpha + \beta\, x_{t-1} + \varepsilon_t, \tag{9a}$$

$$x_t = \phi + \rho\, x_{t-1} + \mu_t, \tag{9b}$$

where r_t is the return during month t and x_{t-1} is the dividend yield at the beginning of the month. As mentioned earlier, all of the predictive regressions use DY calculated for the value-weighted NYSE index. I estimate predictive regressions for NYSE equal- and value-weighted returns and for nominal and excess returns (measured net of the one-month Tbill rate).

The table reports a variety of statistics. The row labeled 'OLS' shows the least-squares slope and standard error, along with the corresponding p-value. These estimates are reported primarily as a reference point. The row labeled 'Stambaugh' reports estimates based on Stambaugh's (1999) small-sample distribution. The slope coefficient is the OLS estimate minus the unconditional bias and the p-value is based on the unconditional distribution. Although Stambaugh derives the exact distribution of the slope estimate, the moments of the distribution are difficult to evaluate analytically. Therefore, the values in Table 2 are actually based on Monte Carlo simulations. The distribution of $\hat{\beta}$ depends on the unknown parameters ρ and Σ, for which I substitute the OLS estimates. Stambaugh notes that the distribution is relatively insensitive to small changes in the parameters, so this substitution should not be too important.

The final row, labeled '$\rho \approx 1$,' reports estimates based on the conditional distribution of $\hat{\beta}$. I assume that the true autocorrelation is approximately one (operationalized as $\rho = 0.9999$). The slope coefficient is the bias-adjusted estimator

$$\hat{\beta}_{adj} = \hat{\beta} - \gamma(\hat{\rho} - \rho)\,, \qquad (10)$$

with $\rho \approx 1$. Recall from Section 2 that $\gamma = \sigma_{\varepsilon\mu} / \sigma_{\mu}^2$. Again, this parameter is required to compute the bias-adjusted estimate, and I substitute the OLS estimate for the true value.[6] The conditional variance of $\hat{\beta}$ equals $\sigma_v^2 (X'X)^{-1}$ and a t-statistic can be calculated in the standard way. Under the null, this t-statistic is truly a t-statistic – that is, it has a Student t distribution with T – 2 degrees of freedom. Therefore, the p-value for '$\rho \approx 1$' can be calculated directly without simulation. This is one of the attractive features of the conditional tests.

Table 2 provides strong evidence of predictability. Consider, first, the nominal return on the value-weighted index. The OLS slope estimate is 1.21 with a standard error of 0.57. Since the standard deviation of DY is 0.28, the point estimate implies that a one-standard-deviation change in DY is associated with a 0.34% (1.21×0.28) change in monthly expected return. Of course, the slope coefficient is biased upward. Based on Stambaugh's (1999) distribution, the bias-adjusted estimate is 0.51 with a one-sided p-value of 0.210. The conditional test, in the third row, gives a much different picture. Assuming that $\rho \approx 1$, the maximum bias in the predictive regression is only 0.49, or approximately 30% lower than the unconditional bias. The bias-adjusted slope is 0.72 and the conditional test rejects the null hypothesis at the 0.000 level. Notice that the conditional standard error of $\hat{\beta}$ is much lower than the standard error estimated from Stambaugh's distribution, 0.18 vs. 0.75.[7] The strong significance in the conditional test is largely attributed to this difference.

These results show that the small-sample distribution analyzed by Stambaugh (1986, 1999) and Mankiw and Shapiro (1986) can greatly understate the significance of DY. Their tests ignore the

---

[6] The appendix shows that estimation error in γ is easily incorporated into the tests. It has little effect on the results because γ can be estimated very precisely. For example, γ equals –88.03 for nominal VWNY; simulations suggest that $\hat{\gamma}$ is unbiased and has a standard error of 1.16. The appendix also shows that $\hat{\gamma}$ does not depend on the assumed value of ρ.

[7] Note that the standard error is relevant only for testing a lower bound on the predictive slope (i.e., for testing the null of no predictability). The standard error on the upside is higher because I only impose a one-sided bound on the sample autocorrelation. Thus, β might be substantially larger than suggested by a symmetric confidence interval around the conditional estimate.

information contained in the observed autocorrelation of DY, which inflates both the bias and the standard error of the slope.[8] Surprisingly, the conditional tests in Table 2 actually find that DY is more significant than suggested by OLS. Although the bias-adjusted slope is substantially lower than the OLS estimate, the additional information conveyed by $\hat{\rho}$ has an even greater effect on the standard error. Indeed, for the regressions in Table 2, we would do better by ignoring small-sample biases than by using the unconditional distribution; the stochastic properties of DY actually strengthen the case for predictability.

The regressions for equal-weighted NYSE returns confirm these findings. The bias-adjusted slope for EWNY, 1.05 (p-value of 0.001), is larger than the estimate for VWNY. The standard error of the estimate is also higher, 0.33, primarily because equal-weighted returns are not as highly correlated with shocks to DY. The table also shows that nominal and excess returns produce very similar estimates. Also, in all regressions, the explanatory power of DY is low as measured by the adjusted R^2. Thus, while DY forecasts significant time-variation in expected returns, it can only explain a small fraction of total variability.

