Threshold Effect and the Predictive Ability of Dividend Yield

Tai-Wei Zhang
Hsing Wu College
Dept of Accounting Information
Taipei County, Taiwan, 244
R.O.C.

An-Sing Chen*
National Chung Cheng University
Department of Finance
Chia-Yi, Taiwan, 621
R.O.C.

Abstract

Whether dividend yield (DY) can forecast aggregate stock returns is the topic of many recent researches. However, statistical evidence of its predictive power is weak, inconsistent across time periods, and different across asset classes. Further, while time-varying risk preferences can induce the standard positive relation between dividend yields and expected returns, time-varying expected dividend growth may induce a negative relation between them, offsetting and reducing the ability of DY to forecast future returns. In this study, we show the existence of a DY threshold point, above which future returns are significantly negatively correlated with DY, but below which future returns are significantly positively correlated with DY. When modeled under the autocorrelation bias-adjusted predictability framework of Lewellen (2004), our threshold controlled forecasting equations show DY has significant ability to forecast future stock returns during the period 1927-2001. Conversely, when the threshold effect is not controlled for, DY fails to show significant ability to forecast future stock returns.

EFMA Classification Code: 760 - Methodological Issues
JEL Classifications: C32; C52; G12
Keywords: Predictive regressions; Bias; Threshold effect; Expected returns

---

*Corresponding author: An-Sing Chen, National Chung Cheng University, Department of Finance, Chia-Yi, Taiwan, 621, R.O.C. Email: finasc@ccu.edu.tw, Fax: 886-5-272-0818, Tel: 886-5-242-8246

Threshold Effect and the Predictive Ability of Dividend Yield

Introduction

Whether or not dividend yield (DY from now on) can predict aggregate stock returns has received much attention in the recent literature. However, statistical evidence of its predictive power is weak, inconsistent across time periods, and different across asset classes. For example, earlier studies such as Fama and French (1988) find that DY predicts monthly NYSE returns from 1941-1986, with t-statistics between 2.20 and 3.21. In contrast, other studies such as Stambaugh (1986) and Mankiw and Shapiro (1986), for example, show that predictive regressions can be severely biased toward finding predictability. Nelson and Kim (1993), applying bootstrap simulations to correct for bias, replicate the Fama and French tests and find that the p-values are actually between 0.03 and 0.33. More recently, Stambaugh (1999), using the exact small-sample distribution of the slope estimate when DY is assumed to follow a first-order autoregressive (AR1) process, reports a one-sided p-value of 0.15 when NYSE returns are regressed on DY over the period 1952-1996. On the other hand, in an even more recent article, Lewellen (2004) shows that the small-sample distribution studied by Stambaugh (1986,1999) and Nelson and Kim (1993), which has become standard in the literature, can substantially understate, in some circumstances, DY's predictive ability.

The standard positive relation between dividend yields and expected returns can be induced by time-varying risk preferences. However, time-varying expected dividend growth may induce a negative relation between them, offsetting and reducing the ability of DY to forecast future returns. Specifically, for a given expected dividend growth, a decrease in risk tolerance increases the equity premia on all assets and increases their DY. That is, the variation in risk preferences induces the standard positive relation between DY and expected returns. However, for a given investors' attitude toward risk, an increase in the asset's expected dividend growth yields both an decrease in its DY and an increase in its equity premium. That is, changes in expected dividend growth induce a negative relation between DY and expected excess returns, contrary to the common wisdom on return predictability. The reason is that an increase in expected dividend growth implies that the asset pays further ahead in the future, making its price more sensitive to shocks to the aggregate discount rate, that is, to fluctuations in the investors' risk preferences. Since this additional volatility of the asset is perfectly correlated with changes in investors' attitude toward risk, it must be priced, and thus the premium is larger. In equilibrium, however, this increase in the premium is not sufficient to offset the decrease in the DY that stems from a higher expected dividend growth, so the negative relation between DY and expected future returns remains, though attenuated. Under this framework, the above offsetting effects can reduce the ability of the dividend yield to forecast future returns. Dividend yield, thus, may not always be a good predictor for future returns since it may not always vary with expected returns with the same sign.

