That is not my dog: why doesn't the log dividend-price ratio seem to predict future log returns or log dividend growths?

Philip H. Dybvig

Washington University in Saint Louis

Huacheng Zhang

Southwest University of Finance and Economics

First draft: August, 2016

Current version: January, 2018

That is not my dog: why doesn't the log dividend-price ratio seem to predict future log returns or log dividend growths? [1]

By Philip H. Dybvig and Huacheng Zhang

Abstract

The "accounting identity" of Campbell and Shiller implies that change in the log dividend-price ratio must predict either future returns or future log dividend growth. It is a well-known puzzle that neither seems predictable. We examine this puzzle step-by-step from theoretical estimation through empirical testing. Stationarity of the log dividend-price ratio is an important assumption behind the accounting identity, but Campbell and Shiller's claimed test of the assumption does not make sense, and a correct test does not reject nonstationarity. Nonetheless, the accounting identity works reasonably well in the existing sample using a truncated version of the identity. Using the truncated identity, we conclude that predictability is in log dividend growth, not returns. Unfortunately, this result does not seem to be robust to subsamples. Also, it seems unwise to rely much on asymptotic properties of estimators when we only have about five non-overlapping observations even in the whole sample.

Key words: return predictability, dividend-price ratio, stationarity test. [JEL G12 G17]

[1]Dybvig: Olin School of Business, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130, USA, (dybvig@wustl.edu); Zhang: Institute of Financial Studies, Southwestern University of Finance and Economics, 55 Guanghuacun Street, Chengdu, 610074, China (zhanghuacheng1@gmail.com). We thank to Michael Brennan, Jason Chen (discussant), Pingyang Gao, Jun Liu, Zhongzhi Song, Rossen Valkanov, Russ Wermers, and participants at the SFS Cavalcade Asia-Pacific 2017 Annual Meeting (Beijing), Southwestern University of Finance and Economics, University of Alabama, and Washington University at Saint Louis for helpful comments.

Clouseau: Does your dog bite?

Innkeeper: No.

Clouseau: Nice doggy.

(Clouseau tries to pet the dog on the floor and is bitten)

Clouseau (angry): I thought you said your dog did not bite.

Innkeeper: That is not my dog. [2]

Many times, we get into trouble because we ask the wrong question. In this paper, we look at the paradox of the failure of the predictability by the log dividend-price ratio (hereafter LDPR) on both future log returns and future log dividend growth. The "accounting identity" of Campbell and Shiller (1988) asserts that current LDPR is approximately equal to a constant plus the sum of present values of future log returns minus the sum of present values of future log dividend growths. This implies that current LDPR should be able to predict future returns, or dividend growth rates, or both, but empirically it seems to predict neither. The puzzle was examined by Cochrane (2008), who sidestepped the puzzle by assuming a just-identified model, but we have a different approach. Because we do not know which question to ask, we go step-by-step through the entire argument to uncover where the problem is. We find that:

(1) The log-linear approximation works very well both in a single period and over 30 years.

(2) Campbell and Shiller use a theoretical assumption that the long-term mean LDPR exists, which they justify by an empirical tests which unfortunately does not test this at all. Our correctly specified test fails to reject the null that LDPR is not a stationary series.

(3) Even though the long-term mean of the LDPR seems not to exist, the approximation works pretty well in the sample to date. In particular, discarding the final term matters, but the relationship is still strong without the final term.

(4) Using a test with many lags as in the theory,[3] future log dividend growth is significantly predictable but expected returns are not. We show that this is hard to uncover because the large prediction error introduced by the unpredictable part of future log returns is a common factor in

[2]Metro-Goldwyn-Mayer, 1976. The Pink Panther Strikes Again. Amjo Productions, Video, 1976.

[3]Using the appropriate Newey-West adjustment for serially correlated errors and Stambaugh adjustment for spurious regression bias.

future log dividend growth but cancels in the accounting identity.

(5) Although the results are significant, they do not seem to be very robust. In particular, the result is reversed significantly on the second half-sample (consistent with Chen (2009)). Also, the estimation is asking a lot of the Newey-West adjustment, since there are only about five non-overlapping observations in the whole sample (and even fewer on subperiods). The lack of evidence for stationarity of the LDPR is also troubling for estimation. If the LDPR is not stationary, the approximation in the accounting identity should get worse and worse over time.

The Campbell-Shiller "accounting identity" can be derived by starting with the single period definition of returns as the sum of dividends plus capital gains. Using algebra, taking logs, and doing a Taylor series expansion around some typical value for LDPR,[4] we obtain a one-period approximation formula linking log returns and log dividend growth with beginning- and end-period LDPR. By telescoping this approximation over many periods, current LDPR can be linearly approximated by the sum of weighted log returns and the sum of weighted dividend growth rates over future periods plus the LDPR in the final period. Using the annual dividend payment and price of $S\&P500$ index from 1871 to 2015, and based on a regression of LDPR on the sums over future 30 years and the final log dividend-price ratio, all coefficients on independent variables are close to the theoretical value one (or minus one) and the R^2 is close to 100% (98.91%). The approximation is worse but not so bad if we drop the final term after 30 years (the R^2 dropped to 82.25%) . Therefore, the source of the puzzle is the lack of power in previous tests, not any intrinsic problem with the theory, at least in the truncated version.

We correct an error in the Campbell-Shiller stationarity test (in which both the null and the alternative contain a trend and therefore neither is stationary), conduct a conventional stationarity test in which the alternative hypothesis is that the underlying series is stationary (without trends), and cannot reject the null of nonstationarity (which may mean the long-term mean does not exist) when we use the entire sample. In principle, this is a big problem for the Campbell-Shiller approximation, which is based on an expansion around the long-term mean, but the truncated version of the approximation still works well in our sample looking out 30 years.

We test the predictability of stock returns and dividend growth rate using an equation that looks

[4] Campbell and Shiller expanded around the long-term mean, but that is not necessary.

like the equation in the model instead of using one or a few lags as is common in the literature. Our estimation uses statistical corrections for the correlation in error terms and for spurious regression bias. The Campbell-Shiller approximation implies that the current LDPR is able to predict either future returns or dividend growth or both; we find that only future log dividend growth is significantly predictable, but future returns are not. The results are robust when we expand the log dividend price ratio around alternative points rather than the sample mean.

As noted by Cochrane (2008), dividends are smooth. He concludes that log dividend growth is not predictable (implying under the model restriction that returns are predictable). However, it is more accurate to assert that predictability of log dividend growth is spread over many maturities and nearby dividends are not very predictable because dividends are smooth. What is happening is that there is small predictability of dividend growth spread over many periods, which is buried by noise in conventional simple regression or vector-autoregressive (VAR) estimation. This issue could be exaggerated in test with one or few lags when the log dividend-price is not stationary.

The limitation of using small lags to find predictability of dividend growth seems to be a solution of the puzzle of why the theory (based on an accounting identity and an approximation that is not so bad in the current sample) has not been verified. In general, the predictability of log dividend growth is a little difficult to find because of the large prediction error introduced by the unpredictable part of future log returns, which is a common factor with future log dividend growth that cancels in the accounting identity.

Although the best evidence based on our whole sample suggests that the LDPR predicts log dividend growth but not log returns, this result seems fragile. The result is not robust to subperiods (consistent with Chen (2009) and explaining an apparent inconsistentcy with a similar regression of Cochrane (2008, Section 7.2). We also worry about the statistical properties of the estimators, both because the whole sample has only about five non-overlapping observations (and subsamples even fewer) and because of the apparent instability over time. If the LDPR is in fact not stationary, there will be bigger problems as the sample gets larger because the Taylor series expansion will become much less accurate as the range of values increases over time. One interesting thing about the accounting identity is that it is not an economic model since its derivation just uses manipulation of identities and approximations. If we think about the economics, Modigliani-Miller suggests that to first order dividends are irrelevant, which is consistent with instability of these relationships over time.

The rest of this paper is organized as the follows. We review the approximation leading to the accounting identity Section 1 and test the quality of the approximation in Section 2. We propose a model-implied novel approach to test the predictability of return and dividend growth in Section 3. We analyze whether the failure of predictability of stock returns is caused by noise in itself or noise introduced by modeling procedure in Section 4 and conduct robustness analyses in Section 5. Section 6 concludes this paper.

