Journal of Banking and Finance 78 (2017) 153–163

Contents lists available at ScienceDirect

Journal of Banking and Finance

journal homepage: www.elsevier.com/locate/jbf

Dividends, earnings, and predictability☆

Stig V. Møller*, Magnus Sander

CREATES, Aarhus University, Fuglesangs Allé 4, 8210 Aarhus, Denmark

ARTICLE INFO

Article history:
Received 25 January 2015
Accepted 22 February 2017
Available online 24 February 2017

JEL classification:
G12
G17
E44

Keywords:
Dividend yield
Earnings yield
Dividend growth predictability
Omitted variable bias

ABSTRACT

We show that the dividend yield and earnings yield jointly are strong predictors of dividend growth. We motivate the joint specification with a theoretical model and show how omitting the earnings yield biases the dividend yield coefficient towards zero, explaining why the dividend yield by itself is a poor predictor of dividend growth. Our empirical results are robust in pre- and post-war U.S. data, in recessions and expansions, in international data, and when controlling for additional predictors.

1. Introduction

Several papers have shown that the ability of the dividend yield to predict dividend growth is weak, see Campbell and Shiller (1988), Cochrane (2008, 2011), among others. We argue that the weak predictive power of the dividend yield for future dividend growth can be explained by a missing variable problem. The background is Lintner's (1956) dividend model from which it can be derived that the dividend yield and the earnings yield should be used jointly as explanatory variables of future dividend growth. Omitting the earnings yield from the equation causes the coefficient on the dividend yield to be biased towards zero. Together with the earnings yield, however, the dividend yield is a strong predictor of dividend growth. We show that this result is remarkably robust in both U.S. and international data.

We are not the first to show that the dividend yield predicts dividend growth when including the earnings yield in the specification. This intriguing result was first discovered by Ang and Bekaert (2007). They find that the dividend yield on its own contains only weak predictive power of dividend growth, but once they control for the earnings yield, the dividend yield is a significant predictor of dividend growth (and so is the earnings yield). Building on the work of Ang and Bekaert (2007), we show that their finding can be explained within the framework of Lintner's (1956) dividend model. Omitting the earnings yield conceals predictability by biasing the dividend yield coefficient towards zero.

We analyze and compare the strong predictive power of the dividend yield-earnings yield ($dy - ey$) model against a number of new findings in the literature. In recent years, an increasing number of papers have challenged the view that dividend yields do not predict dividend growth. Chen (2009) and Golez and Koudijs (2014) show that dividend yields predict dividend growth in the pre-war years, but not in the post-war years. This pre-war vs. post-war effect does not show up when using the $dy - ey$ model to predict dividend growth. While the dividend yield by itself has no predictive power for dividend growth in post-war data, the $dy - ey$ model contains substantial predictive power for dividend growth in both pre- and post-war data. We find that the omitted variable bias from not including the earnings yield is less severe in the pre-war period in part due to a lower correlation between dy and ey. In addition, as Chen et al. (2012) also show, there is less dividend smoothing in pre-war data. These two effects help explain the reversal in dividend growth predictability when using the dividend yield as the only predictor.

Engsted and Pedersen (2010) and Rangvid et al. (2014) show that dividend growth is predictable from dividend yields in countries with small market capitalizations but not in large markets such as the U.S. We examine a cross section of 14 developed countries and show that the $dy - ey$ model contains much more pre-

☆ We would like to thank seminar participants, participants at the 2013 OMI-SoFiE Financial Econometrics Summer School and, in particular, Geert Bekaert and two anonymous referees for useful comments and suggestions. We acknowledge support from CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation.

* Corresponding author.
E-mail addresses: svm@econ.au.dk (S.V. Møller), mdsj@econ.au.dk (M. Sander).

http://dx.doi.org/10.1016/j.jbankfin.2017.02.008
0378-4266/

dictive power for dividend growth than do univariate models with either the dividend yield or earnings yield as the only predictor. In line with the previous literature, we find that the $dy - ey$ model has more predictive power for dividend growth in countries with small market capitalizations compared to countries with large market capitalizations.

Recent literature has produced robust evidence that equity returns in the U.S. are substantially more predictable during economic downturns than during economic expansions, see Rapach et al. (2010) and Henkel et al. (2011). Hence, a natural question is whether the predictive power of the $dy - ey$ model varies with the state of the economy. Unlike return predictability, which seems to be restricted to a few periods around recessions, the $dy - ey$ model strongly predicts dividend growth in both expansions and recessions.

We also demonstrate that commonly used predictive variables such as the short rate, the term spread, the default spread, and the consumption-wealth ratio (cay) of Lettau and Ludvigson (2001) do not add much additional information about expected dividend growth beyond that contained in the $dy - ey$ model.

Overall, despite of its simplicity, the $dy - ey$ model contains robust predictive power for dividend growth in different subsamples, in both recessions and expansions, when controlling for additional variables, and across countries.[2]

The rest of the paper is structured as follows. Section 2 motivates why the dividend yield and earnings yield should be used jointly as predictive variables of dividend growth. Section 3 describes the data. Section 4 examines the U.S. evidence of dividend growth predictability, while Section 5 examines the international evidence. Section 6 concludes.

2. Motivation

This section motivates why the dividend yield and earnings yield should be used jointly as predictors of dividend growth. Consider Lintner's (1956) model of dividend payout in log form

$$\Delta d_{t+1} = a + c\left(d^*_{t+1} - d_t\right) + u_{t+1}, \tag{1}$$

where d_t is the actual log dividend at time t and $\Delta d_{t+1} = d_{t+1} - d_t$ is log dividend growth at time $t+1$. We specify the log target dividend as $d^*_{t+1} = r + e_{t+1}$, where r is the log target payout ratio and e_{t+1} is actual log earnings at time $t+1$.[3] The non-negative parameter c measures the speed of adjustment towards the target and reflects the degree of dividend smoothing. The model can be motivated by a quadratic cost function where managers are penalized for deviations of dividend growth from a normal rate as well as for deviations of realized dividends from target dividends, see Garrett and Priestley (2000). We next assume that e_{t+1} is well approximated by a random walk.[4] After rearranging, we then arrive at the following specification for dividend growth

$$\Delta d_{t+1} = \alpha - c((d_t - e_t) - r) + \nu_{t+1}. \tag{2}$$

Lintner's model can therefore be seen as a theoretical motivation for predicting dividend growth using the log payout ratio, $d_t - e_t$. If the current payout ratio is above the target level, dividend growth is expected to fall. We also see from (2) that the level of predictability is linked to the degree of dividend smoothing. If we then add and subtract $c \times p_t$, where p_t is the log price, and ignore the constant r, we obtain the following model

$$\Delta d_{t+1} = \alpha - c \times dp_t + c \times ep_t + \nu_{t+1}, \tag{3}$$

where $dp_t = d_t - p_t$ is the dividend yield and $ep_t = e_t - p_t$ is the earnings yield. Next, consider a miss-specified model that only includes the dividend yield

$$\Delta d_{t+1} = \alpha + \beta dp_t + \varepsilon_{t+1}, \tag{4}$$

where $\beta = -c$ and $\varepsilon_{t+1} = c \times ep_t + \nu_{t+1}$. If we estimate this model using OLS, we get an omitted variable bias due to not including the earnings yield. The bias is

$$E\left(\widehat{\beta}\right) - (-c) = \gamma c, \tag{5}$$

where γ is the slope coefficient from an auxiliary regression of ep_t on dp_t (and a constant). Rearranging, we get

$$E\left(\widehat{\beta}\right) = -c(1 - \gamma). \tag{6}$$

If the dividend yield and the earnings yield have a high correlation (γ close to 1), regressing dividend growth on the dividend yield could lead us to wrongly conclude that dividend growth is not predictable. The intuition is that dp_t and ep_t have the opposite sign in (3) but are positively correlated. Omitting ep_t pulls the estimated coefficient of dp_t towards zero. This point is not restricted to the Lintner model, but extends to other models of dividend behavior where potential omitted variables correlate with the dividend yield. The models of Marsh and Merton (1987) and Garrett and Priestley (2012) are interesting alternatives.[5]

3. Data

3.1. U.S. data

We use S&P 500 data to compute returns, dividend growth, dividend yields, and earnings yields. In our main regressions, we use a quarterly sample over the period 1927:1 to 2013:4. We compute returns on the S&P 500 index including dividends. To compute the excess return, we subtract a three-month T-bill rate. We derive monthly dividend payments from returns with and without dividends and compute annual dividends as the sum of dividend payments on the S&P 500 index over the past year. We compute dividend growth as the quarterly growth rate in annual dividends and the quarterly dividend yield is then given by the sum of dividends over the past year divided by the end-of-quarter price. In a similar vein, the quarterly earnings yield is defined as earnings over the past year divided by the end-of-quarter price. All variables are in logs.