Menzly, Santos and Veronesi (2004) present a general equilibrium model in which both investors' preferences for risk and their expectation of future dividend growth are time-varying, where time-varying risk preferences induce the standard positive relation between dividend yields and expected returns, while time-varying expected dividend growth induces a negative relation between them in equilibrium. Their model, however, requires that the expected excess rate of return of an asset be a linear function not only of its DY but also of its consumption/price (CP) ratio. Moreover, in their framework, fluctuations in DY must predict changes in dividend growth. Empirically, aggregate dividend growth, though, is not predicted by past DY of the market portfolio.

In this study we describe a threshold-based correction to the standard predictability regressions to disentangle the conflicting effects that variations in expected dividend growth and the preferences for risk have on the expected excess returns of the stock index. Because DY may not always be a good predictor for future returns since it may not always vary with expected returns with the same sign, introducing a threshold effect to the predictability equations allows for the attractive plausibility that investors may have different expected dividend growth prospects with respect to different magnitude of DY. That is, if DY is below a certain threshold consistent with "normal DY" or "normal economic circumstances or times", the standard positive relation between DY and expected future returns maintains. However, if DY is "abnormally high" (above a certain threshold), investors may view it as "abnormal DY" or "abnormal economic circumstances or times", and thus may attribute different (higher) expected dividend growth than in "normal times." If the expected dividend growth is over a certain level, its effect may be large enough to overwhelm the standard positive relation between DY and expected future returns, and, thus, vary with expected returns with the opposite sign instead.

The unobservable DY threshold can be computed empirically from data using standard econometric models. In this study, we derive threshold controlled forecasting equations by modeling the autocorrelation bias-adjusted predictability equations of Lewellen (2004) into the threshold testing methodology equations of Bai and Perron (1998, 2003 (BP)). Using our constructed threshold and autocorrelation bias adjusted equations, we show the existence of a DY threshold point, above which future returns are significantly negatively correlated with DY, but below which future returns are significantly positively correlated with DY. Our threshold and autocorrelation bias controlled forecasting equations show DY has significant ability to forecast future stock returns during the period 1927-2001. On the other hand, when the threshold effect is not controlled for, DY fails to show significant ability to forecast future stock returns under the autocorrelation bias-adjusted predictability equations of Lewellen (2004).

The rest of the paper is organized as follows. Section 2 describes the econometric procedures we use to control for possible threshold effects in predictive regression models. Important issues relating to the time-series properties of regressors appearing in predictive regression models of stock returns are then discussed. Section 3 describes the data and computes the DY threshold test results. Section 4 gives the empirical results. Section 5 concludes and summarizes our main findings.

Methodology

Predictive regressions have often been used in the finance literature to test whether past prices, financial ratios, interest rates, and a variety of other macroeconomic variables can forecast stock and bond returns. For stock market indices, one common form is

$$r_t = \alpha + \beta x_{t-1} + \varepsilon_t \quad (1)$$

where r_t is the return in month t and x_{t-1} is a predictive variable (DY) known at the beginning of the month and is assumed to follow a stationary AR1 process:

$$x_t = \phi + \rho x_{t-1} + \mu_t \quad (2)$$

where $\rho < 1$. When the predictive variable is DY the residuals in (1) and (2) will be negatively correlated since an increase in price leads to a decrease in DY. It follows that ε_t is correlated with x_t in the predictive regression, violating the OLS assumption of requiring independence at all leads and lags. As a consequence, estimation errors in the two equations are closely connected:

$$\hat{\beta} - \beta = \gamma(\hat{\rho} - \rho) + \eta \tag{3}$$

where η is a random error with mean zero and γ is a negative constant. The bias in $\hat{\beta}$ is typically found by taking expectations of both sides of (3). However, this approach implicitly discards any information we have about $(\hat{\rho} - \rho)$. In particular, Lewellen (2004) shows that for stationary predictive variables such as DY, the bias in $\hat{\beta}$ is at most $\gamma(\hat{\rho} - 1)$. This upper bound will be less than the standard bias-adjustment if $\hat{\rho}$ is close to one, and empirical tests that ignore the information in $\hat{\rho}$ will understate DY's predictive power.