1 Dividend-Price Decomposition

We begin by specifying the standard definition relating return, future price, and dividend payment. Define gross investment return over one period as:

$$(1) \qquad \begin{aligned} 1+R_{t+1} &= \frac{P_{t+1}+D_{t+1}}{P_t} \\ &= \frac{P_{t+1}}{P_t}\left(1+\frac{D_{t+1}}{P_{t+1}}\right), \end{aligned}$$

where P_t and P_{t+1} denote respectively the stock prices at the begin and the end of the period, and R_{t+1} and D_{t+1} denote respectively the net return over the period and dividend payment at the end of the period, abstracting from splits and distributions other than dividends.[5] This may seem like a strange way to write the return, since we normally would look at gross capital gains P_{t+1}/P_t and dividend yield D_t/P_t, with information known at the beginning of the period in the denominator, but we are simply manipulating accounting identities, and taking P_t in both denominators would give us a telescoping series in which the final term does not vanish. Taking logs on both sides, equation (1) becomes:

$$(2) \qquad \log(1+R_{t+1}) = \log\left(\frac{P_{t+1}}{P_t}\right) + \log(1+\delta_{t+1}),$$

[5] For empirical work, we might treat all dividend payments during the period as coming at the end of the period, or we might try to do some more correct return calculation taking into account the timing of the dividends and the returns in sub-periods. In practice it probably doesn't matter much.

where $\delta_{t+1} \equiv \log(D_{t+1}/P_{t+1})$. We will approximate equation (2) by a first-order Taylor series expansion around $\delta_{t+1} = \delta$. Traditionally, δ is taken to be the long-term mean of the LDPR $\log(D_t/P_t)$, but we will take a broader view and think about δ being some reasonable value, since we will present some evidence that the long-term average may not exist. Letting $\rho \equiv 1/(1+\exp(\delta))$, then

$$\begin{aligned}
\left.\frac{d\log(1+\exp(\delta_{t+1}))}{d\delta_{t+1}}\right|_{\delta_{t+1}=\delta} &= \left.\frac{\exp(\delta_{t+1})}{1+\exp(\delta_{t+1})}\right|_{\delta_{t+1}=\delta} \\
&= 1-\rho.
\end{aligned}$$

Therefore, letting $\kappa \equiv \log(1+\exp(\delta)) - (1-\rho)\delta$, the Taylor approximation is:

$$\begin{aligned}
(3) \quad \log(1+R_{t+1}) &= \log\left(\frac{P_{t+1}}{P_t}\right) + \log(1+\exp(\delta_{t+1})) \\
&\approx \log(\frac{P_{t+1}}{P_t}) + \log(1+\exp(\delta)) + (1-\rho)(\delta_{t+1}-\delta) \\
&= \log(P_{t+1}) - \log(P_t) + \log(1-\exp(\delta)) \\
&\quad +(1-\rho)(\log(D_{t+1}) - \log(P_{t+1})) - (1-\rho)\delta \\
&= \kappa + \rho\log(P_{t+1}) + (1-\rho)\log(D_{t+1}) - \log(P_t).
\end{aligned}$$

We follow Campbell and Shiller (1988) in being informal about the sense of the approximation; we will take an empirical approach to determine how well the approximation works. We can rewrite equation (3) as

$$(4) \quad \log\left(\frac{D_t}{P_t}\right) \approx -\kappa + \log(1+R_{t+1}) + \rho\log\left(\frac{D_{t+1}}{P_{t+1}}\right) - \Delta\log(D_{t+1}).$$

Substituting the same for $t+1$, $t+2$, and so forth for the $\log(D_{t+1}/P_{t+1})$ on the right-hand side telescopes to imply:

$$(5) \quad \log\left(\frac{D_t}{P_t}\right) \approx -\frac{\kappa}{1-\rho}(1-\rho^{T-t}) + \sum_{s=t+1}^{T} \rho^{s-t-1}(\log((1+R_s) - \Delta\log(D_s)) + \rho^{T-t}\log\left(\frac{D_T}{P_T}\right).$$

This is the essential relationship that we will work with. Since $\rho < 1$ it is at least plausible to argue (as do Campbell and Shiller) that the final term should vanish as T increases, and we have the asymptotic expression

$$\log\left(\frac{D_t}{P_t}\right) \approx -\frac{\kappa}{1-\rho} + \sum_{s=t+1}^{\infty} \rho^{s-t-1}(\log((1+R_s) - \Delta\log(D_s)), \tag{6}$$

often referred to in the literature as *the accounting identity*. This says that, subject to the quality of the approximation, today's log dividend-price ratio $\log(D_t/P_t)$ is *identically equal to* to a linear combination of future log returns $\log(1+R_s)$ and future changes in log dividends $\Delta\log(D_s)$. This implies that the log dividend-price ratio must predict one or both of these. The puzzle in the literature is that the log dividend-price ratio seems to predict neither future log returns nor future log dividend changes.

2 Approximation Test

Before testing whether stock returns are predictable, we test the quality of the LDPR approximation in equation (5) with the annual prices and aggregate dividend payments of the S&P 500 index firms over 1871 to 2015.[6] We focus on annual data because monthly dividend payments are linearly interpolated from annual and quarterly dividend payments, and we do not want to deal with the approximation error this might entail. Over the same period, the average gross return is 10.56% with a standard deviation of 18.17% , and the average annual log gross return on the S&P 500 index is 8.61% with a standard deviation of 17.33%; the average dividend-price ratio is 4.47% with a standard deviation of 1.52% while the average log dividend price ratio is -3.18 with a standard deviation of 0.40; and the average annual log dividend growth rate is 4.37% with a standard deviation of 12.16%. The sample LDPR mean of 4.47% implies a ρ of 0.95.

Campbell and Shiller (1988) suggest a vector autoregression (VAR) approach to test the equation (5) without the final term instead of the conventional predictive regression and find that the LDPR series is persistent and is able to predict both future stock returns and future dividend growth but the associated R^2s in their tests are small. We replicate and confirm their results. Although Campbell and Shiller claim that the VAR procedure does a better job than single linear regressions to "detect

[6]The data is collected from Robert Shiller's website at `http://www.econ.yale.edu/~shiller/data.htm.` We thank Robert Shiller for making it available online.

long-term deviations of stock prices from the 'fundamental value' ", the VAR approach suffers several shortcomings. First, a VAR procedure with limited number of lags does not sufficiently capture the long-term relationship among current dividend-price ratio, future returns and future dividend growth rates. Cochrane (2011) shows that VAR estimates can be biased and significantly different from that in the true linear regressions. Second, the calculations of a VAR procedure will be much more complicated for more than two variables with even very limited lags. Finally, this procedure ignores the final term in the equation. To conclude, the analysis in Campbell and Shiller (1988) does not tell us whether (5) holds empirically.

An improved approach to avoid such shortcomings is to conduct a true linear regression of log dividend-price ratio on $2(T-t)$ terms of discounted log return and dividend growth plus one final term. This procedure, however, is burdensome and may not be implementable when $(T-t)$ is large and the sample period is not sufficiently long. In this study, we propose a parsimonious regression in the form of equation (5), that is regressing current LDPR on the sum of weighted future returns, the sum of weighted dividend growth rates and the log dividend-price ratio in the last period:

$$\text{(7)} \quad \log\left(\frac{D_t}{P_t}\right) = \alpha + \beta_1 \left(\sum_{s=t+1}^{T} \rho^{s-t-1} (\log(1+R_s))\right) + \beta_2 \left(\sum_{s=t+1}^{T} \rho^{s-t-1} \Delta \log(D_s)\right) + \beta_3 \left(\rho^{T-t} \log(\frac{D_T}{P_T})\right) + \varepsilon_t.$$

By construction, this regression overcomes the shortcomings in both conventional linear and VAR estimations and is more parsimonious. If the approximation of equation (5) is effective, we should expect that β_1 and β_3 in equation (7) are close to one and β_2 is close to negative one, and the corresponding R^2 is near 100%.