We also work with annual data over the period 1871 to 2013. We have obtained both the quarterly and annual data from the updated Goyal and Welch (2008) dataset, which is available on Amit Goyal's website.[6]

3.2. International data

We also carry out an international analysis using data on the following 14 countries: Australia, Austria, Belgium, Canada, Denmark, France, Germany, Hong Kong, Japan, the Netherlands, Singapore, South Africa, Switzerland, and United Kingdom. For these

[2] Our results from simple OLS regressions confirm recent evidence of strong dividend growth predictability based on more advanced methods. Golez (2014) extracts a forward looking measure of expected dividend growth from options and futures and shows that it predicts dividend growth, while Binsbergen and Koijen (2010) use state-space models to show that past values of dividend growth help to forecast both returns and dividend growth.

[3] Lintner (1956) originally specified the model with all variables in levels. We follow Garrett and Priestley (2000) and Chen et al. (2012) by specifying the model in logs.

[4] Since the work of Ball and Watts (1972), several empirical studies have shown that earnings are close to a random walk, see the review in Kothari (2001).

[5] The focus of the paper is on predictability of dividend growth from dividend yields and we acknowledge that the dividend behavior model of Lintner (1956) does not provide guidance on predictability of returns from dividend yields. In particular, the model cannot be used to understand the joint dynamics of expected returns and expected dividend growth.

[6] The S&P earnings yield originates from Robert Shiller's website and the S&P Corporation.

Table 1
Dividend growth and return predictability from dividend yields and earnings yields. The table reports results from forecasting regressions, $y_{t+k} = \alpha_k + \beta_k^{dp} dp_t + \beta_k^{ep} ep_t + \varepsilon_{t+k}$, where y_{t+k} is the k-period ahead dividend growth rate or excess return. We use the dividend yield and the earnings yield as regressors, either separately or jointly. The frequency is quarterly and the sample period is from 1927:1 to 2013:4. Panels A and B report results for $k = 1$ and $k = 4$, respectively. Below OLS estimates of β_k are t-statistics in parentheses based on Newey-West standard errors with $k+3$ lags and in brackets are t-statistics based on standard errors from a stationary bootstrap. We resample data in blocks of random size and determine the average block-size using the Politis and White (2004) automatic selection procedure. $\bar{R}^2$ is the adjusted R-squared.

Dividend growth			Returns		
dp	ep	$\bar{R}^2$	dp	ep	$\bar{R}^2$
Panel A: $k=1$					
−0.018			0.021		
(−1.57)		3.4%	(1.27)		0.5%
[−1.12]			[1.13]		
	0.013			0.028	
	(2.09)	1.5%		(1.70)	0.9%
	[1.98]			[1.81]	
−0.059	0.058		0.004	0.025	
(−3.36)	(4.05)	19.5%	(0.12)	(0.75)	0.7%
[−2.42]	[3.11]		[0.11]	[0.81]	
Panel B: $k=4$					
−0.081			0.093		
(−1.56)		7.5%	(1.86)		3.3%
[−1.14]			[1.97]		
	0.029			0.102	
	(0.98)	0.6%		(2.13)	3.7%
	[1.08]			[2.08]	
−0.220	0.197		0.046	0.067	
(−3.07)	(3.85)	26.9%	(0.47)	(0.70)	3.9%
[−2.10]	[3.16]		[0.47]	[0.64]	

countries, we obtain returns, dividend growth, dividend yields, and earnings yields from Datastream over the period 1973:1 to 2013:4. We denote these variables in local currency. Similar to Rangvid et al. (2014), we use the individual country data to generate global aggregate portfolios of returns and dividend growth as well as corresponding global dividend and earnings yields. We use both an equal-weighted global portfolio and a value-weighted global portfolio based on U.S. dollar market capitalizations.

4. Evidence from the United States

Table 1 shows results from forecasting regressions

$$y_{t+k} = \alpha_k + \beta_k^{dp} dp_t + \beta_k^{ep} ep_t + \varepsilon_{t+k}, \tag{7}$$

where y_{t+k} is the k-period ahead dividend growth rate or excess return. We use the dividend yield dp_t and the earnings yield ep_t as regressors, either separately or jointly. The data are quarterly and the sample period is from 1927:1 to 2013:4. Panels A and B report results for $k = 1$ and $k = 4$, respectively.[7] Below OLS estimates of slope coefficients are t-statistics in parentheses based on Newey-West standard errors with $k + 3$ lags and in brackets are t-statistics based on standard errors from a stationary bootstrap. We resample the data in blocks of random size and determine the average block-size using the Politis and White (2004) automatic selection procedure. We set the number of replications to 50,000 and compute the bootstrap standard error on β_k as the standard deviation of the 50,000 simulated slope coefficients.

Focusing first on univariate regressions, the table shows that the dividend yield is statistically insignificant in predicting dividend growth at the 5% level. The earnings yield is marginally significant in predicting dividend growth for $k = 1$ but turns insignificant for $k = 4$. In addition, the two variables only display weak predictive power for returns.

Turning to the bivariate regressions, we find no increase in return predictability by using the dividend yield and earnings yield jointly. However, together they contain substantial predictive power for dividend growth. Consistent with Ang and Bekaert (2007), the dividend yield coefficient is negative and the earnings yield coefficient is positive. Comparing the univariate and bivariate one-quarter ahead regressions, the coefficient on dp_t goes from −0.018 to −0.059 and the coefficient on ep_t goes from 0.013 to 0.058. In addition, both variables are now statistically significant at the 1% level. The impact on the adjusted R^2 is also striking. dp_t on its own generates an $\bar{R}^2$ of 3.4%, ep_t an $\bar{R}^2$ of 1.5%, but together they generate an $\bar{R}^2$ of 19.5%. Similar patterns appear at the one-year ahead forecast horizon.

These empirical results can be explained within the framework of Lintner's (1956) dividend model. From Eq. (3), we see that the model implies a negative dividend yield coefficient and a positive earnings yield coefficient, as we find in the data. The model in (3) restricts the dividend yield and earnings yield coefficients to be equal but of opposite sign. In the bivariate regressions, we have allowed the variables to have separate coefficients. In accordance with the implications of the Lintner model, the estimation results show that the dividend yield and earnings yield coefficients are of opposite sign and of similar magnitude. For instance, with $k = 1$, the dividend yield coefficient is −0.059 and the earnings yield coefficient is 0.058. Statistically, we fail to reject the null hypothesis that these two coefficients are equal in absolute value with t-statistics of −0.06 (Newey-West adjusted) and −0.04 (stationary bootstrap).

Table 2
Dividend growth and return predictability from dividend yields and earnings yields: post-war evidence. The sample period is from 1950:1 to 2013:4. Otherwise see notes to Table 1.

Dividend growth			Returns		
dp	ep	$\bar{R}^2$	dp	ep	$\bar{R}^2$
Panel A: $k=1$					
0.000			0.025		
(0.04)		−0.4%	(1.99)		1.4%
[0.04]			[2.06]		
	0.016			0.015	
	(2.59)	6.7%		(0.96)	0.3%
	[2.52]			[0.96]	
−0.029	0.037		0.031	−0.007	
(−3.53)	(7.94)	15.4%	(1.57)	(−0.35)	1.1%
[−3.05]	[5.36]		[1.39]	[−0.28]	
Panel B: $k=4$					
−0.007			0.106		
(−0.31)		−0.2%	(2.22)		6.6%
[−0.32]			[2.33]		
	0.050			0.063	
	(2.02)	9.1%		(1.33)	2.4%
	[1.98]			[1.33]	
−0.110	0.128		0.128	−0.028	
(−3.47)	(8.42)	26.6%	(1.92)	(−0.61)	6.5%
[−2.93]	[4.94]		[1.70]	[−0.43]	

4.1. Pre-war vs. post-war results

Chen (2009) shows that returns are predictable and dividend growth unpredictable from dividend yields in post-war years, but in pre-war years the opposite pattern applies, see also Golez and Koudijs (2014). Hence, one may wonder whether the strong dividend growth predictability we have found for the 1927:1-2013:4 period is driven mainly by the pre-war years. Over the period 1950:1-2013:4, Table 2 shows that the $dy - ey$ specification

[7] We also report results for non-overlapping annual regressions in Section 4.1.1.