The tests in this paper therefore use the Lewellen (2004) bias-adjusted estimator

$$\hat{\beta}_{adj} = \hat{\beta} - \gamma(\hat{\rho} - \rho) \tag{4}$$

where ρ is assumed to be approximately one (operationalized as $\rho = 0.9999$). The variance of $\hat{\beta}_{adj}$ is $\sigma_v^2 (X'X)^{-1}_{(2,2)}$. To implement the test, we can estimate γ and σ_v from $\varepsilon_t = \gamma\mu_t + v_t$, where ε_t and μ_t are the residuals in (1) and (2). Operationally, Lewellen(2004) used the following equation to estimate $\hat{\beta}_{adj}$ and γ

$$r_t = \alpha + \beta x_{t-1} + \gamma(x_t - .9999 x_{t-1}) + \upsilon_t \tag{5}$$

Equation (5) can be further modified to control for threshold effects by specifying x_{t-1} as the threshold variable (v_t):

$$r_t = \alpha_j + \beta_j x_{t-1} + \gamma_j (x_t - .9999 x_{t-1}) + \upsilon_t, \qquad \tau_{j-1} < v_t < \tau_j \tag{6}$$

for $j = 1,...,m+1$ with the convention that $\tau_0 = -\infty$ and $\tau_{m+1} = +\infty$. Thus, the functional form of equation (6) now depends on the value of the observable variable x_{t-1}.

Applying the Bai and Perron (1998, 2003; hereafter BP) methodology to (6) results in a joint test of (bias adjusted) predictability and threshold effects. That is, we regress aggregate stock return on lag one DY and the difference between current DY and lag one DY which multiplies 0.9999,and then test for threshold effects in the constant term and slope terms. In particular, consider such a regression model with m threshold points $(\tau_1,...,\tau_m)$ and m+1 regimes, i.e.,

$$r_t = z_t' \delta_j + \upsilon_t \qquad\qquad \tau_{j-1} < v_t < \tau_j \tag{7}$$

for $j = 1,...,m+1$, where $z_t = (1, \log DY_{t-1}, \log DY_t - .9999 \log DY_{t-1})'$ and $\delta_j = (\alpha_j, \beta_j, \gamma_j)'$. Let $v' = (v_1,...,v_T)$ and $v^{*\prime} = (v_{t_1},...,v_{t_T})$ be the sorted version of v' such that $v_{t_1} \le v_{t_2} \le ... \le v_{t_T}$. The indices $(t_1,...,t_T)$ are a permutation of the time indices $(1,...,T)$. Now, for $i = 1,..,m$, let T_i be the time index such that $v_{t_j} \le \tau_i$ for all j such that $j \le T_i$ and $v_{t_j} > \tau_i$ for all j such that $j > T_i$. The m-partition $(T_1,...,T_m)$ is the partition that corresponds to the time indices of the sorted vector $v^{*\prime}$ when the variables v_{t_j} reach each of the m thresholds. We can write the model (7) using all variables sorted according to the partition $(T_1,...,T_m)$. Then, we have, for $j = 1,...,T$ and $i = 1,..,m+1$:

$$r_{t_j} = z_{t_j}{}'\delta_i + \upsilon_{t_j} \qquad j = T_{i-1},...,T_i \tag{8}$$

(using $T_0 = 0$ and $T_{m+1} = T$). The BP methodology is used to test for multiple threshold effects in (8) and allows for the threshold points to be explicitly treated as unknowns.[1] Let the estimate of the partition be denoted by $\left(\hat{T}_1,...,\hat{T}_m\right)$; the estimates of the thresholds are then recovered as $\hat{\tau}_j = \nu_{t_r}$ with $r = \hat{T}_j$ for $j = 1,...,m$. One can then recover the estimates of δ_j from (7) by OLS conditioning on the threshold values.

To test the null hypothesis of no threshold effect against the alternative of $m = b$ threshold points, let $(T_1,...,T_b)$ be a partition such that $T_i = [T\lambda_i]$ $(i = 1,...,b)$. Also define R such that $(R\delta)' = (\delta_1{}' - \delta_2{}',...,\delta_b{}' - \delta_{b+1}{}')$. Then the following statistic can be specified:

$$F_T(\lambda_1,...,\lambda_b) = \frac{1}{T}\left(\frac{T-2(b+1)-1}{2b}\right)\hat{\delta}'R'\left[R\hat{V}'(\hat{\delta})R'\right]^{-1}R\hat{\delta} \tag{9}$$

where $\hat{V}(\hat{\delta})$ is an estimate of the variance-covariance matrix for $\hat{\delta}$ that is robust to heteroskedasticity and serial correlation.

