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 and SWUFE IFS

Huacheng Zhang

SWUFE IFS

SWUFE IFS

July 11, 2017

Background

The "Dynamic Accounting Identity" of Campbell and Shiller (1988) says that the log of the dividend-price ratio (LDPR) is linearly related with a weighted linear sum of future log returns and future log dividend growths. It is called an accounting identity because it comes from the accounting definition of returns, log-linearization, and a statistical assumption that the long-term average log dividend-price ratio exists so we can throw away the final term in a telescoping sum.

Their result presents something of a puzzle, since neither log returns nor log dividend growth seems to be predicted by the LDPR. We seek to solve the puzzle.

Cochrane (2008) in his "dog that doesn't bark" paper has a novel but unsatisfactory solution, imposing the model restriction to overfit the model. Using an empirical model more consistent with the data reverses his conclusion.

...not my dog...

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.[1]

Many times, we get into trouble because we ask the wrong question. 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 start with the single-period decomposition of returns, its log-linearization, accumulation over time, stationarity of log dividend-price that implies we can drop the final term in the limit, dropping the final term in the sample we have, and then econometric estimation.

---

[1] Metro-Goldwyn-Myers, 1976, *The Pink Panther Strikes Again*, movie.

Main Results

- The log linearization works very well in a single period, and out over 30 years including the final term ($R^2 \approx 99\%$).

- The Campbell-Shiller test for existence of the long-term mean LDPR doesn't make sense, since the long-term mean does not exist in either the null or the alternative (due to inclusion of a trend).

- A corrected Dickey-Fuller test without a trend does not reject nonstationarity, suggesting the long-term mean LDPR may not exist.

- However, the approximation works pretty well within the sample we have, going out 30 years ($R^2 \approx 83\%$), so that there should still be predictability in our sample even if nonstationarity is a problem in the long-run.

- In our sample, if we use a regression with the form of predictability in the theory (instead of one or a few lags as is usual in the literature), we find that the log dividend growth is significantly predictable and log returns are not.

- The results feature Newey-West adjustment for overlapping data and a Stambaugh adjustment for spurious regression bias.

...also not my dog, did not bark...

Cochrane (2008) in the "dog that didn't bark" paper argues that we should take seriously the model restriction and also the lack of predictability of log dividend growth to conclude that dividends are predictable. This argument is novel, but not very satisfying. We could equally well argue that log returns are not very predictable (as we see in the data and is suggested by theory), and conclude that log dividend growths must be predictable. Or, we could say that lack of evidence of predictability of either is evidence that there is a problem with the model (our usual interpretation if an over-identifying restriction is inconsistent with the data).

One-period definition of returns

$$(1)\ 1 + R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t}$$

$$= \frac{P_{t+1}}{P_t}\left(1 + \frac{D_{t+1}}{P_{t+1}}\right),$$

where R_{t+1} is return, P_{t+1} is price, and D_{t+1} is dividend.

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

where $\delta_{t+1} \equiv \log(D_{t+1}/P_{t+1})$. We will do a power series expansion around $\delta_{t+1} = \delta$.

Taylor Expansion

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

where $\rho \equiv 1/(1+\exp(\delta))$

$$(3)\ \log(1+R_{t+1}) = \log\left(\frac{P_{t+1}}{P_t}\right) + \log(1+\exp(\delta_{t+1}))$$
$$\approx \log(\frac{P_{t+1}}{P_t}) + \log(1+\exp(\delta)) + (1-\rho)(\delta_{t+1}-\delta)$$
$$= \kappa + \rho\log(P_{t+1}) + (1-\rho)\log(D_{t+1}) - \log(P_t),$$

where $\kappa \equiv \log(1-\exp(\delta)) - (1-\rho)\delta$

Rewriting and Telescoping

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

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

$$+ \rho^{T-t} \log\left(\frac{D_T}{P_T}\right).$$

or asymptotically (maybe) we have the "accounting identity"

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

Approximation test

We will focus on the finite approximation (5) rather than the asymptotic approximation (6), since we can evaluate the finite approximation in the finite sample we have. We use the following specification, with and without the final term (corresponding to β_3).

$$(7) \; \log\left(\frac{D_t}{P_t}\right) = \alpha + \beta_1 \left(\sum_{s=t+1}^{T} \rho^{s-t}(\log(1+R_s)) \right)$$

$$+ \beta_2 \left(\sum_{s=t+1}^{T} \rho^{s-t}\Delta\log(D_s) \right) + \beta_3 \left(\rho^{T-t}\log(\frac{D_T}{P_T}) \right) + \varepsilon_t.$$

We use $\delta = -3.18 = \log(4.16\%)$, the sample mean LDPR for the Approximation test and the prediction tests, which corresponds to $\rho = .960$. Our main results are not sensitive to this choice.