Table 3
Dividend growth and return predictability from dividend yields and earnings yields: pre-war evidence. The sample period is from 1927:1 to 1949:4. Otherwise see notes to Table 1.

Dividend growth			Returns		
dp	*ep*	$\bar{R}^2$	*dp*	*ep*	$\bar{R}^2$
Panel A: $k = 1$					
−0.113			0.051		
(−4.16)		25.1%	(0.72)		−0.2%
[−3.88]			[0.64]		
	0.013			0.101	
	(0.56)	−0.7%		(2.05)	3.3%
	[0.52]			[1.98]	
−0.228	0.157		−0.044	0.129	
(−6.79)	(5.09)	56.1%	(−0.39)	(1.44)	2.6%
[−4.56]	[3.71]		[−0.34]	[1.37]	
Panel B: $k = 4$					
−0.529			0.220		
(−5.08)		54.6%	(1.58)		3.4%
[−4.78]			[1.46]		
	−0.100			0.338	
	(−0.96)	0.8%		(2.28)	9.4%
	[−0.84]			[2.36]	
−0.856	0.484		−0.014	0.347	
(−13.39)	(16.81)	79.2%	(−0.05)	(1.21)	8.3%
[−8.98]	[9.74]		[−0.05]	[1.19]	

strongly predicts dividend growth, both at the one-quarter ahead horizon and the one-year ahead horizon. The earnings yield on its own also significantly predicts dividend growth, but the degree of predictive power for future dividend growth increases substantially when the dividend yield and earnings yield are used in a joint specification.

Table 3 shows results for the pre-war period from 1927:1 to 1949:4. Consistent with Chen (2009), the evidence of dividend growth predictability is stronger in this period. The dividend yield on its own significantly predicts dividend growth and generates an $\bar{R}^2$ of 25.1% at the one-quarter ahead horizon. However, the $\bar{R}^2$ increases from 25.1% to 56.1% by controlling for the earnings yield. That is, the missing variable problem is also present in the pre-war period.

4.1.1. Annual frequency

To further judge the pre-war vs. post-war effect, we analyze annual data and run one-year ahead regressions. The use of annual data allows us to extend the sample period back to 1871. The annual results in Table 4 again confirms the tale-of-two-periods finding of Chen (2009). The dividend yield on its own significantly predicts dividend growth over the 1871–1949 period, but not afterwards. Conversely, it significantly predicts returns over the 1950–2013 period, but not before. The dividend growth regressions show that the dividend yield coefficient is −0.435 and highly significant in the 1871–1949 period, while in the 1950–2013 period, it drops to −0.014 and turns insignificant. In a similar vein, the $\bar{R}^2$ drops from 47.1% to −0.6%. However, once we control for the earnings yield, the dividend yield becomes a highly significant predictor of dividend growth in the post-war period and so does the earnings yield. The joint specification gives an $\bar{R}^2$ of 60.1% and 45.7% in the 1871–1949 and 1950–2013 periods, respectively. Hence, the large and puzzling reversal in dividend growth predictability does not show up when using the $dy - ey$ model.

4.2. Omitted variable bias

From the Lintner (1956) model, we have in Section 2 derived that the expected slope estimate from regressing Δd_{t+1} on dp_t is given by $E(\widehat{\beta}) = -c(1-\gamma)$, where $-c$ is the slope coefficient from

Table 4
Long annual sample. The table reports results from forecasting regressions, $y_{t+1} = \alpha + \beta^{dp} dp_t + \beta^{ep} ep_t + \varepsilon_{t+1}$, where y_{t+1} is the one-year ahead dividend growth rate or excess return. We use the dividend yield and the earnings yield as regressors, either separately or jointly. Below OLS estimates of β are t-statistics in parentheses based on Newey-West standard errors with one lag and in brackets are t-statistics based on standard errors from a stationary bootstrap. We resample the data in blocks of random size and determine the average block-size using the Politis and White (2004) automatic selection procedure. $\bar{R}^2$ is the adjusted R-squared. We report results from two sample periods: 1871–1949 (Panel A) and 1950–2013 (Panel B).

Dividend growth			Returns		
dp	*ep*	$\bar{R}^2$	*dp*	*ep*	$\bar{R}^2$
Panel A: 1871–1949					
−0.435			0.057		
(−6.62)		47.1%	(0.70)		−0.7%
[−6.26]			[0.69]		
	−0.047			0.036	
	(−0.76)	−0.5%		(0.59)	−1.0%
	[−0.72]			[0.53]	
−0.593	0.229		0.048	0.014	
(−7.46)	(3.40)	60.1%	(0.41)	(0.15)	−2.1%
[−7.09]	[3.29]		[0.39]	[0.14]	
Panel B: 1950–2013					
−0.014			0.100		
(−0.62)		−0.6%	(1.96)		4.9%
[−0.52]			[2.02]		
	0.047			0.070	
	(1.49)	10.3%		(1.36)	1.5%
	[2.25]			[1.22]	
−0.133	0.151		0.119	−0.024	
(−3.91)	(4.91)	45.7%	(1.56)	(−0.37)	3.5%
[−2.73]	[4.24]		[1.55]	[−0.30]	

regressing Δd_{t+1} on $(dp_t - ep_t)$ or equivalently on $(d_t - e_t)$, and γ is the slope coefficient from a regression of ep_t on dp_t (each regression includes a constant). In the following, we take a closer look at this relation.

Table 5 reports OLS estimates of $\beta, -c$, and γ. Consistent with the results in Table 1, $-\hat{c}$ is much higher in absolute value compared to $\hat{\beta}$. For the full period with a quarterly frequency, the estimates are −0.059 and −0.018, respectively. Since dp_t and ep_t are highly correlated, we expect the dp_t coefficient to be close to zero when omitting ep_t. In fact, the $\widehat{\gamma}$-value indicates that the dp_t coefficient will be reduced by 70% of its true value on average in repeated sampling.

If we use the OLS estimates of γ and c, we can compute the expected dp_t coefficient implied by the Lintner model, $E(\widehat{\beta})$. As Table 5 shows, $E(\widehat{\beta})$ is roughly equal to the univariate sample estimate $\widehat{\beta}$. This is because the multivariate regression slopes on dp_t and ep_t are roughly equal in absolute value as predicted by the model, see Table 1 and Section 2.[8] These results illustrate that an omitted variable bias helps explain the non-predictability of dividend growth from dividend yields.

Comparing the full-period results with the post-war results, we see that the higher level of dividend growth predictability using dp_t for the full period is both due to a weaker relationship between dp_t and ep_t (lower γ) and less dividend smoothing (higher c). Chen et al. (2012) also find that dividends are more smooth in the post-war period. Our analysis extends their work by revealing a direct link between the level of dividend smoothing and the bias in regressions of dividend growth on dp_t.

[8] The univariate and multivariate estimates are linked algebraically as $\widehat{\beta} = \widehat{\beta}^{dp} + \widehat{\beta}^{ep}\widehat{\gamma}$, where $\widehat{\beta}$ is the slope estimate from a univariate regression of Δd_{t+1} on dp_t, and $\left(\widehat{\beta}^{dp}, \widehat{\beta}^{ep}\right)'$ are the multivariate slope estimates from a regression of Δd_{t+1} on $(dp_t, ep_t)'$.