We take $(T-t)$ to be 30 years, which is reasonably long and gives us 115 overlapping observations (years) for analysis. The results are reported in the first column in Table 1 and suggest that the LDPR approximation in equation (5) is effective. The coefficient on the sum of discounted return is positive and close to one, and the coefficient on the sum of discounted dividend growth is approximately equal to minus one. All coefficients are statistically significant at 1% level and the corresponding R^2 is as high as 99%. This regression uses Newey-West standard errors to adjust for

Table 1: Approximation Test

This table reports the empirical results of whether the approximation of log dividend-price ratio in equation (5) is effective. The results are based on the annual prices of and dividend payments on the S&P 500 index from 1871 to 2015. The $(T-t)$ is set to be 30 years. The associated Newey-West standard error with four lags are in parentheses. *** denotes statistical significance at the 1% level.

	Model 1	Model 2
α	-2.96^{***}	-3.64^{***}
	(0.05)	(0.08)
β_1	0.96^{***}	0.88^{***}
	(0.02)	(0.08)
β_2	-0.99^{***}	-1.08^{***}
	(0.02)	(0.08)
β_3	1.19^{***}	
	(0.05)	
N	115	115
Adj-R^2 (%)	98.91	82.25

serial correlation (including that due to overlapping observations) and heteroscedasticity.[7] The high R^2 suggests that current LDPR predicts at least one regressor but not which one(s). Note that we have not adjusted for spurious regression bias (caused by low frequency series on both sides). For now, it suffices to note that the fit is very good. We will correct for spurious regression bias when we conduct predictive regressions.

To address the concern raised by Kleidon (1986), Marsh and Merton (1986), and Merton (1987), we further test whether the final term in equation (5) is small by repeating the analysis on equation (7) after dropping the term. The results are reported in the second column of Table 1 and suggest that the log dividend-price ratio in the final period (30 years from now) is not trivial but also not so large. The coefficients on the sum of discounted returns and dividend growth are still significant but slightly deviate from 1 (or -1), and the R^2 drops significantly by 17%, from 99% to 82%, evidence that equation (7) is a good specification for empirical estimating of log dividend-price ratio.

[7] We report the Newey-West standard errors with 4 lags. In an untabulated table, we find that the Newey-West standard errors are similar based on 10, 20 or 30 lags.

3 Predictability Test

Equation (5) suggests a predictive relationship of current LDPR on cumulative future log returns or cumulative log dividend growth rates (for example, Campbell and Shiller, 1988; Cochrane, 2008). Furthermore, equation (5) suggests that the true predictive tests should be conducted by reversing the dependent and independent variables in equation (7) as:

$$(8) \quad \sum_{s=t+1}^{T} \rho^{s-t-1}(\log(1+R_s)) = \alpha + \beta_1 \log\left(\frac{D_t}{P_t}\right) + \mu_T.$$

This sort of regression was presented in Cochrane (2008), although with a subset of our time period (1926-2004 rather than 1871-2015); Cochrane's time period is similar to the second half of our time period described in our subsample results in Section 5.2.

A significant β_1 and a reasonable level of R^2 would suggest a predictable relationship between the LDPR and the cumulative future returns. This specification applies to the predicability of future log dividend growth rates. The use of cumulative present values of the predicted variable in future periods has advantages against conventional predictive specification, in which one-period leading predicted variable is mostly used, in that it can capture the total predictability of future return or dividend growth. In other words, equation (8) captures both short-run and long-run return predictabilities (if any).

Equation (5) implies that the sum of all coefficients on the log dividend-price ratio (β_1s) across all three predictive tests (i.e. predictability tests of $\sum_{s=t+1}^{T} \rho^{s-t-1}(\log(1+R_s))$, $\sum_{s=t+1}^{T} \rho^{s-t-1} \Delta\log(D_s)$ and $\rho^{T-t}\log(\frac{D_T}{P_T})$) should be one if our predictive specification is appropriate.[8] However, it is not possible to explore this relationship with three conventional predictive regressions. Granger and Newbold (1974), Stambaugh (1999), and Ferson, Sarkissian, and Simin (2003) show that linear regressions with lagged stochastic regressors with finite sample suffer spurious regression bias (SRB). Stambaugh shows that this bias is pronounced in the predictive coefficient but not in the standard error of the predictive coefficient and R^2. By assuming $\log\left(\frac{D_t}{P_t}\right)$ a first-order autoregressive process as $\log\left(\frac{D_t}{P_t}\right) = c + \tau\log\left(\frac{D_{t-1}}{P_{t-1}}\right) + \nu_t$, Stambaugh shows that the magnitude of SRB in the predictive coefficient in equation (8) equals $-\frac{\sigma_{\mu\nu}}{\sigma_\nu^2}(\frac{1+3\tau}{N})$, where $\sigma_{\mu\nu}$ is the covariance of μ_t and ν_t, σ_ν^2 the

[8] This relationship is also perceived by Cochrane (2008, 2011).

variance of v_t, and N the number of observations of the sample. Newey and West (1987) propose an adjustment in the standard error to overcome the serial correlation in the error term μ_T. In this section, we respectively follow Newey and West (1987), and Stambaugh (1999) to calculate the Newey-West standard errors, and the SRB-adjusted coefficients of the predictor (i.e. the LDPR).[9]

Table 2: Predictability Test

This table reports the empirical results of whether current log-dividend-price ratio is able to predict the sum of discounted future returns, the sum of discounted future dividend growths, or discounted log dividend-price 30 years from now. The results are based on the annual data of the S&P 500 index from 1871 to 2015. The spurious regression bias (SRB) is estimated following Stambaugh (1999). The associated Newey-West standard error with four lags are in parentheses. ***, ** and * denote statistical significance at the 1% , 5% and 10% levels, respectively.

Predicted variable	α	β_1	SRB-adjusted β_1	Adj-R^2 (%)
$\sum_{s=t+1}^{T} \rho^{s-t-1}(\log(1+R_s))$	1.98^{**}	0.19	0.17	0.43
	(0.76)	(0.24)	(0.24)	
$\sum_{s=t+1}^{T} \rho^{s-t-1}\Delta\log(D_s)$	-1.26^{*}	-0.60^{***}	-0.56^{***}	14.79
	(0.66)	(0.20)	(0.20)	
$\rho^{T-t}\log(\frac{D_T}{P_T})$	-0.19	0.17^{**}	0.17^{**}	17.31
	(0.19)	(0.07)	(0.07)	

Table 2 shows that the sum of all LDPR coefficients from the three predictability tests of $\sum_{s=t+1}^{T} \rho^{s-t-1}(\log(1+R_s))$, $\sum_{s=t+1}^{T} \rho^{s-t-1}\Delta\log(D_s)$ and $\rho^{T-t}\log(D_T/P_T)$ is close to one $(0.19-(-0.60)+0.17=0.96)$ and the corresponding sum of SRB-adjusted predictor coefficients is similar (0.90), evidence that our predicative specifications are theoretically appropriate and that spurious regression bias is not severe in our analysis. More interestingly, Table 2 shows that the LDPR (predictor) coefficient (row 1) in the predictability test of cumulative future returns is insignificant while it is significant in the predictability tests of cumulative future dividend growth (row 2) and the discounted final-period log dividend-price ratio (row 3). The size of the coefficient on LDPR in the predictability test of future dividend growth rates dominates the other two counterparts and contributes 63% $(0.60/0.96)$ to the sum of all three predictor coefficients. The patterns remain after being adjusted for the spurious bias in the predictive coefficients. The predictor coefficient in the predictability test

[9]We report the Newey-West standard errors with 4 lags in Table 2. The Newey-West standard errors are similar based on 10, 20 or 30 lags.

of future returns contributes 20% to the sum of all three predictor coefficients. Similar patterns hold in the R^2s for both coefficients.

The results provide significant evidence to reject the hypothesis that future dividend growth rates are not predictable but not the hypothesis that future return rates are not predictable. The main reason, as will be soon shown below, is that stock return generating process is too noisy to be predictable, consistent with Shiller (1981), and Poterba and Summers (1988) that stock prices are too volatile to be explained by fundamentals. The unpredictability of stock returns supplements the argument by Lanne (2002), Valkanov (2003), and Boudoukh, Richardson and Whitelaw (2008) that conventional analysis of long-term predictability of stock returns is spurious. Moreover, the results from Table 2 suggest that the relationship among the three predictor coefficients (Cochrane, 2008) is not informative about stock return predictability.