Table 3 reports results for the first and second halves of the sample, 1946 – 1971 and 1972 – 1997 (each 312 months). Even with a fairly short sample, the tests strongly reject the null in many cases. For 1946 – 1971, DY predicts value-weighted, but not equal-weighted, returns. Focusing on the value-weighted index, the bias-adjusted estimates for nominal and excess returns, 0.81 and 1.12 (p-values of 0.002 and 0.000), are similar to the full sample. In the second half of the sample, DY predicts nominal returns on both indices, as well as excess returns on the equal-weighted index (p-values less than 0.001). The slope coefficients are especially large on the equal-weighted index in this period. For example, the bias-adjusted slope for excess EWNY, 1.51, implies that a one-standard-deviation change in dividend

[8] I should mention that conditioning on the sample autocorrelation would not have helped in Stambaugh's paper, primarily because DY is not as highly autocorrelated in his sample. There are two reasons: (1) Stambaugh uses raw DY, which is slightly less persistent than log DY, and (2) DY is not as highly autocorrelated during his sample periods (e.g., 1926 – 1996). I have repeated the tests in Table 2 using raw DY, or using Stambaugh's post-War sample of 1952 – 1996, and find that the conditional tests strongly reject the null in both cases. However, using raw DY *and* 1952 – 1996, the conditional tests reject only for nominal returns. During this period, raw DY has little power to predict excess returns using either the conditional or unconditional distribution.

yield is associated with a 0.44% increase in monthly expected return.

The subperiod results reveal striking differences between the conditional and unconditional tests. Consider, for example, the VWNY regressions in the second half of the sample. The bias-adjusted estimate from the unconditional tests is –0.50 with a standard error of 1.38. In contrast, the estimate from the conditional tests is significantly positive, 0.72, with a standard error of 0.23. The p-values for the two tests are 0.577 and 0.001, respectively. Thus, incorporating the information is $\hat{\rho}$ can be critical when the autocorrelation is close to one and the sample is relatively short. The AR(1) regressions for DY, at the top of the table, show why. The expected bias in $\hat{\rho}$ is approximately –0.017 while the realized 'bias' is at most –0.004. As a consequence, the maximum bias in the predictive slope is nearly 75% lower than the unconditional estimate.

I also emphasize that the conditional tests are quite conservative. The assumption that $\rho \approx 1$ is the weakest assumption we can make for testing the null. Although it is difficult to justify any other upper bound on ρ, it might be useful to evaluate alternative values. Consider, for example, raw EWNY from 1946 – 1971. The p-value drops from 0.335 if we assume $\rho \approx 1$ to 0.008 if we assume instead that $\rho =$ 0.991 (the sample autocorrelation). The p-value would be less than 0.050 for any $\rho < 0.994$. For excess EWNY, the reported p-value is 0.124, but this would drop below 5% for any $\rho < 0.998$. I am not suggesting that DY is truly statistical significant in these regressions, since there is no rigorous justification for imposing a bound on ρ other than one, but the reported p-values do appear to be quite conservative.

Bayesian methods provide a more rigorous way to incorporate uncertainty about ρ.[9] In particular, suppose an investor begins with no information about β, and consider three different beliefs about ρ. If the investor believes ρ equals one with certainty, the conditional tests reported in the table equal the posterior probability that $\beta \leq 0$. If, instead, the investor begins with a flat prior for $\rho \leq 1$, the posterior

[9] The Bayesian posterior probability can be found by integrating the conditional p-value (conditional on a given autocorrelation) over the posterior distribution of ρ. The analysis here is similar to Stambaugh (1999).

probability (for $\beta \leq 0$) drops to 0.041 for nominal EWNY and 0.011 for excess EWNY (compared with 0.335 and 0.124, respectively, if the investor believes $\rho = 1$). Finally, suppose we arbitrarily shift the investor's posterior belief about ρ upward by one standard deviation and truncate at one. This posterior represents fairly strong beliefs that ρ is close to one. It is roughly the lower tail of a normal distribution with mean one and standard deviation 0.008 (the standard error of $\hat{\rho}$). In this case, the posterior probability equals 0.092 for nominal EWNY and 0.027 for excess EWNY. Once again, the p-values in Table 3 seem to be quite conservative.

4.2. The influence of 1995 – 1997

The tests in Tables 2 and 3 include data for 1995 – 1997. Returns during these years moved opposite to the predictions of the model: DY was extremely low, but the NYSE indices performed far above average. For example, DY at the end of 1994 was 2.9% and dropped to 1.7% by the end of 1997. Over the same period, the value-weighted NYSE index returned 118%. Because this period is so unusual, I briefly consider its effect on the results.

Table 4 shows regressions for both the full sample (the same as Table 2) and the truncated sample 1946 – 1994. I report only regressions with nominal returns for simplicity. Focusing on the value-weighted index, the OLS slope coefficient in the truncated sample is 82% greater than it is for the entire period, 2.19 compared with 1.21. The slope coefficient is significant at the 0.061 level of the truncated sample, but only at the 0.210 level for the full regression, using the small-sample distribution of Stambaugh (1999). Interestingly, however, the estimate does not change much in the conditional tests. The bias-adjusted slope drops from 0.95 to 0.72 and remains significant at the 0.000 level. The t-statistic only drops from 4.65 to 4.09. Thus, the statistical significance of DY remains strong even though the model has performed terribly from 1995 – 1997.

The relative insensitivity of the conditional tests is explained by the sample autocorrelation of DY. Table 4 shows that the sample autocorrelation of DY increases from 0.986 to 0.994. The increase in $\hat{\rho}$ means that the sampling error in ρ must have gone up (become less negative or more positive). The conditional tests implicitly recognize that the sampling error in β has correspondingly decreased. Although the OLS estimate declines by 0.99, the conditional test attributes 76% of the drop to sampling error, which shows up in a much lower estimate of the bias.

Table 4 also reports a remarkable result: the bias-adjusted slope coefficient for EWNY does not change with the addition of 1995 – 1997 and the statistical significance actually increases. Again, this counterintuitive result can be explained by the sharp increase in the sample autocorrelation of DY. Given the sharp drop in DY from 1995 – 1997, and the associated increase in $\hat{\rho}$, we would *expect* to see contemporaneously high returns.