Table 5
Explaining the non-predictability of dividend growth as an omitted variable bias. $-\widehat{c}$ is the OLS estimate from regressing Δd_{t+k} on $d_t - e_t$, $\widehat{\gamma}$ is the OLS estimate from regressing ep_t on dp_t, and $\widehat{\beta}$ is the OLS estimate from regressing Δd_{t+k} on dp_t. $E(\widehat{\beta})$ is calculated as $-\widehat{c}(1-\widehat{\gamma})$. The quantiles are from simulations of the system in Eqs. (8) and (9). For the coefficients, we use equation-by-equation Maximum Likelihood estimates. For each simulation, we draw a value from the empirical likelihood of c where the likelihood is calculated with the other parameters maxed out. We use 50,000 simulations and draw $(\widehat{\varepsilon}_{t+1}, \widehat{\eta}^{dp}_{t+1}, \widehat{\eta}^{ep}_{t+1})'$ jointly to preserve cross-correlation. In each simulated sample we regress Δd_{t+k} on dp_t and report quantiles of the resulting distribution of slope coefficients. We also report probabilities of rejecting the null of no predictability (power) for the univariate regression of Δd_{t+k} on dp_t (t-statistic is $t^{\widehat{\beta}}$) and for the multivariate regression of Δd_{t+k} on dp_t and ep_t (t-statistic is $t^{\widehat{\beta}^{dp}}$).

	Data			Model				
	$-\widehat{c}$	$\widehat{\gamma}$	$\widehat{\beta}$	$E(\widehat{\beta})$	$\widehat{\beta}_{2.5\%}$	$\widehat{\beta}_{97.5\%}$	$P(\lvert t^{\widehat{\beta}}\rvert>1.96)$	$P(\lvert t^{\widehat{\beta}^{dp}}\rvert>1.96)$
$k=1$								
1927:12013:4	−0.059	0.698	−0.018	−0.018	−0.037	0.006	0.491	0.978
1950:1–2013:4	−0.034	0.794	0.000	−0.007	−0.024	0.009	0.217	0.841
$k=4$								
1927:1–2013:4	−0.210	0.702	−0.081	−0.063	−0.164	0.005	0.652	0.990
1950:1–2013:4	−0.121	0.799	−0.007	−0.024	−0.109	0.025	0.349	0.901

We also report quantiles of the distribution of $\widehat{\beta}$ coefficients implied by the Lintner model. We compute them by simulating from the equation-by-equation Maximum Likelihood (ML) estimates of the following system[9]

$$\Delta d_{t+1} = \alpha - c(d_t - e_t) + \varepsilon_{t+1} + \rho_1\varepsilon_t + \rho_2\varepsilon_{t-1} + \rho_3\varepsilon_{t-2} + \rho_4\varepsilon_{t-3} \tag{8}$$

$$x_{t+1} = \mu + \Phi x_t + \eta_{t+1}. \tag{9}$$

We incorporate uncertainty about the true c by drawing a new value for each simulation from the empirical likelihood. We map out the likelihood by reestimating the model for a grid of c-values going from $-2\widehat{c}$ to 0 with increments of 0.0005, where $\widehat{c}$ is the ML estimate. Cochrane (2008) uses a similar procedure to model the uncertainty about the dividend-price ratio process when simulating returns and dividend growth.

Table 5 shows that the empirical $\widehat{\beta}$ value is included in the 95% confidence interval implied by the model. We also report the probability of finding significant dividend growth predictability ("power") from estimating the wrong regression using only dp_t and from our main regression using dp_t and ep_t jointly.[10] For the $k=1$ horizon, the probability of rejecting the null of no predictability increases from 0.49 to 0.98 using dp_t and ep_t together instead of only dp_t for the full period and from 0.22 to 0.84 for the post-war period. As Table 5 shows, substantial gains in power are also achieved for the $k=4$ horizon. The probabilities are calculated with a critical value for the Newey-West t-statistic of 1.96. Using a 1% critical value of 2.58 gives similar results. The significant increase in power highlights the importance of using dividends and earnings jointly when predicting dividend growth.

4.3. Statistical significance

We compute dividend growth as quarterly log changes in annual dividend payments, which is likely to induce a moving average structure in the error term in our regressions. We use the Newey-West estimator with a relatively high lag number and a stationary bootstrap to conduct inference. However, our time-series regressions may still suffer from finite sample biases, see, e.g., Hodrick (1992), Stambaugh (1999), Valkanov (2003), and Ang and Bekaert (2007). To verify robustness, we therefore now bootstrap under the null of no predictability as in, e.g., Nelson and Kim (1993) and Kothari and Shanken (1997). Specifically, we assume an MA(4) process for dividend growth and a bivariate VAR(1) for dp_t and ep_t. The system is

$$\Delta d_{t+1} = \alpha + \varepsilon_{t+1} + \rho_1\varepsilon_t + \rho_2\varepsilon_{t-1} + \rho_3\varepsilon_{t-2} + \rho_4\varepsilon_{t-3} \tag{10}$$

$$x_{t+1} = \mu + \Phi x_t + \eta_{t+1}, \tag{11}$$

where $x_t = (dp_t,\ ep_t)'$. We first estimate the system equation-by-equation using ML and save the estimated coefficients and residuals. We then create 50,000 bootstrap samples with $T = 348$ observations corresponding to the length of our full sample period. To generate the samples, we draw $\left(\widehat{\varepsilon}_{t+1},\ \widehat{\eta}'_{t+1}\right)'$ jointly from the ML residuals to preserve cross-correlation. We initiate the simulations at random draws of x_t, $\widehat{\varepsilon}_t$, $\widehat{\varepsilon}_{t-1}$, $\widehat{\varepsilon}_{t-2}$, and $\widehat{\varepsilon}_{t-3}$.

For each simulated sample, we reestimate the regression model in (7) with $k=1$ and $k=4$ and compute the Newey-West t-statistics with $k+3$ lags. Fig. 1 plots histograms of the simulated t-statistics. Here we focus on the t-statistics of $\widehat{\beta}^{dp}_k$ but results are similar for $\widehat{\beta}^{ep}_k$. Less than one percent of the simulated t-statistics are as extreme as the t-statistic we observe in the data. Hence, it is very unlikely that our results are driven by small-sample biases in the t-statistics.

4.4. Economic significance

To give a sense of the economic significance, we analyze standard deviations of fitted values $\sigma\left(\widehat{y}_{t+k}\right)$ from the forecasting regressions in (7). Cochrane (2008) finds standard deviations of about 5% for expected returns and about 0% for expected dividend growth when using dp_t as a sole predictor in annual regressions. As we show in Table 6, this roughly matches our results for the post-war period (see Panel B). However, when using dp_t and ep_t together, the standard deviation of expected dividend growth rises from close to 0% to 3.8% on an annual horizon. Comparing the standard deviation of fitted values and the mean value, $\sigma\left(\widehat{y}_{t+k}\right)/E\left(y_{t+k}\right)$, we see that expected dividend growth vary almost as much as its mean level.

If we consider the full sample from 1927:1 to 2013:4, the economic significance of dividend growth predictability is even stronger (see Panel A). The standard deviation of expected dividend growth using dp_t on its own is 3.6% at the annual horizon and it increases to 6.8% when including ep_t in the regression, which implies that expected dividend growth actually varies by more than its mean level. These results show that the economic effect of controlling for ep_t is sizeable when forecasting dividend growth.

[9] We include lagged MA terms to address the issue of overlapping observations. The results are robust towards the lag length specification.

[10] We do not include the MA terms in these regressions.