It may be surprising that the predictability of the LDPR 30 years out in Table 2 is both economically and statistically significant, and may lead us to ask what we know now about what will happen 30 years from now.[10] This view is compelling if we take the dividend process as an exogenous, but as Modigliani and Miller (1958) point out, dividends are somewhat arbitrary. Although our new information today may be primarily about cash flows in the coming ten years, this cash may go into repurchasing shares rather than dividends, with the actual dividend increase spread slowly over decades. Predictability of the LDPR 30 years out only depends on (1) predictability of cash flows over a short horizon, and (2) firm policies that imply it takes a very long time for these increased cash flows to appear in dividends. All of this is consistent with the smoothness of dividends as noted by Lintner (1956) and others over time.

In the meantime, it is worth further exploring why the coefficient on stock returns in Table 1 is consistently close to one and statistically significant but the predictor coefficient (on LDPR) in stock return predictability test in Table 2 is small and insignificant. Our explanation is that the coefficient of one in the first case is caused by high collinearity between stock return and dividend growth rather than information innovation in stock return generating process. To illustrate our argument, let start with equation (5) as $log(D_t/P_t) \approx \alpha + \beta_1(\sum_{s=t+1}^{T} \rho^{s-t-1}(\log(1+R_s))) + \beta_2(\sum_{s=t+1}^{T} \rho^{s-t-1}\Delta\log(D_s)) + \varepsilon_t$. Assume that the first term only contains noise (denoted as Z_t) and the second term contains both

[10] We must, of course, grant that the entire economic system may change over 30 years and the stock market may not exist.

information and noise as:

$$(9) \quad \sum_{s=t+1}^{T} \rho^{s-t-1}(\log(1+R_s)) \approx Z_t.$$

$$(10) \quad \sum_{s=t+1}^{T} \rho^{s-t-1}\Delta\log(D_s) \approx \log(\frac{D_t}{P_t}) + Z_t.$$

Then the covariance matrix between LDPR and sum of discounted future dividend growth rates becomes:

$$(11) \quad \mathrm{var}\left(\log\left(\frac{D_t}{P_t}\right), \sum_{s=t+1}^{T} \rho^{s-t-1}\Delta\log(D_s)\right) = \begin{pmatrix} \sigma_\delta^2 & -\sigma_\delta^2 \\ \sigma_\delta^2 & \sigma_\delta^2 + \sigma_Z^2 \end{pmatrix}.$$

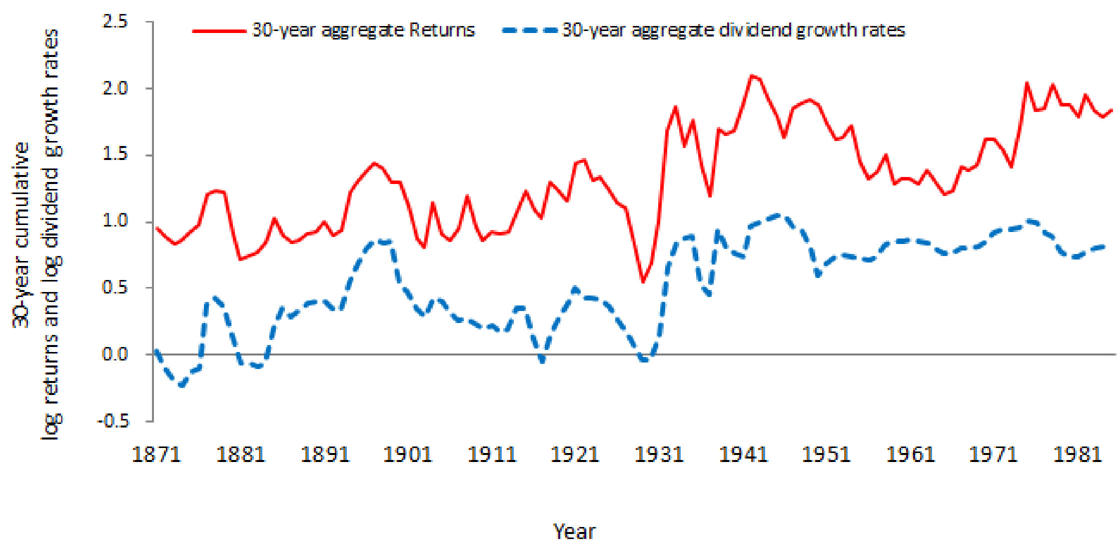
When we run current LDPR on the sum of discounted future dividend growth rates as $\log(\frac{D_t}{P_t}) = \alpha + \beta(\sum_{s=t+1}^{T} \rho^{s-t-1}\Delta\log(D_s)) + \mu_t$, the coefficient on the independent variable is given as: $\beta = \frac{-\sigma_\delta^2}{\sigma_\delta^2+\sigma_Z^2}$. When the noise σ_Z^2 in stock return is large, then the coefficient will be downward biased. Moreover, the correlation between the sum of discounted future returns and the sum of discounted future dividend growth also becomes large.

$$(12) \quad \mathrm{var}\left(\sum_{s=t+1}^{T} \rho^{s-t-1}(\log(1+R_s)), \sum_{s=t+1}^{T} \rho^{s-t-1}\Delta\log(D_s)\right) = \begin{pmatrix} \sigma_\delta^2 & \sigma_Z^2 \\ \sigma_Z^2 & \sigma_\delta^2 + \sigma_Z^2 \end{pmatrix}.$$

$$(13) \quad \mathrm{corr}\left(\sum_{s=t+1}^{T} \rho^{s-t-1}(\log(1+R_s)), \sum_{s=t+1}^{T} \rho^{s-t-1}\Delta\log(D_s)\right) = \frac{\sigma_Z^2}{\sigma_Z^2 + \sigma_\delta^2}.$$

Equation (9) implies that when we conduct a stock return predictability test as $\log(\frac{D_t}{P_t}) = \alpha + \beta(\sum_{s=t+1}^{T} \rho^{s-t-1}(\log(1+R_s))) + \mu_t$, the β coefficient should be close to zero and our untabulated empirical results confirm this. However, equation (13) suggests that when the noise is large, we may end up with a spurious coefficient significantly different from zero, which suggests that the approximation test of LDPR specified in equation (5) is different from the predictive tests specified in equation (8). Our explanation is consistent with the empirical data. Figure 1 shows that the evolutions of cumulative log returns and dividend growth rates are closely correlated. In fact, the correlation between the log return and the log dividend growth is 0.63, and the correlation between $(\sum_{s=t+1}^{T}\rho^{s-t-1}(\log(1+R_s)))$ and $(\sum_{s=t+1}^{T}\rho^{s-t-1}\Delta\log(D_s))$ is 0.84. Moreover, the standard deviations of the cumulative returns and log dividend growth rates are respectively 38.1% and 34.3% which implies that sum of variance of the two terms is as high as 37.2%, or 61.0% in terms of standard deviation, while the standard deviation of the log dividend-price ratio over the same period is 23.6%. This spurious relationship is also observed by Ferson, Sarkissian and Simin (2003) and Valkanov (2003) with simulated data.

Figure 1: Time Series of Cumulative Discounted Returns and Dividend Growth Rates.