To test the null hypothesis of l threshold points against the alternative hypothesis of $l+1$ threshold points, the statistic labeled $SupF_T(l+1|l)$ by BP can be used. Operationally, the global minimized sum of squared residuals for the model

[1] See technical appendix

with l threshold points is first computed. Each of the intervals defined by the l threshold points is then analyzed for an additional threshold point. From all of the intervals, the partition allowing for an additional threshold point that results in the largest reduction in the sum of squared residuals is considered as the model with $l+1$ threshold points. The $SupF_T(l+1|l)$ statistic tests whether the additional threshold point leads to a significant reduction in the sum of squared residuals.

The BP methodology allows for quite general specifications when computing confidence intervals and test statistics for the threshold points and regression coefficients including autocorrelation and heteroskedasticity in the residuals, and different moment matrices for the regressors in the different regimes.

Data, Descriptive Statistics and Threshold Test Results

Prices and dividends come from the Center for Research in Security Prices (CRSP) database. Our tests focus on NYSE value-weighted index. DY is defined as dividends paid over the prior year divided by current level of the index and is calculated monthly on the value-weighted NYSE index. We use value-weighted DY to predict returns on the NYSE index. The predictive regressions use the natural log of DY. The analyses focus on the period January 1927 to December 2001.

Table 1 provides summary statistics for the data. Log DY averages 1.35% over the full sample with a standard deviation of 0.36%. The table also shows that our predictor variable is extremely persistent. The first-order autocorrelation is 0.991. The autocorrelations tend to diminish as the lag increases. That the log DY is found to be highly autocorrelated is important for the empirical tests since the bias-adjustment depends on $\hat{\rho}-1$. The table also provides corresponding summary statistics for the returns and the excess returns on the value-weighted NYSE index, VWNY and EVWNY, respectively.

Table 2 reports BP statistics for tests of threshold effect of the parameters in our predictive regression model (i.e. equation 8). For this test, both double maximum statistics are significant at conventional significance levels, giving us strong evidence of threshold effects in the predictive regression model. The $F(2|1)$ statistic is insignificant at the 5% level, while the $F(3|2)$ is also insignificant, indicating just only one threshold point (two regimes) for this model. The threshold point for the predictive variable DY is found to occur at 1.6376.

Empirical Results

Table 3 explores the predictive ability of log DY in the full sample. The Table reports a variety of statistics. The row labeled 'OLS' shows the least-squares slope and standard error. These estimates ignore bias and are reported primarily as a benchmark. The row labeled '$\rho \approx 1$' reports estimates based on the conditional distribution of $\hat{\beta}$. The slope coefficient is the Lewellen (2004) bias-adjusted estimator (equation (4)). The results, as a whole, show no evidence of predictability for both the nominal return on the value-weighted index (VWNY) and its corresponding excess returns (EVWNY) series. For the nominal return on the value-weighted index (VWNY), the conditional test of Lewellen (2004) provides a bias-adjusted estimate of –0.081 with a p-value of 0.587. For EVWNY, the bias-adjusted estimate is 0.072 and the p-value is 0.630.

Table 4 reports the results of our joint test for predictability and threshold effect under the Lewellen (2004) bias adjusted estimator. Panel A shows the results for VWNY and Panel B shows the results for EVWNY. Both VWNY and EVWNY are shown to be highly predictable using DY. Panel A shows that when DY is below

1.6376, the standard positive relation between DY and return results. The bias-adjusted slope for VWNY is 0.631 with a p-value of 0.002. However, Panel A also shows that when DY is above 1.6376, the relation between DY and return reverses. The slope coefficient and p-values are now –3.235 and 0.000, respectively. Panel B shows similar results for EVWNY. When DY is below 1.7017, the bias-adjusted slope for EVWNY is 0.519 with a p-value of 0.007. When DY is above 1.7017, the slope coefficient and p-values becomes –3.80 and 0.000, respectively.