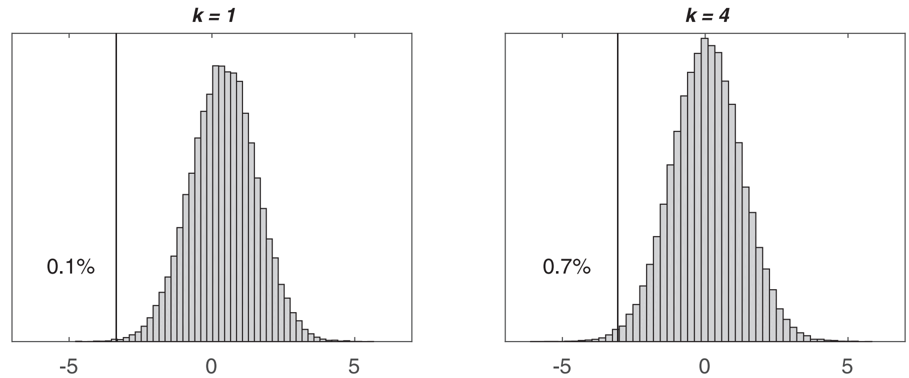

Fig. 1. Distribution of dp_t regression coefficients assuming no predictability. We simulate 50,000 bootstrap samples of length $T = 348$ using $\Delta d_{t+1} = 0.012 + \widehat{\varepsilon}_{t+1} + 0.489\widehat{\varepsilon}_t + 0.499\widehat{\varepsilon}_{t-1} + 0.433\widehat{\varepsilon}_{t-2} - 0.286\widehat{\varepsilon}_{t-3}$, $dp_{t+1} = -0.096 + 0.951dp_t + 0.026ep_t + \widehat{\eta}^{dp}_{t+1}$, and $ep_{t+1} = -0.162 + 0.016dp_t + 0.922ep_t + \widehat{\eta}^{ep}_{t+1}$. The coefficients are the full-sample (1927:1-2013:4) equation-by-equation Maximum Likelihood estimates. We draw $(\widehat{\varepsilon}_{t+1}, \widehat{\eta}^{dp}_{t+1}, \widehat{\eta}^{ep}_{t+1})'$ jointly to preserve cross-correlation. In each simulated sample, we regress Δd_{t+k} on $(dp_t, ep_t)'$ for $k = 1$ and $k = 4$. The histograms show the distribution of the simulated t-statistics on dp_t. The black lines are the original full-sample t-statistics and the percentages indicate the fraction of simulations more extreme than the original t-statistics.

Table 6
Economic significance. We run forecasting regressions, $y_{t+k} = \alpha_k + \beta^{dp}_k dp_t + \beta^{ep}_k ep_t + \varepsilon_{t+k}$, where y_{t+k} is the k-period ahead dividend growth rate or excess return. We use the dividend yield and the earnings yield as regressors, either separately or jointly. We report $\sigma(\widehat{y}_{t+k})$ and $\sigma(\widehat{y}_{t+k})/E(y_{t+k})$, where $\widehat{y}_{t+k}$ is the fitted values from the regressions. All values are in percent. We report results from two sample periods: 1927:1-2013:4 (Panel A) and 1950:1-2013:4 (Panel B).

		Dividend growth			Returns		
		dp	*ep*	Joint	*dp*	*ep*	Joint
Panel A: 1927:1–2013:4							
$\sigma(\widehat{y}_{t+k})$	$k=1$	0.81	0.57	1.88	0.94	1.18	1.19
$\sigma(\widehat{y}_{t+k})/E(y_{t+k})$	$k=1$	0.73	0.51	1.70	0.64	0.81	0.81
$\sigma(\widehat{y}_{t+k})$	$k=4$	3.63	1.25	6.79	4.16	4.38	4.60
$\sigma(\widehat{y}_{t+k})/E(y_{t+k})$	$k=4$	0.83	0.28	1.54	0.74	0.77	0.81
Panel B: 1950:1–2013:4							
$\sigma(\widehat{y}_{t+k})$	$k=1$	0.01	0.71	1.07	1.07	0.67	1.09
$\sigma(\widehat{y}_{t+k})/E(y_{t+k})$	$k=1$	0.01	0.53	0.79	0.68	0.42	0.69
$\sigma(\widehat{y}_{t+k})$	$k=4$	0.32	2.23	3.77	4.46	2.82	4.53
$\sigma(\widehat{y}_{t+k})/E(y_{t+k})$	$k=4$	0.06	0.42	0.71	0.74	0.47	0.75

4.5. Additional variables

Movements in expected returns that have positive correlation with movements in expected dividend growth will have offsetting effects on the dividend yield. In particular, other variables may predict both returns and dividend growth but in an offsetting way so that the dividend yield does not change, which will make the dividend yield highly persistent and unable to uncover return and dividend growth predictability, see Menzly et al. (2004), Lettau and Ludvigson (2005), Binsbergen and Koijen (2010), and Golez (2014).

To examine the importance of including additional variables, Table 7 reports results from forecasting regressions with quarterly observations

$$y_{t+1} = \alpha + \beta^{dp} dp_t + \beta^{ep} ep_t + \beta^z z_t + \varepsilon_{t+1}, \tag{12}$$

where z_t is either the short rate, the term spread, the default spread, or *cay*.[11] These variables are among the most commonly used predictive variables in the literature. Comparing the results from Tables 1 and 7, we see that controlling for these variables adds little additional predictive power for future dividend growth. However, especially the short rate and *cay* help in predicting future returns, which is in line with the findings of Ang and Bekaert (2007) and Lettau and Ludvigson (2001). These results highlight that it is not the same variables that drive expected dividend growth and expected returns.

4.6. Recessions vs. expansions

In a recent study, using the dividend yield among other predictive variables, Henkel et al. (2011) show that return predictability exists in recessions only. Similarly, Rapach et al. (2010); Dangl and Halling (2012), and Rapach and Zhou (2013) also find that return predictability is primarily present in recessions. Zhu (2015) introduces regime switching into to the present value model of Binsbergen and Koijen (2010) and finds evidence of time-varying return and dividend growth predictability.

To investigate whether the predictive power of the $dy - ey$ specification varies across the business cycle, we run simple OLS switching regressions of the following form

$$\begin{aligned} y_{t+1} = &\left(\alpha_R + \beta^{dp}_R dp_t + \beta^{ep}_R ep_t\right) \times I_t \\ &+\left(\alpha_E + \beta^{dp}_E dp_t + \beta^{ep}_E ep_t\right) \times (1 - I_t) + \varepsilon_{t+1}. \end{aligned} \tag{13}$$

Because the NBER methodology provides the most common definition of recessions and expansions, we use the NBER recession indicator, I_t, to classify business cycles. Investors do not know NBER recession dates in real time but this is not an issue. Our goal is not to set up an implementable investment strategy.

Table 8 shows the results of estimating (13) with quarterly observations. We classify recessions using the quarterly turning points from the NBER. Consistent with Henkel et al. (2011), the dividend yield significantly predicts returns in recessions, but not in expansions. However, this result only holds for the post-war period, as the dividend yield is insignificant in both expansions and recessions over the full period from 1927:1 to 2013:4.

Interestingly, conditioning on the state of the business cycle does not matter much with respect to the predictive power of the $dy - ey$ specification for future dividend growth. When the two variables are used together, the dividend yield and earnings yield coefficients are generally highly significant in both states of the economy.[12] Hence, while return predictability seems concentrated to a few recession periods, dividend growth predictability from the $dy - ey$ model holds in both recessions and expansions.

[11] The short rate is proxied by the three-month T-bill rate, the term spread is defined as the long-term yield on government bonds minus the short rate, and the default spread is the difference in yields between BAA and AAA rated corporate bonds. We use same data and definitions as in Goyal and Welch (2008). The consumption-wealth ratio (*cay*) of Lettau and Ludvigson (2001); 2005) is obtained from the website of Martin Lettau.

[12] In unreported results, we find in general that the differences in coefficients are not statistically different across recessions and expansions.

Table 7
Controlling for additional variables. The table reports results from forecasting regressions, $y_{t+1} = \alpha + \beta^{dp} dp_t + \beta^{ep} ep_t + \beta^{z} z_t + \varepsilon_{t+1}$, where y_{t+1} is the one-quarter ahead dividend growth rate or excess return. z_t is either the short rate, the term spread, the default spread, or *cay*. We report results from two sample periods: 1927:1-2013:4 (Panel A) and 1950:1-2013:4 (Panel B). *cay* is available from 1952:1. Below OLS estimates of slope coefficients are *t*-statistics in parentheses based on Newey-West standard errors with 4 lags and in brackets are *t*-statistics based on standard errors from a stationary bootstrap. We resample the data in blocks of random size and determine the average block-size using the Politis and White (2004) automatic selection procedure. $\overline{R}^2$ is the adjusted *R*-squared.