The approximation does very well with the final term ($R^2 = 99\%$), and well enough to be useful ($R^2 = 82.5\%$) without the final term.

Approximation test: results
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	1.01***	0.93***
	(0.02)	(0.08)
β_2	-1.04***	-1.14***
	(0.02)	(0.09)
β_3	1.19***	
	(0.05)	
N	115	115
R^2 (%)	98.94	82.57

Predictability Tests

We use tests matching the model (finite version), for example, for returns we have:

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

with similar equations except with log dividend growth or final term on the left-hand side. Using Newey-West corrected estimators and Stambaugh's correction for spurious regression bias, we find that the LDPR significantly predicts log dividend growth and the error terms, but not log returns.

Table 2: Predictability Tests

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% and 5% levels, respectively.

Predicted variable	α	β_1	SRB-adjusted β_1	R^2 (%)
$\sum_{s=t+1}^{T} \rho^{s-t}(\log(1+R_s))$	1.88**	0.19	0.17	1.30
	(0.72)	(0.23)		
$\sum_{s=t+1}^{T} \rho^{s-t}\Delta \log(D_s)$	-1.20^*	-0.57^{***}	-0.61	15.54
	(0.62)	(0.20)		
$\rho^{T-t}\log(\frac{D_T}{P_T})$	-0.19	0.17***	0.17	18.03
	(0.18)	(0.06)		

Why is the R^2 so different in the two directions?

Figure 1: Time Series of Cumulative Discounted Returns and Dividend Growth Rates
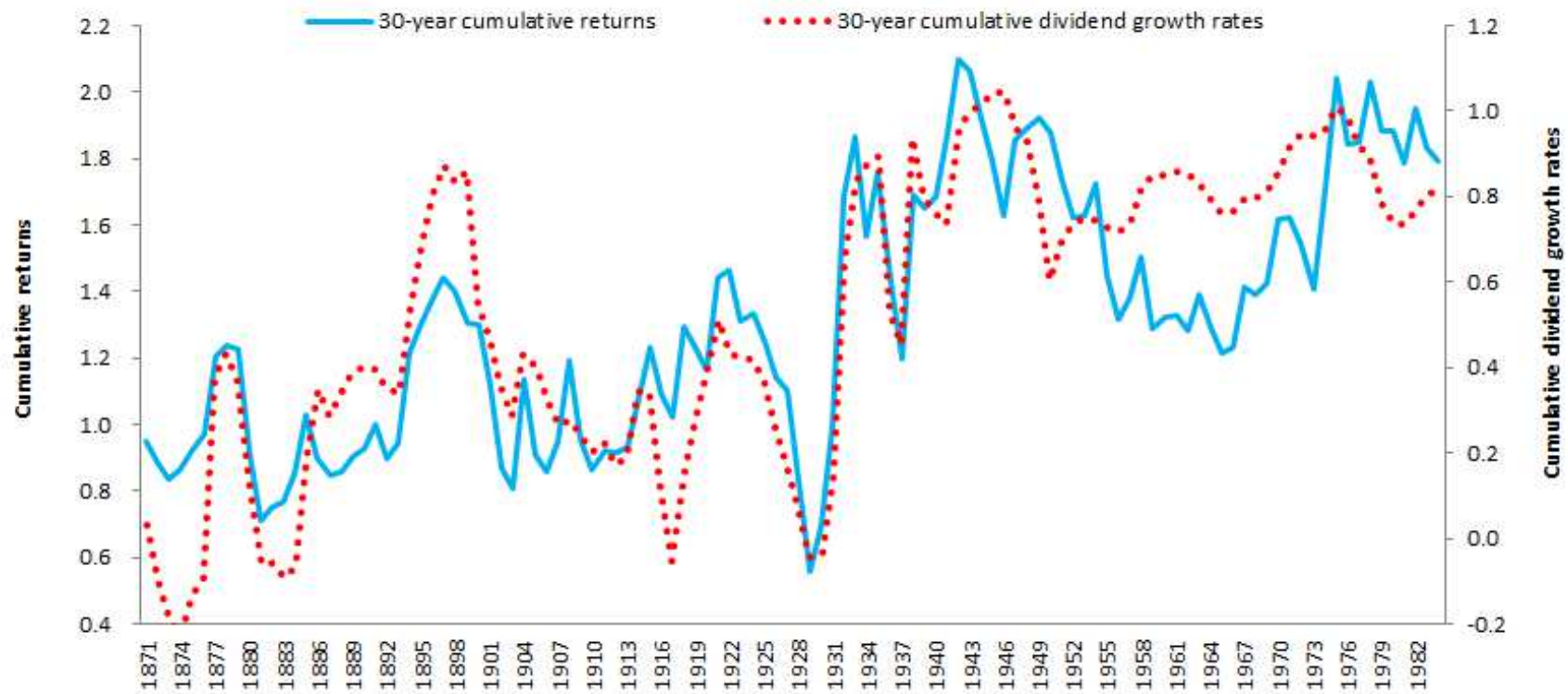
answer: because the cumulative log returns and the cumulative log dividend growth are highly correlated