Summary and Conclusion

Whether DY can forecast aggregate stock returns has evolved considerably over the last 20 years. Initial tests produced strong evidence that market returns are predictable using financial variables such as DY. On the other hand, more recent evidence suggests that DY has at best, weak power to predict returns. Under standard predictability equations, DY is positively related with future returns. However, time-varying expected dividend growth may induce a negative relation between DY and future returns, offsetting and reducing the ability of DY to forecast returns under standard predictability equations.

In this study we describe a threshold-based correction to the standard predictability regressions to disentangle the conflicting effects that variations in expected dividend growth and the preferences for risk have on the expected excess returns of the stock index. Introducing a threshold effect to the predictability equations allows for the attractive plausibility that investors may have different expected dividend growth prospects with respect to different magnitude of DY. Methodologically, we showed that we could feasibly compute the unobservable DY threshold from data by modeling the autocorrelation bias-adjusted predictability equations of Lewellen (2004) into the threshold testing methodology equations of Bai and Perron (1998, 2003 (BP)). Using our constructed threshold and autocorrelation bias adjusted equations, we show the existence of a DY threshold point, above which future returns are significantly negatively correlated with DY, but below which future returns are significantly positively correlated with DY. This DY threshold shows up for both the VWNY and EVWNY (excess return) series. Our overall results show that DY has significant ability to forecast both nominal and excess future stock returns during the period 1927-2001 under the threshold and autocorrelation bias controlled forecasting equations. We also show that when the threshold effect is not controlled for, DY fails to forecast future stock returns under the autocorrelation bias-adjusted predictability equations of Lewellen (2004).

Our model is intuitive from an economic viewpoint since, in real life, investors may have different expected dividend growth prospects with respect to different magnitude of DY. We show that when DY is below a certain threshold consistent with “normal DY” or “normal economic circumstances or times”, the standard positive relation between DY and expected future returns maintains. Alternatively, when DY is “abnormally high” (above the threshold), investors may view it as “abnormal DY” or “abnormal economic circumstances or times”, and thus may attribute different (higher) expected dividend growth than in “normal times.” If the expected dividend growth is over a certain level, its effect may be large enough to overwhelm the standard positive relation between DY and expected future returns, and, thus, vary with expected returns with the opposite sign instead.

Technical Appendix

For each m-partition $(T_1, ..., T_m)$, the associated least-squares estimates of δ_i can be obtained by minimizing the sum of squared residuals,

$$S_T(T_1, ..., T_M) = \sum_{i=1}^{m+1} \sum_{t_j = T_{i-1}+1}^{T_i} \left(r_{t_j} - z'_{t_j} \delta_i \right)^2 \qquad (1A)$$

Let the regression coefficient estimates based on a given m-partition $(T_1, ..., T_m)$ be denoted by $\hat{\delta}(\{T_1, ..., T_m\})$, where $\delta = \left(\delta_1', ..., \delta_{m+1}' \right)'$. Substituting these into equation (1A), the estimated threshold points are given by

$$\left(\hat{T}_1, ..., \hat{T}_m \right) = \arg\min_{T_1, ..., T_M} S_T\left(T_1, ..., T_M \right) \qquad (2A)$$

From equation (2A), it is clear that the threshold point estimators correspond to the global minimum of the sum of squared residuals objective function. With the breakpoint estimates in hand, the corresponding least-squares regression parameter estimates $\hat{\delta} = \hat{\delta}\left(\{ \hat{T}_1, ..., \hat{T}_m \} \right)$ can be computed. Dynamic programming can be used to efficiently compute these estimates.

To test the null hypothesis of no threshold points against the alternative hypothesis of an unknown number of threshold points given an upper bound M, two “double maximum” statistics developed by BP can be used. Firstly, consider the following maximum F -statistic corresponding to equation (9) in the main body,

$$SupF_T(b) = F_T\left(\hat{\lambda}_1,...,\hat{\lambda}_b\right) \tag{3A}$$

where $\left(\hat{\lambda}_1,...,\hat{\lambda}_b\right)$ minimize the global sum of squared residuals, $S_T\left(T\hat{\lambda}_1,...,T\hat{\lambda}_b\right)$, under the restriction that $\left(\hat{\lambda}_1,...,\hat{\lambda}_b\right) \in \Theta_\pi$, where $\Theta_\pi = \{(\lambda_1,...,\lambda_b);|\lambda_{i+1} - \lambda_i| \geq \pi, \lambda_1 \geq \pi, \lambda_b \leq 1-\pi\}$ for some arbitrary small positive number π (the trimming parameter). The first double maximum statistic can then be written as

$$UD_{\max} = \max_{1 \leq m \leq M} SupF_T(m) \tag{4A}$$

The second double maximum statistic, $WD_{\max}$, applies different weights to the individual $SupF_T(m)$ statistics so that the marginal p -values are equal across values of m.