	Dividend growth				Returns			
	dp	*ep*	*z*	$\overline{R}^2$	*dp*	*ep*	*z*	$\overline{R}^2$
Panel A: 1927:1–2013:4								
Short rate	−0.063	0.064	−0.146		−0.006	0.038	−0.349	
	(−3.54)	(4.12)	(−1.73)	20.3%	(−0.19)	(1.04)	(−1.95)	1.3%
	[−2.12]	[2.62]	[−1.03]		[−0.19]	[1.15]	[−1.99]	
Term spread	−0.059	0.059	0.052		0.000	0.035	0.801	
	(−3.42)	(4.17)	(0.29)	19.3%	(0.00)	(1.01)	(1.88)	1.3%
	[−2.39]	[2.97]	[0.21]		[0.00]	[1.12]	[1.80]	
Default spread	−0.041	0.046	−1.462		−0.004	0.031	0.659	
	(−3.76)	(4.79)	(−2.52)	23.9%	(−0.13)	(0.88)	(0.38)	0.5%
	[−2.13]	[3.07]	[−2.51]		[−0.16]	[1.07]	[0.33]	
Panel B: 1950:1–2013:4								
Short rate	−0.029	0.041	−0.126		0.030	0.010	−0.546	
	(−3.71)	(7.37)	(−1.65)	16.7%	(1.45)	(0.45)	(−3.28)	4.2%
	[−3.06]	[5.44]	[−1.61]		[1.27]	[0.36]	[−3.13]	
Term spread	−0.029	0.039	0.201		0.030	0.003	0.863	
	(−3.79)	(7.93)	(1.51)	16.1%	(1.52)	(0.16)	(2.36)	2.8%
	[−3.26]	[5.49]	[1.29]		[1.33]	[0.13]	[2.24]	
Default spread	−0.028	0.036	−0.246		0.031	−0.007	−0.011	
	(−3.26)	(7.64)	(−0.72)	15.2%	(1.51)	(−0.35)	(−0.01)	0.7%
	[−2.79]	[4.74]	[−0.60]		[1.38]	[−0.28]	[−0.01]	
cay	−0.030	0.035	0.065		0.020	−0.006	0.653	
	(−3.63)	(8.47)	(0.99)	16.5%	(0.95)	(−0.29)	(2.38)	2.4%
	[−3.01]	[5.88]	[0.88]		[0.86]	[−0.23]	[2.38]	

Table 8
Recessions vs. expansions. The table reports results from running forecasting regressions of the following form: $y_{t+1} = (\alpha_R + \beta_R^{dp} dp_t + \beta_R^{ep} ep_t) \times I_t + (\alpha_E + \beta_E^{dp} dp_t + \beta_E^{ep} ep_t) \times (1 - I_t) + \varepsilon_t$, where I_t is the NBER recession indicator and y_{t+1} is the one-quarter ahead dividend growth rate or excess return. We report results from two sample periods: 1927:1-2013:4 (Panel A) and 1950:1-2013:4 (Panel B). Below OLS estimates of slope coefficients are *t*-statistics in parentheses based on Newey-West standard errors with 4 lags and in brackets are *t*-statistics based on standard errors from a stationary bootstrap. We resample the data in blocks of random size and determine the average block-size using the Politis and White (2004) automatic selection procedure. $\overline{R}^2$ is the adjusted *R*-squared.

Dividend growth					Returns				
$dp \times I$	$dp \times (1-I)$	$ep \times I$	$ep \times (1-I)$	$\overline{R}^2$	$dp \times I$	$dp \times (1-I)$	$ep \times I$	$ep \times (1-I)$	$\overline{R}^2$
Panel A: 1927:1–2013:4									
−0.044	−0.003				0.058	0.014			
(−1.63)	(−0.34)			11.7%	(1.09)	(1.17)			0.4%
[−1.19]	[−0.31]				[0.79]	[1.25]			
		0.018	0.013				0.037	0.023	
		(2.28)	(1.80)	10.0%			(1.03)	(1.65)	0.4%
		[1.35]	[2.04]				[0.89]	[1.72]	
−0.099	−0.036	0.064	0.046		0.044	−0.008	0.016	0.030	
(−3.35)	(−2.48)	(3.02)	(3.33)	25.3%	(0.50)	(−0.43)	(0.28)	(1.37)	0.2%
[−2.36]	[−2.08]	[1.97]	[3.28]		[0.40]	[−0.40]	[0.24]	[1.28]	
Panel B: 1950:1–2013:4									
0.009	0.002				0.101	0.011			
(1.21)	(0.29)			4.6%	(2.67)	(0.88)			3.4%
[0.82]	[0.31]				[2.04]	[0.90]			
		0.018	0.015				0.013	0.016	
		(3.85)	(2.20)	11.5%			(0.48)	(1.11)	−0.1%
		[2.82]	[2.34]				[0.34]	[1.12]	
−0.026	−0.028	0.029	0.040		0.169	−0.003	−0.056	0.020	
(−4.68)	(−2.76)	(8.25)	(5.22)	17.3%	(3.67)	(−0.17)	(−2.56)	(0.84)	4.7%
[−1.40]	[−2.57]	[2.01]	[4.35]		[1.65]	[−0.16]	[−0.67]	[0.77]	

4.7. Other specifications of dividends, earnings, and prices

When predicting dividend growth it is ultimately an empirical question what the optimal combination of dividends, earnings, and prices is. The dividend yield $(d_t - p_t)$, the payout ratio $(d_t - e_t)$, and our main $dy - ey$ specification $(d_t - p_t,\ e_t - p_t)$ all correspond to different sets of weights on d_t, p_t, and e_t. Garrett and Priestley (2012) use cointegration analysis to estimate these weights. Ignoring the constant, their dpe_t cointegration residual is computed as $(d_t - \delta_1 p_t - \delta_2 e_t)$, where the coefficients are es-

Table 9
Other specifications of dividends, earnings, and prices. The table reports results from k-quarter ahead forecasting regressions, $\Delta d_{t+k} = \alpha + \beta' x_t + \varepsilon_{t+k}$. We report results for $x_t = (dp_t, ep_t)'$, $x_t = d_t - e_t$, $x_t = dpe_t = (d_t - \delta_0 - \delta_1 p_t - \delta_2 e_t)$, and $x_t = (d_t, p_t, e_t)'$. The cointegration coefficients $(\delta_0, \delta_1, \delta_2)'$ are estimated from a regression of d_t on a constant, p_t, and e_t. In parentheses are t-statistics using Newey-West standard errors with $k+3$ lags and in brackets are t-statistics using standard errors from a stationary bootstrap. We resample the data in blocks of random size and determine the average block-size using the Politis and White (2004) automatic selection procedure. $\overline{R}^2$ is the adjusted R-squared.

	$k=1$				$k=4$			
	β_1	β_2	β_3	$\overline{R}^2$	β_1	β_2	β_3	$\overline{R}^2$
Panel A: 1927:1–2013:4								
dp, ep	−0.059	0.058			−0.220	0.197		
	(−3.36)	(4.05)		19.5%	(−3.07)	(3.85)		26.9%
	[−2.42]	[3.11]			[−2.10]	[3.16]		
$d-e$	−0.059				−0.210			
	(−3.69)			19.7%	(−3.32)			26.7%
	[−2.78]				[−2.65]			
dpe	−0.089				−0.364			
	(−3.69)			17.1%	(−3.30)			30.2%
	[−2.52]				[−2.14]			
d, p, e	−0.089	0.017	0.063	22.5%	−0.363	0.099	0.223	34.9%
	(−3.25)	(1.59)	(3.35)		(−3.14)	(1.91)	(3.10)	
Panel B: 1950:1–2013:4								
dp, ep	−0.029	0.037			−0.110	0.128		
	(−3.53)	(7.94)		15.4%	(−3.47)	(8.42)		26.6%
	[−3.05]	[5.36]			[−2.93]	[4.94]		
$d-e$	−0.034				−0.121			
	(−5.95)			14.3%	(−5.63)			25.8%
	[−4.27]				[−3.85]			
dpe	−0.032				−0.126			
	(−1.99)			4.7%	(−1.85)			10.7%
	[−1.74]				[−1.68]			
d, p, e	−0.032	−0.006	0.037	15.0%	−0.125	−0.008	0.128	26.4%
	(−3.14)	(−0.89)	(7.66)		(−2.90)	(−0.30)	(8.44)	

timated from a regression of d_t on a constant, p_t, and e_t.[13] An alternative approach is to impose no structure a priori and let the data decide by including all three variables on their own in a multivariate regression. Lettau and Ludvigson (2005) use this approach in the context of their cointegration analysis of consumption, dividends, and labor income.