Does the long-term mean LDPR exist?

The original analysis by Campbell and Shiller (1988) takes δ to be the long-term mean log dividend price ratio, but does it exist? There Campbell-Shiller tested for this, but their test doesn't make sense, since both the null and alternative include a trend, implying the long-term mean does not exist in either case. Specifically, they perform an augmented Dickey-Fuller test with the null being nonstationary plus a trend and the alternative being stationary plus a trend.

We perform a corrected test without a trend in both the null and the alternative. Using the full sample, we cannot reject a nonstationary process. Interestingly, given the data they had at the time, Campbell and Shiller would have rejected nonstationarity with their sample. The difference is that log dividend yields are overall lower in the subsequent years.

We view the apparent failure of their assumption that the long-term mean LDPR exists as more of a future problem rather than a current problem, since the finite approximation without the final term is adequate for the existing sample.

LDPR: having a look

Figure 2: Time Series of $\frac{D}{P}$ and $\log(\frac{D}{P})$
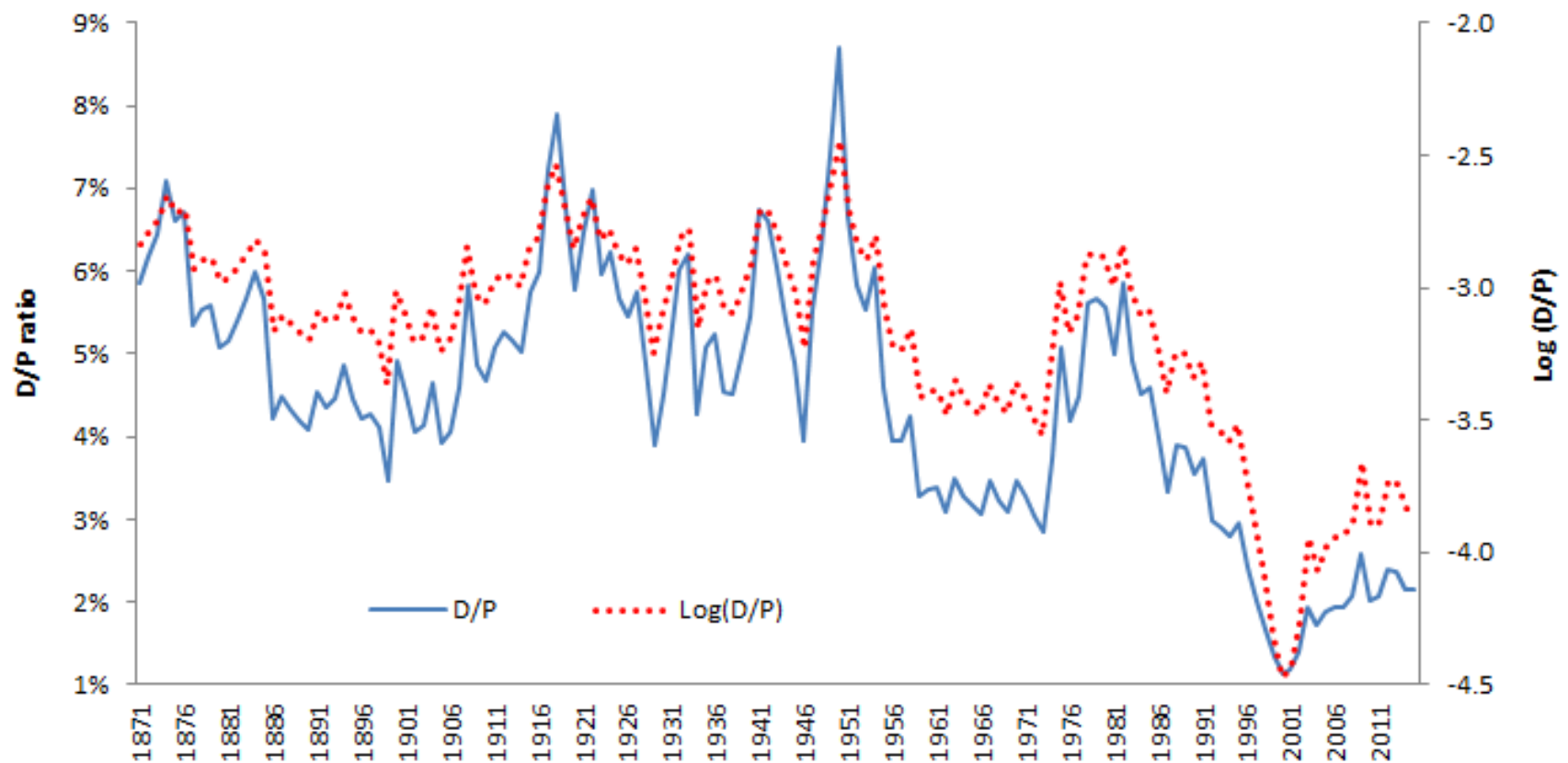
Dickey-Fuller Test