The truncated sample is interesting for another reason. The unconditional bias-adjusted slopes are higher than the conditional estimates, but their statistical significance is much lower. Focusing on value-weighted returns, the unconditional bias-adjusted slope is 1.53 with a p-value of 0.061; the conditional bias-adjusted slope is 0.95 with a p-value of 0.000. This combination is a bit awkward. The higher unconditional estimate suggests that the conditional tests are too convervative. The true autocorrelation is probably lower than one, which would imply that the conditional estimates of β are too low. But without the extreme assumption, we cannot reject the null as strongly. This finding points out an odd property of the tests, and it suggests one advantage of a Bayesian approach.

4.3.Predicting with book-to-market and earnings-price

Table 5 explores the predictive power of B/M. I report predictive regressions both for the full sample, 1963 – 1997, and the truncated sample ending in 1994. The regressions suggest that B/M forecasts nominal, but not excess, returns. Focusing on nominal returns in the full sample, the OLS slope coefficient is 0.89 for the value-weighted index and 1.61 for the equal-weighted index. The bias-adjusted slopes, 0.46 and 1.13, are significant with p-values of 0.062 and 0.008, respectively. In the truncated sample, the statistical significance is similar but the slope estimates are a bit larger. The bias-adjusted estimate is 0.77 for VWNY (p-value, 0.012) and 1.20 for EWNY (p-value, 0.014). The evidence for excess returns is uniformly weaker, with marginal significance (0.077 p-value) only for excess EWNY in the full sample.

Like the results for DY, Table 5 shows that conditional tests provide much stronger evidence of predictability than unconditional tests. The differences are dramatic for nominal returns in the full sample. Incorporating the information in $\hat{\rho}$, the p-value for nominal VWNY drops from 0.476 to 0.062 and the p-value for EWNY drops from 0.288 to 0.008. For both indices, the conditional bias is less than half the unconditional bias. The difference between the two tests is also revealed by comparing the full and truncated samples. The regressions for excess EWNY are especially interesting. The OLS slope estimate drops from 1.93 for the sample ending in 1994 to 1.16 for the sample ending in 1997. However, the conditional bias-adjusted slope actually *increases* slightly, from 0.65 to 0.67. That result mirrors the evidence for DY in Table 4.

Table 6 replicates the tests using E/P. The results are remarkably similar to those of B/M, both qualitatively and quantitatively. E/P appears to forecast nominal returns, but again there is little evidence that it forecasts excess returns. The p-values for nominal returns range from 0.023 to 0.053 for the different time periods and stock returns. The bias-adjusted slope coefficients are relatively large, implying economically meaningful changes in expected returns (a one-standard-deviation increase in E/P maps into a 0.19% increase in expected return for VWNY and a 0.30% increase for EWNY). Table 6 also confirms the influence of 1995 – 1997 on the regressions. The OLS slope estimate declines when the final three years are included in the regressions, but the conditional bias-adjusted slope remains approximately the same. For both nominal and excess EWNY, the addition of 1995 – 1997 strengthens the case for predictability.

5. Summary and conclusions

The literature on stock return predictability has evolved considerably over the last twenty years. Empirical tests initially produced strong evidence that returns are predictable, especially over long horizons. Research later showed that predictive regressions are subject to small-sample biases. Correcting for bias weakens, and often reverses, conclusions about predictability. The accumulated evidence suggests that DY does not predict monthly returns, although it may predict annual and multiple-year returns. Existing results for B/M and E/P are similar.

This paper presents new evidence on the predictive power of the three financial ratios. The paper makes four main points:

(a) Stambaugh (1999) and Mankiw and Shapiro (1986) consider the ‘unconditional’ distribution of $\hat{\beta}$. Unlike standard OLS, this distribution does not condition on the observed regressors. Most importantly, the tests do not condition on the sample autocorrelation of the predictive variable, which throws out useful information when the autocorrelation is close to one. The unconditional distribution can substantially understate the significance of DY, B/M, and E/P.

(b) The conditional tests in this paper are intuitive. If we know ρ and γ, then the best estimate of β is the bias-adjusted estimator $\hat{\beta}_{adj} = \hat{\beta} - \gamma(\hat{\rho} - \rho)$. This estimator is normally distributed with mean β and variance $\sigma_v^2 (X'X)^{-1}$. The parameter γ can be estimated very precisely from the data and, consequently, treated as known. Although we do not know true autocorrelation, $\rho \approx 1$ is the most conservative assumption we can make because it yields the lowest estimate. The conditional tests are also easy to apply: all of the necessary statistics can be estimated from OLS and, under the null, the test statistic has a Student t distribution. When $\hat{\rho}$ is close to one, the conditional bias-adjusted slope can be higher than the unconditional estimate. Further, when $\hat{\beta}$ and $\hat{\rho}$ are highly correlated, the conditional variance is much lower than the unconditional variance. Both of these effects produce more powerful tests of predictability.

(c) Empirically, incorporating the information $\hat{\rho}$ can be quite important. I find strong evidence that DY predicts returns. The tests examine NYSE equal- and value-weighted indices over the period 1946 – 1997. In the full sample and various subsamples, DY is typically significant at the 0.001 level, with many t-statistics greater than 3.0 or 4.0. The evidence for B/M and E/P ratios is weaker than for DY, but stronger than previous studies. B/M and E/P appear to forecast nominal returns on both the equal- and value-weighted indices, but not excess returns. Even when the statistics cannot reject the null, the conditional bias-adjusted slopes look much different than the unconditional estimates.