Table 9 shows results for the various alternatives described above. Panels A and B report results for the 1927:1-2013:4 and 1950:1-2013:4 periods, respectively. For the full period, all the models we consider provide similar explanatory power with $\overline{R}^2$s around 20% at the quarterly horizon reflecting strong dividend growth predictability. For the post-war period, the explanatory power remains roughly intact for the $dy-ey$ model. However, the $\overline{R}^2$ and t-statistic for the dpe cointegration residual drops substantially. In Table 10, we report results for the long annual sample and the conclusions are similar. Interestingly, for the free specification, in which we do not impose any structure on the weights of dividends, prices, and earnings, we get slopes and $\overline{R}^2$s that are not too different from the main $dy-ey$ specification.[14] This is true for both pre- and post-war data even though the loadings change significantly across the two subsamples. The $dy-ey$ model thus provides flexibility when modeling dividend growth and it does not rely on the estimation of cointegration vectors.

Looking at the pre-war period, Table 10 shows how the less restricted $dy-ey$ model and the free multivariate specification give slope estimates on dividends, earnings, and prices that are similar to those of the dpe cointegration residual, which has a nonzero weight on prices. The $\overline{R}^2$s are also quite close. The post-war period shows a different pattern, as the $dy-ey$ model and the free specification give results that mimic the payout ratio, $d-e$. For the $dy-ey$ model, the post-war coefficients on the dividend yield and earnings yield are very close in absolute value canceling out the effect of prices. For the free specification, the post-war price coefficient is very close to zero and is insignificant. These results suggest that dividend payments have become less responsive to price levels over time, which goes hand in hand with the finding of Chen et al. (2012) that dividends have become more smooth in the post-war period. That is, as prices are more volatile than earnings, we should expect to see more smooth dividends payments if they respond more to earnings than to prices, other things being equal.

5. International evidence

To examine whether the pattern of strong dividend growth predictability from the $dy-ey$ model holds outside the U.S., we now analyze data from 14 developed countries. We start out running panel regressions using all 14 countries to make a statement about the average predictive relationship across countries. We forecast one-quarter ahead and the sample runs from 1973:1 to 2013:4. Similar to the international predictability studies of Ang and Bekaert (2007) and Hjalmarsson (2010), we constrain the slope coefficients to be the same across countries but allow for heterogeneous intercepts. Following the procedure in Thompson (2011), we compute standard errors that are robust to heteroskedasticity as well as correlation in both the country and time dimension. In addition, we compute standard errors from a stationary bootstrap where the random time-series blocks are drawn commonly for all countries to preserve cross-correlation.

Table 11 shows that in international equity markets the dividend yield significantly predicts dividend growth as well as re-

[13] In our quarterly sample 1927:1-2013:4, we estimate the cointegration relation to be $(d_t - 0.40p_t - 0.43e_t)$ and we get similar values for the post-war period. Garrett and Priestley (2012) use annual real data, whereas we use quarterly nominal data. In unreported results, we find that the $dp-ep$ model also strongly predicts real dividend growth.

[14] For the unrestricted d, p, e regression, the right hand side variables are not stationary. However, conditional on the three variables being cointegrated, we can rely on standard asymptotics, see Sims et al. (1990).

Table 10
Other specifications of dividends, earnings, and prices: evidence from long annual sample. We use annual observations. Otherwise see notes for Table 9.

	β_1	β_2	β_3	$\bar{R}^2$	β_1	β_2	β_3	$\bar{R}^2$
	Panel A: 1871–1949				**Panel B: 1950–2013**			
dp, ep	−0.593	0.229			−0.133	0.151		
	(−7.46)	(3.40)		60.1%	(−3.91)	(4.91)		45.7%
	[−7.09]	[3.29]			[−2.73]	[4.24]		
$d-e$	−0.331				−0.142			
	(−3.92)			29.6%	(−4.25)			44.9%
	[−3.85]				[−3.82]			
dpe	−0.578				−0.146			
	(−6.51)			52.1%	(−1.82)			19.6%
	[−6.06]				[−1.74]			
d, p, e	−0.577	0.383	0.219	60.1%	−0.146	−0.012	0.153	45.2%
	(−6.84)	(6.88)	(3.19)		(−2.96)	(−0.45)	(5.11)	

Table 11
International evidence. The table reports results from international one-quarter ahead forecasting regressions of dividend growth and returns using the dividend yield and earnings yield as predictive variables. We use data from 14 developed countries over the period 1973:1 to 2013:4. We forecast equal-weighted (EW) and value-weighted (VW) global portfolios. For each regression, we report slope estimates, Newey-West t-statistics with 4 lags in parentheses, and bootstrap t-statistics in brackets. We also run cross-section fixed-effect panel regressions from which we report Thompson (2011) two-way clustered robust t-statistics with 4 lags in parentheses and bootstrap t-statistics in brackets. We use a stationary bootstrap where the random time-series blocks are drawn commonly for all countries to preserve cross-correlation. For the panel regressions, the $\bar{R}^2$ column reports the within R^2.

	Dividend growth						
	dp	$\bar{R}^2$	ep	$\bar{R}^2$	dp	ep	$\bar{R}^2$
Equal-weighted	−0.024		−0.008		−0.131	0.129	
	(−1.94)	4.7%	(−0.75)	−0.1%	(−3.50)	(3.39)	19.9%
	[−1.55]		[−0.73]		[−3.23]	[3.29]	
Value-weighted	−0.004		0.002		−0.073	0.078	
	(−0.63)	0.1%	(0.30)	−0.5%	(−3.39)	(3.83)	15.2%
	[−0.56]		[0.30]		[−2.56]	[3.07]	
Panel	−0.038		−0.011		−0.076	0.049	
	(−3.49)	3.7%	(−2.39)	0.3%	(−4.13)	(3.20)	6.1%
	[−3.59]		[−1.99]		[−3.95]	[3.03]	
	Returns						
	dp	$\bar{R}^2$	ep	$\bar{R}^2$	dp	ep	$\bar{R}^2$
Equal-weighted	0.040		0.044		0.023	0.021	
	(1.68)	1.0%	(1.55)	1.0%	(0.28)	(0.22)	0.5%
	[1.51]		[1.46]		[0.31]	[0.25]	
Value-weighted	0.017		0.014		0.056	−0.044	
	(0.96)	0.1%	(0.76)	−0.2%	(1.04)	(−0.77)	−0.2%
	[0.89]		[0.72]		[0.95]	[−0.73]	
Panel	0.043		0.039		0.034	0.012	
	(2.86)	1.8%	(2.50)	1.5%	(2.10)	(0.66)	1.9%
	[2.55]		[2.38]		[1.87]	[0.74]	

turns, although the explanatory power obtained from the return regression is relatively weak.[15] The panel regressions also reveal that the $dy - ey$ model gives much stronger evidence of dividend growth predictability than do univariate models with the dividend yield or the earnings yield as the only predictor. As with the U.S. evidence, the magnitudes of the dividend and earnings yield coefficients increase substantially when the two variables are used jointly: the dp_t coefficient changes from −0.038 to −0.076 and the ep_t coefficient changes from −0.011 to 0.049. These results again confirm our omitted variable interpretation.

Ang and Bekaert (2007) analyze countries with large market capitalizations (U.S., U.K., Germany, and France) and find that the dividend yield does not contain much predictive power for dividend growth. Similarly, the international studies of Engsted and Pedersen (2010) and Rangvid et al. (2014) also find evidence of weak dividend growth predictability from dividend yields in countries with large market capitalizations but show that dividend yields significantly predict dividend growth in small markets. These findings also relate to the work of Maio and Santa-Clara (2015), who examine the U.S. cross section of equities and show that there is more dividend growth predictability for small firms than for large firms.

To analyze the impact of market capitalization, Table 11 shows results for two global portfolios: an equal-weighted portfolio that gives a weight of 1/14 to each country and a value-weighted global portfolio based on market capitalizations. Unlike the panel regressions, we do not find any evidence of significant return predictability for these global portfolios, which may be attributed to lack of power due to the small sample size.[16] In addition, the dividend yield and earnings yield on their own do not seem to contain much predictive power for dividend growth. In agreement with the results of Rangvid et al. (2014), we find that the dividend yield by it-

[15] Other variables may help predict international equity returns such as the short interest rate, see Ang and Bekaert (2007) and Hjalmarsson (2010).

[16] Ang and Bekaert (2007) show how pooling cross-country data increases the power in international return regressions.