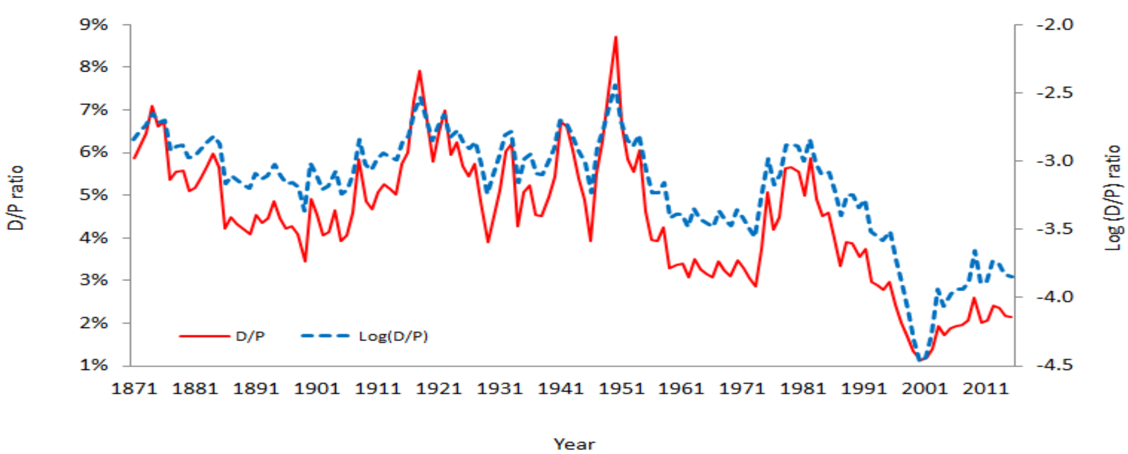
One concern about the results in this section is that our sample includes only about five non-overlapping observations of the weighted average of log differenced dividinds. Although it is impressive that the estimates (with Newey-West and Stambaugh corrections) are significant in spite of this, this puts a lot of demand on the Newey-West adjustment and we are far from its asymptotic justification.

4 Further Analysis

4.1 The Long-Term Mean Log Dividend-Price Ratio

The empirical analyses in the previous section are based on expanding LDPR around its sample mean. The literature typically assumes that the long-term mean exists and that the expansion is around the long-term mean. Unfortunately, Campbell and Shiller do not correctly test the existence of the long-term mean, because the long-term mean does not exist either under the null or under the alternative of their test because both include trends. In a corrected version of their test without a trend, we cannot reject the null of non-stationarity under which the long-term mean LDPR does not exist. This certainly weakens the interpretation of the sample mean as the long-term mean, but it doesn't necessarily invalidate the analysis using the current data, as we will discuss in the next Section. The time series of dividend-price ratio of S&P 500 index (solid line) and the corresponding log ratio (dash line) are plotted in Figure 2.

Figure 2: Time Series of $\frac{D}{P}$ and $log(\frac{D}{P})$.



Both show a strong declining trend. Specifically, the dividend-price ratio over Campbell-Shiller period (1871-1986) is around 5%, and declines to around 2% over post Campbell-Shiller period (1987-2015), and becomes lower than 2% over the last 15 years in the 21^{st} century. Figure 2 suggests that the long-term mean of dividend-price or log dividend-price may not exist. To formally test the existence of LDPR long-term mean, we follow the conventional stationarity test in which the null is that LDPR is a non-stationary process. Although Campbell and Shiller claim to conduct a stationarity test, both their null and alternative hypotheses contain trends and hence theirs is not a test of stationarity. Our stationarity test is specified as $\log(D_t/P_t) = \alpha + \beta \log(D_{t-1}/P_{t-1}) + \varepsilon_t$ and the trend assumption is removed. The results are reported in Table 3.

Table 3: Stationarity Tests

This table reports the empirical results of whether the annual series of log dividend-price ratio of the S&P 500 index is stationary over the whole sample period, the Campbell-Shiller period, and the post Campbell-Shiller period. The stationarity test is specified as $\log(D_t/P_t) = \alpha + \beta \log(D_{t-1}/P_{t-1}) + \varepsilon_t$. The whole sample period is from 1871 to 2015 and the Campbell-Shiller period is from 1871 to 1986.

$log(\frac{D}{P})_t$	$1871-2015$ (Entire sample)	$1871-1986$ Campbell-Shiller	$1987-2015$ Post Campbell-Shiller
α	-0.16	-0.39	-0.34
	(0.06)	(0.09)	(0.15)
β	0.89	0.71	0.82
	(0.04)	(0.07)	(0.09)
Dicky-Fuller stat	-15.78	-32.95	-4.76
Dicky-Fuller critical	-16.30	-16.30	-14.60
N	139	111	26
Adj$-R^2$	77.14	49.55	72.37
Reject unit root	No	Yes	No

Consistent with Figure 2 that the dividend payment on the S&P 500 index declines over time, the LDPR time series is stationary over Campbell-Shiller period but non-stationary over the whole sample period and post-Campbell-Shiller period. The Dicky-Fuller statistic is -15.8 over the whole sample period, -33.0 over the Campbell-Shiller sample period, and -4.8 over the post Campbell-Shiller period, and the corresponding critical values at 5% level are respectively -16.3, -16.3 and -14.6. In short, we fail to reject the hypothesis that LDPR is a non-stationary series (without a long term mean). If LDPR does not have a long-term mean, the sample mean cannot be an estimate of the long-term mean (which does not exist). Practically speaking the failure to reject non-stationarity says that we don't know whether the long-term mean exists - perhaps it exists but the test does not have enough power or is misspecified - but even if it does exists, lack of power in the test suggests we do not have a long enough data series to get a good estimate. It will certainly be a problem over time if the long-term mean does not exist and the LDPR gets more and more dispersion that will make the Taylor approximation worse and worse. However, a Taylor expansion around the sample mean may still be useful with the sample we have, and that is what we test next. In fact, we will find that our main results are not sensitive to the choice of δ in a reasonable range. This also says our results are not being caused by a look-ahead bias due to constructing the covariates using the sample mean LDPR for the whole period.

4.2 Approximation Error in a Single Period

Previous analyses show that stock return predictability is buried by noise but the LDPR approximation around its sample mean is helpful and appropriate. The seemly controversial results suggest that the noise contained in stock return series is endogenous as illustrated in Section (3) rather than introduced by the LDPR approximation. It is interesting to understand whether and how much noise is introduced in the LDPR series generating process in equation (5). For this purpose, we start examining the noise in one-period approximation and then the cumulative noise in equation (5).

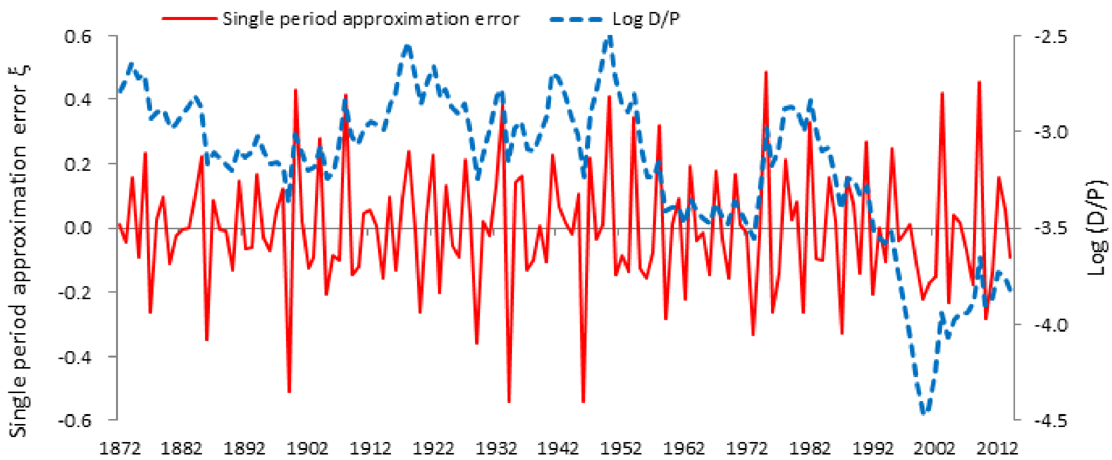
Note that equation (3) is a Taylor expansion of log return around the LDPR long-term mean. The expansion around sample mean in our empirical analyses may introduce a significant amount of noise when the long-run mean does not exist. We examine whether exogenous noise is added to our empirical approximation with equation (3) and define the noise term as:

$$(14) \quad \xi_t = \kappa + log(\frac{D_t}{P_t}) - \rho \log(\frac{D_{t+1}}{P_{t+1}}) + \log(\frac{D_t}{D_{t-1}}) - log(1+R_t),$$