(d) The last few years of the sample, 1995 – 1997, have a large impact on the results. Adding these years to the VWNY regressions reduces the OLS slope on DY by 45%, the slope on B/M by 53%, and the slope on E/P by 21%. However, the bias-adjusted estimates are much less sensitive to the recent data, and the statistical significance of the three variables remains strong. Remarkably, the bias-adjusted estimates for EWNY actually increase with the addition of 1995 – 1997, even though returns over this period move strongly against the model. These results reveal an important property of the conditional tests: the tests recognize that changes in the sample autocorrelation of the financial ratios convey a lot of information about the predictive slope.

The conditional tests presented in this paper are based on a frequentist approach. The methodology is similar to the Bayesian tests of Stambaugh (1999). Both approaches condition on the observed sample when deriving the distribution of β (in Bayesian tests) or $\hat{\beta}$ (in frequentist tests). The main advantage of the frequentist approach is that it assumes only that ρ is less than one; a Bayesian approach requires additional assumptions about investors' beliefs. On the other hand, the Bayesian approach is applicable even when the sample autocorrelation is not close to one.

The conditional tests show that information about the sampling error in ρ can be important. The only information we used here is the stationarity of the predictive variable. The approach could be generalized to incorporate additional information about ρ into the tests. For example, suppose we have a sample of DY beginning prior to the period in which we want to test for predictability. If the autocorrelation is constant over the entire history, we can use the earlier data to help us infer whether the within-sample autocorrelation is above or below the true value. This approach, similar to Stambaugh's (1997) analysis of returns with differing histories, could be used when the sample autocorrelation is not close to one.

Appendix

Estimation error in γ is easily incorporated into the conditional tests. Estimating γ absorbs one degree-of-freedom and slightly increases the standard error of the estimate. To show this, we can think about estimating β in a slightly different way than presented in the text. In particular, re-write the predictive regression using $\varepsilon_t = \gamma\,\mu_t + \nu_t$:

$$r_t = \alpha + \beta\, x_{t-1} + \gamma\,\mu_t + \nu_t. \tag{A.1}$$

Given a value of ρ, this equation can be estimated because μ_t is observable.[10] I show below that the estimate of β from (A.1) is identical to the bias-adjusted estimator in eq. (10). That observation immediately provides the sampling distribution of $\hat{\beta}_{adj}$: eq. (A.1) satisfies OLS since $E[\nu_t \mid x_{t-1}, \mu_t] = 0$, so the estimate of β from (A.1) has all the usual properties.

From regression analysis, the multiple-regression estimate of β from (A.1), denoted $\hat{\beta}^M$, and the simple-regression estimate from the predictive regression, denoted $\hat{\beta}^S$, are related as follows:

$$\hat{\beta}^M = \hat{\beta}^S - \hat{\gamma}\ \lambda, \tag{A.2}$$

where λ is the slope coefficient in an auxiliary regression of μ_t on x_{t-1}:

$$\mu_t = c + \lambda\, x_{t-1} + \omega_t. \tag{A.3}$$

λ is the second element of $(X'X)^{-1}X'\mu$. Comparing this to eq. (4b) in the paper, we find that $\lambda = \hat{\rho} - \rho$. Substituting into (A.2) yields

$$\hat{\beta}^M = \hat{\beta}^S - \hat{\gamma}(\hat{\rho} - \rho). \tag{A.4}$$

Also, the estimate of γ from (A.1) is the same as the estimate of γ from a regression of $\hat{\varepsilon}_t$ (the sample residuals from the predictive regression) on $\hat{\mu}_t$ (the sample residuals from an AR1 model for DY). Therefore, $\hat{\beta}^M$ is identical to the bias-adjusted estimator in eq. (10). These results imply that the test statistic has a t-distribution with T – 3 degrees of freedom. Notice, also, that if γ is unknown, the t-

[10] More precisely, if we know ρ, then we observe $\mu_t + \phi = x_t - \rho\, x_{t-1}$. The value of ϕ does not affect the estimate of β or γ (it affects only the intercept), and I ignore it in the remainder of the discussion.

statistic should use the standard error of β from (A.1).

I should emphasize two other features of (A.1). First, $\hat{\gamma}$ does not depend on the assumed value of ρ. As we change ρ, we simply add or subtract x_{t-1} from the value of μ_t that enters the multiple regression. Doing so affects only β, not γ or the residual variance. This observation is important because it means that the t-statistic is linearly related to ρ. In turn, this implies that any $\rho < 1$ yields stronger significance than $\rho = 1$, as claimed in the text.

Second, the variance of β from eq. (A.1) will typically be very close to the variance used in the text (which ignores uncertainty about γ). The variance from (A.1) is scaled up by $1 / (1 - R^2)$, where R^2 is the explained variance when μ_t is regressed on x_{t-1}. That R^2 is approximately equal to $(1 - \hat{\rho}) / (1 + \hat{\rho})$, which is close to zero when $\hat{\rho}$ is close to one. Empirically, the two standard errors are almost identical, differing by about 0.1%. In Table 2, for example, the standard error for nominal VWNY is 0.1761. The standard error from (A.1) is 0.1763.

References

Blanchard, Olivier and Mark Watson, 1982. Bubbles, rational expectations, and financial markets. In: Paul Watchel, ed., *Crises in the Economic and Financial Structure*. Lexington Books, Lexington, MA, 295-315.

Campbell, John, 1987. Stock returns and the term structure. *Journal of Financial Economics* 18 (2), 373-399.

Campbell, John and Robert Shiller, 1988a. The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies* 1, 195-228.

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

DeJong, David and Charles Whiteman, 1991. The temporal stability of dividends and stock prices: Evidence from the likelihood function. *American Economic Review* 81, 600-617.

Diba, Behzad and Herschel Grossman, 1988. Explosive rational bubbles in stock prices?. *American Economic Review* 78 (3), 520-530.