To test the null hypothesis of l threshold points against the alternative hypothesis of $l+1$ threshold points, the statistic labeled $SupF_T(l+1|l)$ by BP can be used. To determine the number of threshold points, BP recommends the following strategy. First, examine the double maximum statistics to determine if any threshold points are present. If the double maximum statistics are significant, then examine the $SupF_T(l+1|l)$ statistics to decide on the number of threshold points, choosing the $SupF_T(l+1|l)$ statistic that rejects for the largest value of l. BP (2001) recommend using a trimming parameter π of at least 0.15 (corresponding to 5 threshold points). We follow this recommendation.

References

Bai, J., Perron, P.,1998. Estimating and testing linear models with multiple structure changes. Econometrica 66, 47-78.

Bai, J., Perron, P., 2003. Computation and analysis of multiple structure change models. Journal of Applied Econometrics 18, 1-22.

Fama, E., French, K., 1988. Dividend yields and expected stock returns. Journal of Financial Economics 22, 3-25.

Lewellen, J., 2004. Predicting returns with financial ratios. Journal of Financial Economics 74, 209-235.

Mankiw, N.G., Shapiro, M., 1986. Do we reject too often? small sample properties of tests of rational expectations modes. Economics Letters 20, 139-145.

Menzly, L., Santos, T., Veronesi, P., 2004. Understanding predictability. Journal of Political Economy 112, 1-47.

Nelson, C. R., Kim, M.J., 1993. Predictable stock returns: the role of small sample bias. Journal of Finance 48, 641-661.

Stambaugh, R., 1986. Bias in regressions with lagged stochastic regressors. Unpublished Manuscript, University of Chicago, Chicago, IL.

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

Table 1: Summary statistics

The table reports summary statistics for stock returns and dividend yield for the period covering January 1927 to December 2001 (900 months). Observations are monthly and the variables are expressed in percent. VWNY is the returns on the value-weighted NYSE index. EVWNY is the corresponding excess returns. Excess returns are calculated as VWNY minus the one-month T-bill rate. The dividend yield (*DY*) equals dividends paid over the prior year divided by the current level of the index.

Variable	Mean	S.D.	Skew.	Autocorrelation		
				ρ_1	ρ_{12}	ρ_{24}
VWNY	0.97	5.44	0.24	0.097	-0.001	0.026
EVWNY	0.67	5.46	0.29	0.100	-0.0004	0.027
Log(*DY*)	1.35	0.36	-0.35	0.991	0.841	0.730

Table 2: Bai and Perron statistics for tests of multiple thresholds in the parameters of the predictive regression model

Predictive Model	UD*max*[a]	WD*max*[b]	*F*(2|1)[c]	*F*(3|2)[c]	*F*(4|3)[c]	*F*(5|4)[c]
VWNY	52.7508***	52.7508***	7.1372	10.1818	7.2161	1.3136
EVWNY	36.9939***	36.9939***	7.2347	3.0907	6.4695	1.5027

Notes: The statistics are used to test for structural change in the parameters, α and β, in the predictive regression model,

$r_t = \alpha + \beta \text{Log}(DY_{t-1}) + \gamma (\text{Log}(DY_t) - 0.9999*\text{Log}(DY_{t-1})) + \varepsilon_t$,

$er_t = \alpha + \beta \text{Log}(DY_{t-1}) + \gamma (\text{Log}(DY_t) - 0.9999*\text{Log}(DY_{t-1})) + \varepsilon_t$,

where r_t is the returns on the value-weighted NYSE index (VWNY) , er_t is the excess returns on the value-weighted NYSE index (EVWNY), $\text{Log}(DY_{t-1})$ is the natural logarithm of dividend yield on the value-weighted NYSE index, and ε_t is a disturbance term(all at time t);

*** indicates significance at the 1 percent level. The minimal length of any threshold regime is required to 15% of the full sample.