Table 3: Dickey-Fuller Test

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.

$\log(\frac{D}{P})_t$	$1871 - 2015$ (full 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

Robustness: expanding around different values of δ

Predicted variable	α	β_1	SRB-adjusted β_1	R^2 (%)
Panel A: Expanding point around $\delta = \log(0.03)$ ($\rho \approx 0.97$)				
$\sum_{s=t+1}^{T} \rho^{s-t}(\log(1+R_s))$	1.82**	0.03	0.01	0.03
	(0.86)	(0.28)		
$\sum_{s=t+1}^{T} \rho^{s-t}\Delta\log(D_s)$	-1.37^*	-0.68^{***}	-0.72	15.90
	(0.76)	(0.24)		
$\rho^{T-t}\log(\frac{D_T}{P_T})$	-0.35	0.31^{***}	0.31	18.03
	(0.34)	(0.12)		
Panel B: Expanding around $\delta = \log(0.02)$ ($\rho \approx 0.98$)				
$\sum_{s=t+1}^{T} \rho^{s-t}(\log(1+R_s))$	1.74^*	-0.08	-0.07	0.14
	(0.96)	(0.31)		
$\sum_{s=t+1}^{T} \rho^{s-t}\Delta\log(D_s)$	-1.45^*	-0.75^{***}	-0.80	15.58
	(0.85)	(0.27)		
$\rho^{T-t}\log\left(\frac{D_T}{P_T}\right)$	-1.45^{***}	0.75^{***}	0.76	15.58
	(0.50)	(0.16)		

Conclusion

- The log linearization works very well in a single period, and out over 30 years including the final term ($R^2 \approx 99\%$).

- The Campbell-Shiller test for existence of the long-term mean LDPR doesn't make sense, since the long-term mean does not exist in either the null or the alternative (due to inclusion of a trend).

- A corrected Dickey-Fuller test without a trend does not reject nonstationarity, suggesting the long-term mean LDPR may not exist.

- However, the finite approximation works pretty well within the sample we have, going out 30 years ($R^2 \approx 83\%$), so there should still be predictability in our sample even if nonstationarity is a problem in the long-run.

- In our sample, if we use a regression with the form of predictability in the theory (instead of one or a few lags as is usual in the literature), we find that the log dividend growth is significantly predictable and log returns are not.

- The results feature Newey-West adjustment for overlapping data and a Stambaugh adjustment for spurious regression bias.