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

Flood, Robert and Robert Hodrick, 1986, Asset price volatility, bubbles, and process switching, *Journal of Finance* 41 (4), 831-842.

Flood, Robert and Robert Hodrick, 1990, On testing for speculative bubbles, *Journal of Economic Perspectives* 4 (2), 85-101.

Goetzmann, William and Philippe Jorion, 1993. Testing the predictive power of dividend yields. *Journal of Finance* 48, 663-679.

Hamilton, James and Charles Whiteman, 1985, The observable implications of self-fulfilling expectations, *Journal of Monetary Economics* 16, 353-373.

Hodrick, Robert, 1992. Dividend yields and expected stock returns: Alternative procedures for inference and measurement. *Review of Financial Studies* 5, 357-386.

Kendall, Maurice, 1953. The analysis of economic time series, part I: Prices. *Journal of the Royal Statistical Society* 96, 11-25.

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

Kothari, S.P. and Jay Shanken, 1997. Book-to-market, dividend yield, and expected market returns: A time-series analysis. *Journal of Financial Economics* 44, 169-203.

Kothari, S.P., Jay Shanken, and Richard Sloan, 1995. Another look at the cross-section of expected stock returns. *Journal of Finance* 50, 185-224.

Lamont, Owen, 1998. Earnings and expected returns. *Journal of Finance* 53, 1563-1587.

Loewenstein, Mark and Gregory Willard, 1999. Rational equilibrium asset-pricing bubbles in continuous trading models, forthcoming in *Journal of Economic Theory*.

Mankiw, N. Gregory and Matthew Shapiro, 1986. Do we reject too often? Small sample properties of tests of rational expectations models. *Economic Letters* 20, 139-145.

Nelson, Charles and Myung Kim, 1993. Predictable stock returns: The role of small sample bias. *Journal of Finance* 48, 641-661.

Nelson, Charles and Charles Plosser, 1982. Trends and random walks in macroeconomic time series. *Journal of Monetary Economics* 10, 139-162.

Pontiff, Jeffrey and Lawrence Schall, 1998. Book-to-market ratios as predictors of market returns. *Journal of Financial Economics* 49, 141-160.

Stambaugh, Robert, 1986. Bias in regressions with lagged stochastic regressors. Working paper (University of Chicago, Chicago, IL).

Stambaugh, Robert, 1999. Predictive regressions. *Journal of Financial Economics* 54, 375-421.

Stambaugh, Robert, 1997, Analyzing investments whose histories differ in length, *Journal of Financial Economics* 45, 285-331.

Tirole, Jean, 1982, On the possibility of speculation under rational expectations, *Econometrica* 50, 1163-1181.

Tirole, Jean, 1985, Asset bubbles and overlapping generations, *Econometrica* 53, 1071-1100.

West, Kenneth, 1988, Bubbles, fads, and stock price volatility tests: A partial evaluation, Journal of Finance 43, 639-656.