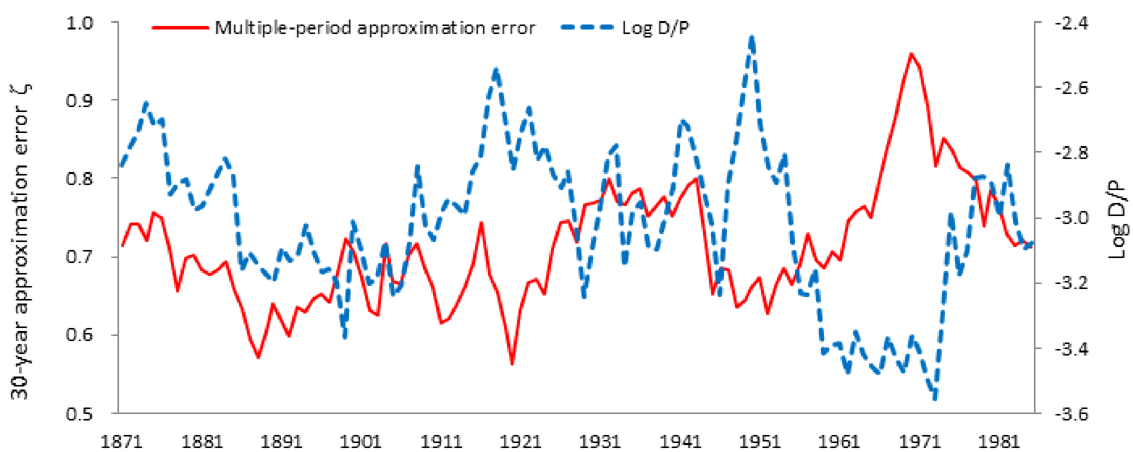
where ρ and κ are derived with the LDPR sample mean.

The time series of this single-period noise (ξ_t) is plotted in Figure 3.A and seems volatile. Although the mean of this noise term is close to zero (-0.37%) over the sample period from 1871 throughout 2015, its standard error is as high as 19.25%. An *ARMA* test suggests that the ξ_t series is an $AR(1)$ process. Although the magnitude of this noise is small relative to LDPR, which has a mean of -317.84% and a standard deviation of 40.44% over the same period, it is large relative to the log gross return, which has a mean of 8.61% and a standard deviation of 17.33%. Moreover, this noise will contaminate mainly stock return process because dividend growth process is more persistent. The results provide an explanation for findings in existing studies that short-term returns are not predictable by dividend-price ratio. In short, the approximation in equation (3) with the LDPR sample mean introduces a significant amount of noise in the process of returns and stock prices.

Figure 3: Time Series of Approximation Error



A single-period approximation error



B Multiple-period approximation error

4.3 Approximation Error in Multiple Periods

We further examine the cumulative approximation error in equation (5), which is defined as the following:

$$\zeta_t = -\frac{\kappa}{1-\rho}(1-\rho^{T-t}) + \sum_{s=t+1}^{T} \rho^{s-t-1}(\log((1+R_s) - \Delta \log(D_s)) + \rho^{T-t}\log(\frac{D_T}{P_T}) - log(\frac{D_t}{P_t}), \tag{15}$$

where ρ and κ are derived with the LDPR sample mean.

To be consistent with previous analyses, we take $(T-t)$ to be 30 and plot the evolution of ζ_t in Figure 3.B, which shows that the evolution of this noise term is smoother than the single-period noise and increases overtime. It has a sample mean of -1.5%, which is almost four times as large as that of one-period noise, and sample standard deviation of 3.5%, higher than that of the single-period noise either. An ARMA test shows that ζ_t is an $ARMA(1,1)$ process. In the meantime, the means of $log(\frac{D_t}{P_t})$,$\sum_{s=t+1}^{T} \rho^{s-t-1}(\log(1+R_s))$, $\sum_{s=t+1}^{T} \rho^{s-t-1}(\Delta \log(D_s))$ are respectively -3.02, 1.33 and 0.53, and their standard errors are 0.24, 0.38 and 0.34. Compared with the single-period noise, these numbers suggest that the approximation generated noise can be reduced by accumulating short-term returns. However, combining Figure 3 and Table 2 shows that our accumulating approach cannot reduce the endogenous noise in stock return generating process, suggesting that the lack of predictability of stock returns is caused by endogenous noise not by exogenously introduced noise.

5 Robustness Analysis

5.1 Alternative Expanding Point

In the log linear approximation in equation 3, we approximate $\log(1+\exp(\delta_{t+1}))$ around some value δ using a first-order Taylor expansion. Using a $3^{rd}-$order Taylor expansion, we have that $\log(1+\exp(\delta_{t+1})) \approx \log(1+\exp(\delta)) + (1-\rho)(\delta_{t+1}-\delta) + (1/2)\rho(1-\rho)(\delta_{t+1}-\delta)^2 + (1/6)\rho(1-\rho)(1-2\rho)(\delta_{t+1}-\delta)^3$, where the term beyond the first two are approximation error. Intuitively, the approximation is least severe if δ is in the middle of the range of δ_{t+1}'s, but how sensitive is the error to our choice? Figure 2 shows that the sample mean of dividend-price ratio is smaller than but close to 5% over the whole sample period and decreases to lower than 2% over years in the 21^{st} century. We first test whether the LDPR approximation in equation (5) is effective when it is expanded around alternative points. We consider four expanding points to take into account the declining trend in LDPR: 2%, 3%, 7% and 8% and two cases of approximation: with and without the final-period log dividend-price ratio. The empirical test is specified in the same way as equation (7) and results are reported in Table 4.

When LDPR is expanded around 2%, which is close to the payout ratio over years in the 21^{st} century, Panel A in Table 4 shows that the precision of the LDPR approximation is slightly impacted but still reasonably good. The coefficients are close to one (negative one) and the adjusted R^2 is around 94%. However, dropping the final term reduces the adjusted R^2 by 39% to a level of 55.6%, suggesting that final term is important when dividend payment is, relative to stock price, small. When LDPR is expanded around 3% (Panel *B*), the coefficients on sums of discounted future returns and dividend growth rates are very close to that in Table 1 and the adjusted R^2 is slightly reduced by 0.8%. The adjusted R^2 is reduced by about 15% when the final-period log dividend-price ratio in equation (7) is dropped from the LDPR approximation test. Panels *C* and *D* shows that the adjusted R^2 and the coefficients on cumulative discounted log returns, cumulative discounted log dividend growth rate, and final term are similar to that in the based case in Table 2 when LDPR is expanded around a relatively large point. These findings suggest that the approximation in equation (5) is effective when firm's payout is around a reasonable level, such as 5%, and that the log dividend-price ratio in the final period should not be dropped. In short, Table 4 suggests that expanding LDPR around its sample mean is meaningful and effective with current sample even though its long-term mean does not exist.

We further test whether the findings on the predictability tests of future returns and dividend growth rates in Section (3) hold for different discounting rates (corresponding to different expanding points). The empirical results are reported in Table 5 and support that the predictability of future dividend growth rates and the unpredictability of future stock returns are not impacted by expanding points. When LDPR is expand around the point of 0.02 (Panel *A*), which is close to the level in 2100s, the coefficient on current LDPR becomes negative and insignificant in the stock return predictability test suggesting that stock return is not predictable at all when LDPR becomes small. The magnitude and significance of the coefficient on LDPR in the predictability test of dividend growth rates are slightly improved, suggesting that dividend growth rate is still predictable even when LDPR is small.

Table 4: Approximation Tests: Alternative Expanding Points

This table reports the linear regression results of the Taylor expansion of log dividend-price ratio around alternative points. The regression is specified as: $\log(D_t/P_t) = \alpha + \beta_1(\sum_{s=t+1}^{T} \rho^{s-t-1}(\log(1+R_s))) + \beta_2(\sum_{s=t+1}^{T} \rho^{s-t-1} \Delta\log(D_s)) + \beta_3(\rho^{T-t}\log(D_T/P_T)) + \varepsilon_t$. The results are based on the annual data of the S&P 500 index from 1871 to 2015. The $(T-t)$ is set to be 30 years. The associated Newey-West standard error with four lags are in parentheses. *** denotes statistical significance at the 1% level.