[a]One-sided (upper-tail) test of the null hypothesis of no threshold effect against the alternative hypothesis of an unknown number of threshold points given an upper bound of 5. The 1 percent critical value is equal to 18.26.

[b]One-sided (upper-tail) test of the null hypothesis of no threshold effect against the alternative hypothesis of an unknown number of threshold points given an upper bound of 5. The 1 percent critical value is equal to 19.86.

[c]One-sided (upper-tail) test of the null hypothesis of l threshold points against the alternative hypothesis of $l+1$ threshold points; $F(2|1)$, $l = 1;\ldots;$ $F(5|4)$, $l = 4$; 10, 5, and 1 percent critical values are: $F(2|1)$, 13.91, 15.72, and 19.77 ; $F(3|2)$, 14.96, 16.83 , and 20.75; $F(4|3)$, 15.68, 17.61, and 21.98 ; $F(5|4)$, 16.35, 18.14, and 22.46.

Table 3: Dividend yield and expected returns without controlling for threshold effects

The table reports an AR1 regression for dividend yield and predictive regressions for stock returns for the period January 1927-December 2001 (900 months). Observations are monthly. *DY* is the dividend yield on the value-weighted NYSE index and Log(*DY*) is the natural logarithm of *DY* . VWNY is the returns on the value-weighted NYSE index. EVWNY is the corresponding excess returns. Excess returns are calculated as VWNY minus the one-month T-bill rate. Returns are expressed in percent. For the predictive regressions, 'OLS' reports the standard OLS estimates, and '$\rho \approx 1$' reports the bias-adjusted estimate and *p*-value assuming that ρ is approximately one.

$Log(DY_t) = \phi + \rho\, Log(DY_{t-1}) + \mu_t$							
		ρ	S.E.(ρ)	*p*-value	Adj. R^2	S.D.(μ)	
AR(1)	OLS	0.991	0.005	0.000	0.976	0.014	
$r_t = \alpha + \beta\, Log(DY_{t-1}) + \varepsilon_t$							
		β	S.E.(β)	*p*-value	Adj. R^2	S.E.(ε)	cor(ε, μ)
VWNY	OLS	0.773	0.501	0.062	0.003	5.437	-0.955
	$\rho \approx 1$	-0.081	0.149	0.587			
EVWNY	OLS	0.927	0.502	0.032	0.003	5.445	-0.954
	$\rho \approx 1$	0.072	0.150	0.630			

Table 4: Dividend yield and expected returns accounting for threshold effects

Panels A and B reports the results of our joint test for predictability and threshold effects under the Lewellen (2004) bias adjusted estimator that accounts for AR(1) by fixing the slope of the dividend yield autoregressive parameter to approximately equal 1 (.9999). Observations are monthly. *DY* is the dividend yield on the value-weighted NYSE index and Log(*DY*) is the natural logarithm of *DY* . VWNY is the returns on the value-weighted NYSE index. Excess returns (EVWNY) are calculated as VWNY minus the one-month T-bill rate. Returns are expressed in percent.

Panel A: VWNY

$r_t = \alpha + \beta \text{Log}(DY_{t-1}) + \gamma (\text{Log}(DY_t) - 0.9999*\text{Log}(DY_{t-1})) + \varepsilon_t$

Threshold Value	Observations	α (*p*-value)	β (*p*-value)	γ (*p*-value)
$Log(DY) \leq 1.6376$	720	0.171 (0.498)	0.631 (0.002)	-88.37 (0.000)
$Log(DY) > 1.6376$	180	6.419 (0.000)	-3.235 (0.000)	-99.08 (0.000)

Panel B: EVWNY

$er_t = \alpha + \beta \text{Log}(DY_{t-1}) + \gamma (\text{Log}(DY_t) - 0.9999*\text{Log}(DY_{t-1})) + \varepsilon_t$

Threshold Value	Observations	α (*p*-value)	β (*p*-value)	γ (*p*-value)
$Log(DY) \leq 1.7017$	760	-0.052 (0.832)	0.519 (0.007)	-88.78 (0.000)
$Log(DY) > 1.7017$	140	7.474 (0.000)	-3.80 (0.000)	-99.44 (0.000)