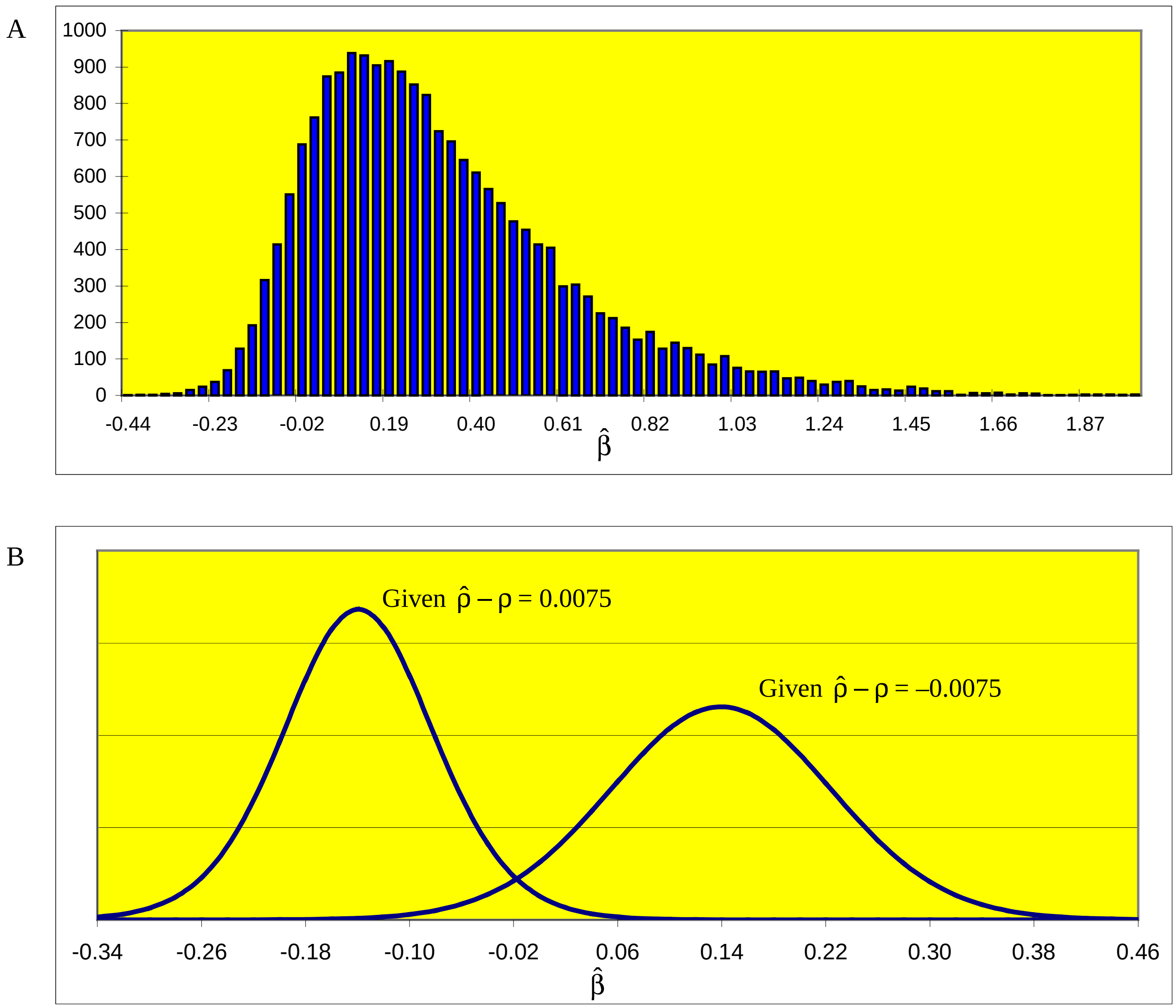

Figure 1
Distribution of the predictive slope, $\hat{\beta}$

The figure shows the distribution of the OLS slope estimate based on 20,000 Monte Carlo simulations. Panel A plots the marginal, or unconditional, distribution of $\hat{\beta}$, and Panel B plots the conditional distribution of $\hat{\beta}$ given two different values of $\hat{\rho}$ ($\hat{\rho}-\rho$ = +/–0.0075). The true coefficients are $\beta = 0$ and $\rho = 0.99$, the correlation between ε_t and μ_t is –0.92, and T = 300.

Table 1
Summary statistics, 1/46 – 12/97

The table reports summary statistics for monthly observations on stock returns, dividend yield, book-to-market, and the earnings-price ratio. Market value and dividend data come from CRSP and accounting data come from Compustat. The variables are expressed in percent. EWNY and VWNY are returns on the equal- and value-weighted NYSE indexes, respectively. The financial ratio are all calculated for the value-weighted index. DY equals the dividends paid over the prior year divided by the current level of the index. B/M is the ratio of book-to-market equity. E/P is the ratio of operating earnings defore depreciation to market equity. Log(·) denotes the natural logarithm.

				Autocorrelation			
Variable	Mean	S.D.	Skew.	ρ_1	ρ_{12}	ρ_{24}	ρ_{36}
			Returns and dividend yield				
Full sample: 1/46 – 12/97							
VWNY	1.05	4.05	-0.35	0.036	0.029	0.025	-0.013
EWNY	1.17	4.84	-0.18	0.135	0.066	0.035	0.026
DY	3.92	1.12	0.62	0.989	0.852	0.751	0.635
Log(DY)	1.33	0.28	-0.00	0.994	0.892	0.797	0.705
1st half: 1/46 – 12/71							
VWNY	0.97	3.72	-0.38	0.080	0.023	0.065	0.007
EWNY	1.06	4.43	-0.28	0.146	0.015	0.024	0.001
DY	4.09	1.20	0.82	0.991	0.867	0.762	0.625
Log(DY)	1.37	0.28	0.55	0.991	0.858	0.773	0.670
2nd half: 1/72 – 12/97							
VWNY	1.13	4.36	-0.34	0.002	0.031	0.003	-0.032
EWNY	1.28	5.23	-0.13	0.123	0.095	0.057	0.002
DY	3.75	1.01	0.13	0.986	0.802	0.702	0.676
Log(DY)	1.28	0.29	-0.46	0.996	0.906	0.808	0.789
			Book-to-market and the earnings-price ratio				
Compustat: 6/63 – 12/97							
B/M	56.03	16.97	0.47	0.988	0.855	0.768	0.706
Log(B/M)	3.98	0.30	0.04	0.995	0.914	0.825	0.752
E/P	20.78	6.73	0.49	0.987	0.835	0.712	0.625
Log(E/P)	2.98	0.32	0.09	0.990	0.856	0.713	0.623

Table 2
Dividend yield and expected returns, 1/46 – 12/97

The table reports an AR(1) regression for dividend yield and predictive regressions for stock returns for the full sample period, Jan. 1946 – Dec. 1997 (624 months). Observations are monthly. DY is the dividend yield on the value-weighted NYSE index, equal to dividends paid over the prior year divided by the current level of the index. Log(DY) is the natural logarithm of DY. EWNY and VWNY are monthly returns on the equal- and value-weighted NYSE indexes, respectively. Excess returns are calculated as EWNY and VWNY minus the one-month Tbill rate. All data come from CRSP and returns are reported in percent.

$\text{Log}(DY_t) = \phi + \rho\ \text{Log}(DY_{t-1}) + \mu_t$

		ρ	S.E.(ρ)	Bias	$-(1+3\rho)/T$	Adj. R^2	S.D.(μ)
AR(1)	OLS	0.994	0.006	-0.008	-0.006	0.976	0.044

$r_t = \alpha + \beta\ \text{Log}(DY_{t-1}) + \varepsilon_t$

		β	S.E.(β)	p-value	Adj. R^2	S.D.(ε)	cor(ε, μ)
VWNY	OLS	1.206	0.573	0.018	0.005	4.035	-0.952
	Stambaugh	0.506	0.748	0.210			
	$\rho \approx 1$	0.720	0.176	0.000			
EWNY	OLS	1.581	0.684	0.010	0.007	4.817	-0.879
	Stambaugh	0.809	0.869	0.153			
	$\rho \approx 1$	1.046	0.326	0.001			
Excess VWNY	OLS	1.191	0.576	0.019	0.005	4.056	-0.950
	Stambaugh	0.478	0.762	0.215			
	$\rho \approx 1$	0.704	0.179	0.000			
Excess EWNY	OLS	1.566	0.686	0.011	0.007	4.833	-0.879
	Stambaugh	0.805	0.875	0.154			
	$\rho \approx 1$	1.029	0.327	0.001			

Table 3
Dividend yield and expected returns, 1/46 – 12/71 and 1/72 – 12/97

The table reports AR(1) regressions for dividend yield and predictive regressions for stock returns for two subperiods, Jan. 1946 – Dec. 1971 and Jan. 1972 – Dec. 1997 (each 312 months). Observations are monthly. DY is the dividend yield on the value-weighted NYSE index, equal to dividends paid over the prior year divided by the current level of the index. Log(DY) is the natural logarithm of DY. EWNY and VWNY are monthly returns on the equal- and value-weighted NYSE indexes, respectively. Excess returns are calculated as EWNY and VWNY minus the one-month Tbill rate. All data come from CRSP and returns are reported in percent.