	α	β_1	β_2	β_3	Adj-R^2 (%)
Panel A: Expanding point: 0.02($\rho \approx 0.98$)					
Model 1	-2.71^{***}	0.92^{***}	-0.87^{***}	0.82^{***}	93.79
	(0.08)	(0.04)	(0.04)	(0.05)	
Model 2	-3.60^{***}	0.62^{***}	-0.80^{***}		55.62
	(0.07)	(0.08)	(0.08)		
Panel B: Expanding point: 0.03($\rho \approx 0.97$)					
Model 1	-2.82^{***}	0.95^{***}	-0.93^{***}	0.92^{***}	98.12
	(0.04)	(0.02)	(0.02)	(0.05)	
Model 2	-3.66^{***}	0.74^{***}	-0.93^{***}		67.37
	(0.09)	(0.07)	(0.08)		
Panel C: Expanding point: 0.07($\rho \approx 0.94$)					
Model 1	-3.00^{***}	0.95^{***}	-1.02^{***}	1.53^{***}	96.17
	(0.08)	(0.04)	(0.04)	(0.14)	
Model 2	-3.58^{***}	0.93^{***}	-1.13^{***}		85.67
	(0.07)	(0.09)	(0.08)		
Panel D: Expanding point: 0.08($\rho \approx 0.93$)					
Model 1	-3.02^{***}	0.95^{***}	-1.03^{***}	1.80^{***}	94.36
	(0.09)	(0.05)	(0.05)	(0.22)	
Model 2	-3.54^{***}	0.94^{***}	-1.14^{***}		85.83
	(0.06)	(0.09)	(0.08)		

The results are almost unchanged when the expanding point is 0.03 (Panel *B*). When the value of expanding point increases to 0.07 (Panel *C*), or 0.08 (Panel *D*), the coefficient on current LDPR becomes positive but small and statistically insignificant in the predictability test of future returns, and it is large and significant in the predictability test of future dividend growth rates.

5.2 Subsample Period Analysis

Our second robustness test considers whether our findings over the whole sample period (from 1871 to 2015) hold over shorter subsample periods. After computing the 30-year averages, we split the whole sample period into two equally long subsample periods and repeat our predictability tests each period. Table 6 sums the predictability results for the two subsample periods, corresponding to the results in Table 2 for the whole sample. In the first subsample from 1871 to 1928 (Panel A), we see that cumulative discounted returns are not significantly predictable by current LDPR but cumulative discounted dividend growth is significantly predictable, consistent with the test over the whole sample period. In this subperiod, coefficient of the LDPR for the predictability of LDPR in 30 years is negative and insignificant, in contrast to the significant positive coefficient in the whole sample. The results on the second subsample from 1929 to 1985 (Panel B) are much different from the results in the whole sample. In this subsample, the coefficient of the LDPR in the predictability of the sum of discounted dividend growth is small and insignificant, while cumulative discounted log returns are significantly predictable. We repeat the Newey-West correction substantially from 5 to 30 lags and the standard errors are almost unchanged. This is consistent with a similar test of Cochrane (2008, Section 7.2) on a similar sample period. Unfortunately, the half-periods have even fewer nonoverlapping observations (about 2 1/2 instead of about 5) than the whole sample period regressions presented in Section 3.

It seems unrealistic to think that economic relationships will remain stable over 30 years (Chen, 2009) left alone so on 100, which is a weakness of the whole literature. Certainly dividend policy has changed over time, we saw that on the test of stationarity of the LDPR in Section 4.1 and ”disappearing dividends” have been documented by DeAngelo, DeAngelo and Skinner (2004), and Brav, Gramham, Harvey and Michaely (2005). This could be why the final term is significant in the second period but not the first. As Modigliani and Miller (1958) emphasize, in a frictionless world, the mix between

Table 5: Predictability Tests: Alternative Expanding Points

This table reports the empirical results of whether current log dividend-price ratio is able to predict sums of discounted future returns or discounted dividend growth rates, or the discounted final-period log dividend-price ratio. The results are based on the annual data of the S&P 500 index from 1871 to 2015. The spurious regression bias (SRB) is estimated following Stambaugh (1999). The associated Newey-West standard error with four lags are in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Predicted variable	α	β_1	SRB-adjusted β_1	Adj-R^2 (%)
Panel A: Expanding point: 0.02 ($\rho \approx 0.98$)				
$\sum_{s=t+1}^{T} \rho^{s-t-1}(\log(1+R_s))$	1.78*	-0.08	-0.07	-0.75
	(0.98)	(0.32)	(0.32)	
$\sum_{s=t+1}^{T} \rho^{s-t-1}\Delta\log(D_s)$	-1.48^{*}	-0.76^{***}	-0.70^{***}	14.83
	(0.86)	(0.27)	(0.27)	
$\rho^{T-t}\log\left(\frac{D_T}{P_T}\right)$	-0.48	0.43**	0.41**	17.31
	(0.47)	(0.17)	(0.17)	
Panel B: Expanding point: 0.03 ($\rho \approx 0.97$)				
$\sum_{s=t+1}^{T} \rho^{s-t-1}(\log(1+R_s))$	1.87**	-0.08	-0.08	-0.87
	(0.90)	(0.29)	(0.29)	
$\sum_{s=t+1}^{T} \rho^{s-t-1}\Delta\log(D_s)$	-1.42^{*}	-0.71^{***}	-0.65^{***}	15.14
	(0.79)	(0.25)	(0.25)	
$\rho^{T-t}\log(\frac{D_T}{P_T})$	-0.36	0.32***	0.30***	17.31
	(0.35)	(0.12)	(0.12)	
Panel C: Expanding point: 0.07 ($\rho \approx 0.94$)				
$\sum_{s=t+1}^{T} \rho^{s-t-1}(\log(1+R_s))$	2.01*	0.28	0.26	2.36
	(0.67)	(0.21)	(0.21)	
$\sum_{s=t+1}^{T} \rho^{s-t-1}\Delta\log(D_s)$	-1.14^{*}	-0.53^{***}	-0.51^{***}	13.80
	(0.58)	(0.18)	(0.18)	
$\rho^{T-t}\log\left(\frac{D_T}{P_T}\right)$	-0.14^{***}	0.10***	0.10***	17.31
	(0.18)	(0.04)	(0.04)	
Panel D: Expanding point: 0.08 ($\rho \approx 0.93$)				
$\sum_{s=t+1}^{T} \rho^{s-t-1}(\log(1+R_s))$	2.01***	0.31	0.28	3.48
	(0.64)	(0.20)	(0.20)	
$\sum_{s=t+1}^{T} \rho^{s-t-1}\Delta\log(D_s)$	-1.08^{**}	-0.49^{***}	-0.46^{***}	13.14
	(0.54)	(0.17)	(0.17)	
$\rho^{T-t}\log\left(\frac{D_T}{P_T}\right)$	-0.09	0.08***	0.08***	17.31
	(0.08)	(0.03)	(0.03)	

dividends and share repurchase would be irrelevant, and probably we should not expect dividend policy to be stable in the actual economy. However, it seems unlikely that markets were very efficient during the unstable times in the first half of the sample but inefficient later. The reversal of the results on the subperiods makes us wonder about the size of all the tests (on the whole sample as well as on the subperiods). Of our whole sample of 145 years and discounted weighted sums over 30 years, we only have about five non-overlapping observations in the whole sample and even fewer over subperiods. The lack of robustness to subperiods may also be due to model instability over time.

Table 6: Predictability Test over Subsample Period

We equally split the whole sample period into two non-overlapping periods. This table reports the empirical results of whether current log dividend-price ratio is able to predict sums of discounted future returns or discounted dividend growth rates over future 30 years, or the discounted final-period log dividend-price ratio over each subsample period. The results are based on the annual data of the S&P 500 index from 1871 to 2015. The spurious regression bias (SRB) is estimated following Stambaugh (1999). The associated Newey-West standard error with four lags are in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Predicted variable	α	β_1	SRB-adjusted β_1	Adj-R^2 (%)
Panel A: First Subsample Period (1871-1928)				
$\sum_{s=t+1}^{T}\rho^{s-t-1}(\log(1+R_s))$	1.64^{*}	0.18	0.17	0.66
	(0.86)	(0.29)	(0.29)	
$\sum_{s=t+1}^{T}\rho^{s-t-1}\Delta\log(D_s)$	-2.21^{**}	-0.85^{***}	-0.81^{***}	33.01
	(0.91)	(0.31)	(0.31)	
$\rho^{T-t}\log\left(\frac{D_T}{P_T}\right)$	-0.74^{***}	-0.03	-0.03	0.33
	(0.16)	(0.05)	(0.05)	
Panel B: The Second Subsample Period (1929-1985)				
$\sum_{s=t+1}^{T}\rho^{s-t-1}(\log(1+R_s))$	4.00^{***}	0.75^{***}	0.69^{***}	34.15
	(0.39)	(0.13)	(0.13)	
$\sum_{s=t+1}^{T}\rho^{s-t-1}\Delta\log(D_s)$	0.59^{**}	-0.08	-0.07	-1.01
	(0.23)	(0.07)	(0.07)	
$\rho^{T-t}\log(\frac{D_T}{P_T})$	-0.25	0.17^{**}	0.16^{**}	25.09
	(0.20)	(0.07)	(0.07)	

6 Conclusion

Whether stock returns are predictable is important and challenging to both academia and industry. Campbell and Shiller (1988) present an interesting argument, based on accounting definitions and some approximations, that the log dividend-price ratio must predict future returns, future log dividend growth, or both. However, in past literature neither prediction is economically or statistically significant, and this has been viewed as an important puzzle. We check each step of this argument, from the accounting definition through the approximation to the statistical tests. The source of the failure to find a significant relationship arises from a mismatch between the small lags in the traditional tests and the many terms in the theoretical expression. When we conduct a test closer to the theoretical expression, with appropriate correction for serial correlation due to overlapping data, the possible heteroscedasticity, and spurious regression bias, we find that future log dividend growth is significantly predictable but future returns are not, thus resolving the puzzle.

While this is the best conclusion given the data currently available, this result does not seem to be robust for several reasons. There only a few (about five) non-overlapping observations of the truncated identity for the whole period, so we are asking a lot of the Newey-West adjustment. Also, the results are different on subperiods, which calls into question the reliance on asymptotic properties of the statistical estimates. Perhaps we should not expect stability of the dividend process over time, since according to Modigliani and Miller dividends are irrelevant, and we expect this to be true to first order even if other considerations such as taxes and transaction costs mean Modigliani-Miller should not be taken too literally. Possible nonstationarity, which we cannot reject on the whole sample or the second half of the sample, is a serious problem for the theory because the Taylor series approximation gets worse and worse as the range of the LDPR increases. For these reasons, it seems that the limitations of this approach may be intrinsic, and there are reasons to believe the accounting identity will not tell us any more about return predictability even as we get more and more data.

References

Ang, Andrew, and Geert Bekaert. 2007. Stock Return Predictability: Is it There? The Review of Financial Studies 20 (3): 651-707.

Binsbergen, Jules H. van, and Ralph S. J. Koijen. 2010. Predictive Regressions: A Present-Value Approach. The Journal of Finance 65 (4): 1439-1471.

Boudoukh, Jacob, Matthew Richardson, and Robert F. Whitelaw. 2008. The Myth of Long-Horizon Predictability. The Review of Financial Studies 21 (4): 1577-1605.

Brav, Alon, John R. Graham, Campbell R. Harvey, and Roni Michaely. 2005. ”Payout Policy in the 21st Century.” Journal of Financial Economics 77: 483-527.

Bris, David Le, William N., Goetzmann, and Sebastien Pouget. 2016. Testing Asset Pricing Theory on Six Hundred Years of Stock Returns: Prices and Dividends for the Bazacle Company from 1372 to 1846. Working Paper, Toulouse Business School, Yale University, University of Touluse.

Campbell, John Y., and John Ammer. 1993. ”What Moves the Stock and Bond Markets? A Variance Decomposition for Long-Term Asset Returns.” The Journal of Finance 65 (1):3-37.

Campbell, John Y., and Robert J. Shiller. 1988. The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors. The Review of Financial Studies 1 (3): 195-228.

Chen, Long. 2009. On the Reversal of Return and Dividend Growth Predictability: A Tale of Two Periods. Journal of Financial Economics 92: 128-151.

Cochrane, John H., 2011. Presidential Address: Discount Rates. The Journal of Finance 66 (4): 1047-1108.

Cochrane, John H. 2008. The Dog That Did Not Bark: A Defense of Return Predictability. The Review of Financial Studies 21 (4):1533-1575.

DeAngelo, Harry, Linda DeAngelo, and Douglus J. Skinner. 2004. ”Are Dividends Disappearing? Dividend Concentration and the Consolidation of Earnings.” Journal of Financial Economics 72: 425-456.

Golez, Benjamin, and Peter Koudijis. 2016. Four Century of Return Predictability. Working Paper, University of Notre Dame and Stanford University.

Fama, Eugene F., and Kenneth R. French. 1988. Dividend Yields and Expected Stock Returns. Journal of Financial Economics 22: 3-25.

Ferson, Wayne E., Sergei Sarkissian, and Timothy T. Simin. 2003. Spurious Regressions in Financial Economics? The Journal of Finance 58 (4): 1393-1413.

Goetzmann, William N., and Philippe Jorion. 1993. Testing the Predictive Power of Dividend Yields. The Journal of Finance 48 (2):663-679.

Goetzmann, William N., and Philippe Jorion. 1995. A Longer Look at Dividend Yields. The Journal of Business 68 (4):483-508.

Goyal, Amit, and Ivo Welch. 2003. Predicting the Equity Premium with Dividend Ratios. Management Science 49 (5): 639-654.

Goyal, Amit, and Ivo Welch. 2008. A Comprehensive Look at The Empirical Performance of Equity Premium Prediction. The Review of Financial Studies 21 (4):1455-1508.

Granger, C.W.J., and P. Newbold. 1974. ”Spurious Regressions in Econometrics.” Journal of Econometrics 2 :111-120.

Kleidon, Allan W., 1986. Variance Bounds Tests and Stock Price Valuation Models. The Journal of Political Economy 94 (5), 953-1001.

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.

Lanne, Markku. 2002. Testing the Predictability of Stock Returns. The Review of Economics and Statistics 84 (3): 407-415.

Lettau, Martin, and Sydney C. Ludvigson. 2005. Expected Returns and Expected Dividend Growth. Journal of Financial Economics 76: 583-626.

Lintner, John. 1956. ”Distribution of Incomes of Corporations Among Dividends, Retained Earnings, and Taxes.” The American Economic Review 46(2): 97-113.

Mankiw, Gregory N., David Romer, and Matthew D. Sharpiro. 1991. Stock Market Forecastability and Volatility. The Review of Economic Studies 58 (3), Special Issue: The Econometrics of Financial Markets, 455-477.

Marsh, Terry A., and Robert C. Merton. 1986. Dividend Variability and Variance Bounds Tests for the Rationality of Stock Market Prices. The American Economic Review 76 (3): 483-498.

Marsh, Terry A., and Robert C. Merton. 1987. Dividend Behavior for the Aggregate Stock Market. Journal of Business 60 (1): 1-40.

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

Modigliani, Frnaco, and Merton H. Miller, 1958. The Cost of Capital, Corporation Finance and the Theory of Investment." The American Economic Review 48(3): 261-297.

Nelson, Charles R., and Myung J. Kim. 1993. Predictable Stock Returns: The Role of Small Sample Bias. The Journal of Finance 48 (2): 641-661.

Poterba, James M., and Lawrence H. Summers. 1988. Mean Reversion in Stock Prices: Evidence and Implications. Journal of Financial Economics 22: 27-59.

Shiller, Robert J., 1981. Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? The American Economic Review 71 (3):421-436.

Stambaugh, Robert F., 1999. Predictive Regressions. Journal of Financial Economics 52: 375-421.

Valkanov, Rossen. 2003. Long-Horizon Regressions: Theoretical Results and Applications. Journal of Financial Economics 68: 201-232.