$\text{Log}(DY_t) = \phi + \rho\, \text{Log}(DY_{t-1}) + \mu_t$

		1946 – 1971						1972 – 1997					
		ρ	S.E.(ρ)	Bias	$-(1+3\rho)/T$	Adj. R^2	S.D.(μ)	ρ	S.E.(ρ)	Bias	$-(1+3\rho)/T$	Adj. R^2	S.D.(μ)
AR(1)	OLS	0.991	0.008	-0.016	-0.013	0.979	0.040	0.996	0.009	-0.017	-0.013	0.973	0.047

$r_t = \alpha + \beta\, \text{Log}(DY_{t-1}) + \varepsilon_t$

		1946 – 1971						1972 – 1997					
		β	S.E.(β)	p-value	Adj. R^2	S.D.(ε)	cor(ε,μ)	β	S.E.(β)	p-value	Adj. R^2	S.D.(ε)	cor(ε,μ)
VWNY	OLS	1.563	0.762	0.020	0.010	3.694	-0.933	1.030	0.875	0.120	0.001	4.360	-0.965
	Stambaugh	0.154	1.428	0.373				-0.498	1.383	0.577			
	$\rho \approx 1$	0.808	0.275	0.002				0.720	0.228	0.001			
EWNY	OLS	1.035	0.908	0.127	0.001	4.406	-0.883	2.350	1.042	0.012	0.013	5.196	-0.878
	Stambaugh	-0.571	1.712	0.565				0.694	1.582	0.263			
	$\rho \approx 1$	0.182	0.425	0.335				2.014	0.498	0.000			
Exc.VWNY	OLS	1.876	0.763	0.007	0.016	3.701	-0.931	0.532	0.881	0.273	-0.002	4.390	-0.966
	Stambaugh	0.455	1.461	0.299				-0.984	1.407	0.761			
	$\rho \approx 1$	1.120	0.279	0.000				0.220	0.229	0.168			
Exc.EWNY	OLS	1.347	0.909	0.069	0.004	4.412	-0.882	1.852	1.050	0.039	0.007	5.231	-0.878
	Stambaugh	-0.299	1.732	0.483				0.216	1.528	0.360			
	$\rho \approx 1$	0.494	0.428	0.124				1.514	0.502	0.001			

Table 4
Influence of 1995 – 1997

The table reports AR(1) regressions for dividend yield and predictive regressions for stock returns for two periods, Jan. 1946 – Dec. 1994 (588 months) and Jan. 1946 – Dec. 1997 (624 months). Observations are monthly. DY is the dividend yield on the value-weighted NYSE index, equal to dividends paid over the prior year divided by the current level of the index. Log(DY) is the natural logarithm of DY. EWNY and VWNY are monthly returns on the equal- and value-weighted NYSE indexes, respectively. All data come from CRSP and returns are reported in percent.

Without recent data: **Jan. 1946 – Dec. 1994**

$\text{Log}(DY_t) = \phi + \rho \,\text{Log}(DY_{t-1}) + \mu_t$

		ρ	S.E.(ρ)	Bias	$-(1+3\rho)/T$	Adj. R^2	S.D.(μ)
AR(1)	OLS	0.986	0.007	-0.008	-0.007	0.970	0.044

$r_t = \alpha + \beta \,\text{Log}(DY_{t-1}) + \varepsilon_t$

		β	S.E.(β)	p-value	Adj. R^2	S.D.(ε)	cor(ε,μ)
VWNY	OLS	2.191	0.662	0.000	0.017	4.058	-0.951
	Stambaugh	1.528	0.893	0.061			
	$\rho \approx 1$	0.954	0.206	0.000			
EWNY	OLS	2.430	0.800	0.001	0.014	4.897	-0.881
	Stambaugh	1.672	1.106	0.081			
	$\rho \approx 1$	1.047	0.378	0.003			

With recent data: **Jan. 1946 – Dec. 1997**

$\text{Log}(DY_t) = \phi + \rho \,\text{Log}(DY_{t-1}) + \mu_t$

		ρ	S.E.(ρ)	Bias	$-(1+3\rho)/T$	Adj. R^2	S.D.(μ)
AR(1)	OLS	0.994	0.006	-0.008	-0.006	0.976	0.044

$r_t = \alpha + \beta \,\text{Log}(DY_{t-1}) + \varepsilon_t$

		β	S.E.(β)	p-value	Adj. R^2	S.D.(ε)	cor(ε,μ)
VWNY	OLS	1.206	0.573	0.018	0.005	4.035	-0.952
	Stambaugh	0.506	0.748	0.210			
	$\rho \approx 1$	0.720	0.176	0.000			
EWNY	OLS	1.581	0.684	0.010	0.007	4.817	-0.879
	Stambaugh	0.809	0.869	0.153			
	$\rho \approx 1$	1.046	0.326	0.001			

Table 5
Book-to-market and expected returns, 6/63 – 12/94 and 6/63 – 12/97

The table reports AR(1) regressions for book-to-market and predictive regressions for stock returns for two subperiods, June 1963 – Dec. 1994 (379 months) and June 1963 – Dec. 1997 (415 months). Observations are monthly. B/M is the ratio of book equity to market equity on the value-weighted NYSE index. Log(B/M) is the natural logarithm of B/M. EWNY and VWNY are monthly returns on the equal- and value-weighted NYSE indexes, respectively. Excess returns are calculated as EWNY and VWNY minus the one-month Tbill rate. Market equity and return data come from CRSP, book equity data come Compustat, and returns are expressed in percent.

$Log(B/M_t) = \phi + \rho \, Log(B/M_{t-1}) + \mu_t$

		1963 – 1994						1963 – 1997					
		ρ	S.E.(ρ)	Bias	$-(1+3\rho)/T$	Adj. R^2	S.D.(μ)	ρ	S.E.(ρ)	Bias	$-(1+3\rho)/T$	Adj. R^2	S.D.(μ)
AR(1)	OLS	0.986	0.009	-0.013	-0.010	0.971	0.047	0.995	0.008	-0.013	-0.010	0.977	0.046

$r_t = \alpha + \beta \, Log(B/M_{t-1}) + \varepsilon_t$

		1963 – 1994						1963 – 1997					
		β	S.E.(β)	p-value	Adj. R^2	S.D.(ε)	cor(ε,μ)	β	S.E.(β)	p-value	Adj. R^2	S.D.(ε)	cor(ε,μ)
VWNY	OLS	1.882	0.805	0.010	0.012	4.270	-0.906	0.889	0.688	0.098	0.002	4.209	-0.901
	Stambaugh	0.866	1.231	0.203				-0.146	1.028	0.476			
	$\rho \approx 1$	0.767	0.340	0.012				0.459	0.299	0.062			
EWNY	OLS	2.469	0.994	0.006	0.014	5.274	-0.834	1.607	0.838	0.028	0.006	5.124	-0.828
	Stambaugh	1.287	1.550	0.178				0.467	1.200	0.288			
	$\rho \approx 1$	1.201	0.548	0.014				1.126	0.470	0.008			
Exc.VWNY	OLS	1.342	0.811	0.049	0.005	4.304	-0.908	0.439	0.693	0.263	-0.001	4.235	-0.902
	Stambaugh	0.303	1.268	0.335				-0.630	1.079	0.702			
	$\rho \approx 1$	0.217	0.340	0.262				0.006	0.299	0.497			
Exc.EWNY	OLS	1.929	1.001	0.027	0.007	5.311	-0.836	1.156	0.843	0.085	0.002	5.154	-0.829
	Stambaugh	0.751	1.535	0.259				-0.016	1.269	0.413			
	$\rho \approx 1$	0.651	0.549	0.118				0.672	0.472	0.077			

Table 6
Earnings-price ratios and expected returns, 6/63 – 12/94 and 6/63 – 12/97

The table reports AR(1) regressions for earnings-price ratios and predictive regressions for stock returns for two subperiods, June 1963 – Dec. 1994 (379 months) and June 1963 – Dec. 1997 (415 months). Observations are monthly. E/P is the ratio of operating earnings to market equity on the value-weighted NYSE index. Log(E/P) is the natural logarithm of E/P. EWNY and VWNY are monthly returns on the equal- and value-weighted NYSE indexes, respectively. Excess returns are calculated as EWNY and VWNY minus the one-month Tbill rate. Market equity and return data come from CRSP, earnings data come Compustat, and returns are expressed in percent.

$\text{Log}(E/P_t) = \phi + \rho\ \text{Log}(E/P_{t-1}) + \mu_t$

		1963 – 1994						1963 – 1997					
		ρ	S.E.(ρ)	Bias	$-(1+3\rho)/T$	Adj. R^2	S.D.(μ)	ρ	S.E.(ρ)	Bias	$-(1+3\rho)/T$	Adj. R^2	S.D.(μ)
AR(1)	OLS	0.987	0.008	-0.011	-0.010	0.977	0.049	0.990	0.007	-0.011	-0.010	0.977	0.049

$r_t = \alpha + \beta\ \text{Log}(E/P_{t-1}) + \varepsilon_t$

		1963 – 1994						1963 – 1997					
		β	S.E.(β)	p-value	Adj. R^2	S.D.(ε)	cor(ε,μ)	β	S.E.(β)	p-value	Adj. R^2	S.D.(ε)	cor(ε,μ)
VWNY	OLS	1.621	0.676	0.008	0.012	4.269	-0.860	1.279	0.642	0.023	0.007	4.197	-0.854
	Stambaugh	0.776	1.087	0.196				0.460	0.990	0.269			
	$\rho \approx 1$	0.639	0.345	0.032				0.540	0.334	0.053			
EWNY	OLS	2.045	0.836	0.007	0.013	5.276	-0.794	1.793	0.782	0.011	0.010	5.114	-0.784
	Stambaugh	1.064	1.346	0.190				0.887	1.199	0.197			
	$\rho \approx 1$	0.925	0.508	0.034				0.965	0.485	0.023			
Exc.VWNY	OLS	1.165	0.682	0.044	0.005	4.303	-0.862	0.833	0.647	0.099	0.002	4.228	-0.856
	Stambaugh	0.309	1.133	0.321				0.030	0.998	0.410			
	$\rho \approx 1$	0.173	0.346	0.308				0.087	0.334	0.398			
Exc.EWNY	OLS	1.589	0.842	0.030	0.007	5.312	-0.795	1.347	0.787	0.044	0.005	5.147	-0.786
	Stambaugh	0.622	1.352	0.263				0.467	1.174	0.293			
	$\rho \approx 1$	0.459	0.510	0.184				0.513	0.487	0.146			