Table 12
International evidence: subsample analysis. The table reports results from international one-quarter ahead forecasting regressions of dividend growth using the dividend yield and earnings yield as predictive variables. We use data from 14 developed countries and show results for two subperiods: 1973:1 to 1992:4 and 1993:1 to 2013:4. We estimate cross-section fixed-effect panel regressions from which we report Thompson (2011) two-way clustered robust t-statistics with 4 lags in parentheses and bootstrap t-statistics in brackets. We use a stationary bootstrap where the random time-series blocks are drawn commonly for all countries to preserve cross-correlation.

	dp	R^2_{within}	ep	R^2_{within}	dp	ep	R^2_{within}
1973:1-1992:4	−0.038		−0.022		−0.055	0.021	
	(−3.41)	4.3%	(−2.21)	1.4%	(−2.62)	(1.04)	4.8%
	[−5.08]		[−3.11]		[−3.68]	[1.44]	
1993:1-2013:4	−0.077		−0.015		−0.143	0.093	
	(−3.35)	8.0%	(−1.50)	0.3%	(−5.36)	(6.35)	13.2%
	[−4.44]		[−0.93]		[−5.86]	[4.72]	

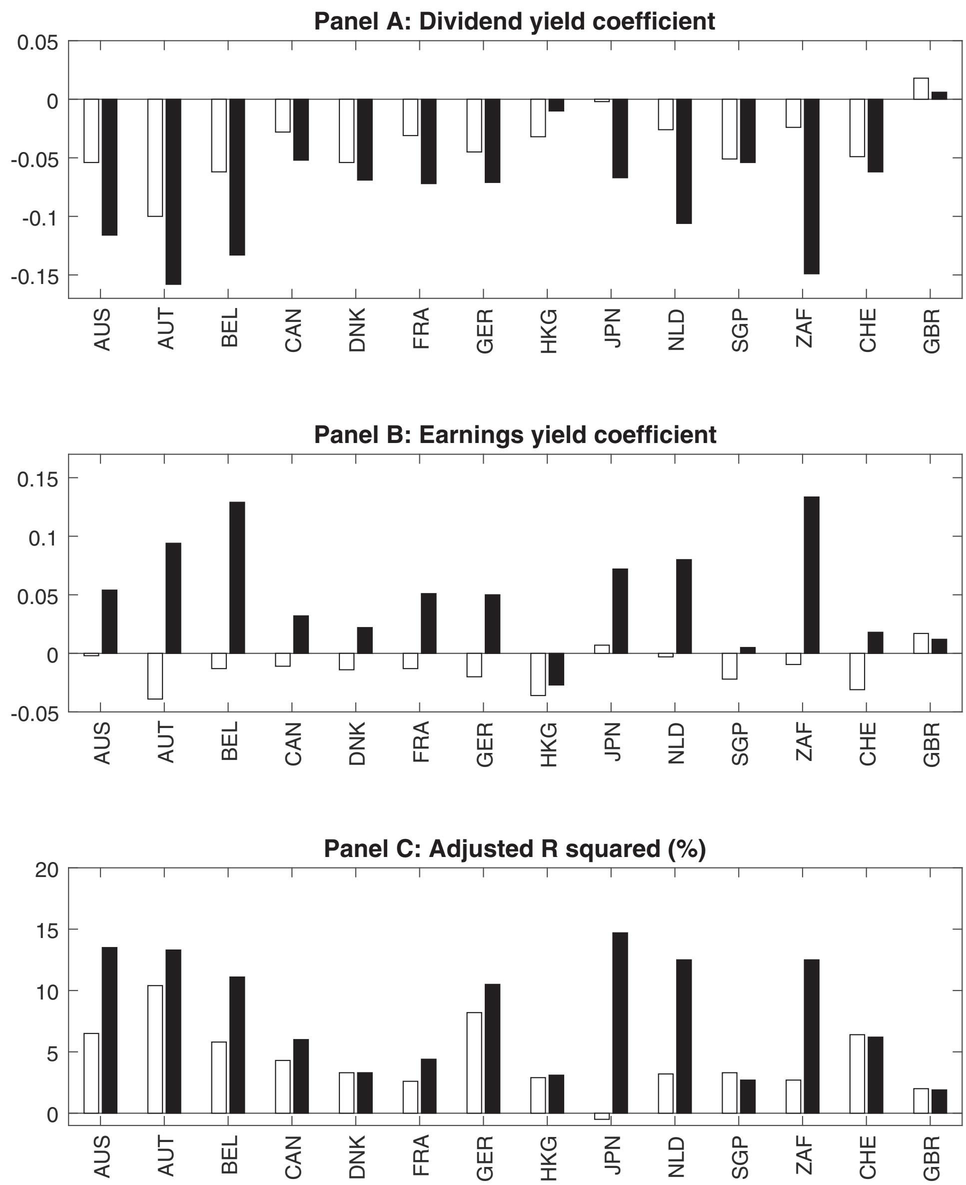

Fig. 2. Dividend growth regressions for individual countries. The figure shows results from forecasting regressions, $y_{i,t+1} = \alpha_i + \beta_i^{dp} dp_{i,t} + \beta_i^{ep} ep_{i,t} + \varepsilon_{i,t+1}$, where $y_{i,t+1}$ is the one-period ahead dividend growth rate for country i. We use the dividend yield and the earnings yield as regressors, either separately or jointly. The frequency is quarterly and the sample period is from 1973:1 to 2013:4. Panel A shows the estimates of β_i^{dp} obtained from univariate regressions using only $dp_{i,t}$ as regressor (white bars) and from bivariate regressions using $dp_{i,t}$ and $ep_{i,t}$ as regressors (black bars). Similarly, Panel B shows estimates of β_i^{ep} obtained from univariate regressions using only $ep_{i,t}$ as regressor (white bars) and from bivariate regressions using $dp_{i,t}$ and $ep_{i,t}$ as regressors (black bars). Finally, Panel C shows adjusted R^2s from the univariate dividend yield regressions (white bars) and bivariate regressions (black bars). For illustrative purposes, we have divided the slope coefficients for South Africa (ZAF) by two.

self generates a larger $\bar{R}^2$ for the equal-weighted portfolio.[17] However, once we use the dividend and earnings yields jointly in the predictive regressions, there is strong evidence of dividend growth predictability for both portfolios.

There is empirical evidence that financial market integration has increased significantly in recent decades, see, e.g., Pukthuanthong and Roll (2009) and Bekaert et al. (2011). Specifically, Bekaert et al. (2011) find that a large set of developed markets has been effectively integrated since 1993. To account for the changes in the level of market integration, Table 12 shows results from two sub-periods: 1973:1-1992:4 and 1993:1-2013:4. In the period were the markets in our sample were less financially integrated (before 1993), we find less support for the $dy - ey$ specification. In the 1973:1-1992:4 subperiod, the dy_t coefficient changes from −0.038 in the univariate regression to −0.055 in the joint specification and the ey_t coefficient changes sign from negative to positive, which lends partial support to the $dy - ey$ specification. However, ey_t is not statistically significant. In the period with high market integration (after 1993), we obtain stronger evidence in favor of using dy_t and ey_t jointly in forecasting regressions of dividend growth. As Table 12 shows, the magnitudes of the dy and ey coefficients are much larger in the joint specification relative to the univariate regressions. Moreover, the dividend yield by itself generates an explanatory power of 8.0%, but this number increases to 13.2% when controlling for the earnings yield.

To examine the cross-country patterns in dividend growth predictability, we next run forecasting regressions for each of the 14 individual countries (i.e., we do not constrain the slope coefficients to be the same across countries).[18] Fig. 2 summarizes the results. Consistent with the U.S. evidence, both the dividend yield coefficient and the earnings yield coefficient increase substantially in bivariate regressions using dp_t and ep_t as regressors compared to univariate regressions with either dp_t or ep_t as the only regressor. In addition, the explanatory power increases substantially by controlling for the earnings yield. This pattern is observed in the majority of countries. However, there is no obvious pattern in the degree of omitted variable bias across countries. The pattern seems complicated and does not appear to be related to market capitalization. The main conclusion we draw from the international results is that the omitted variable problem is not just restricted to the U.S. but exists in many countries.

6. Conclusion

We argue that the dividend yield by itself is a poor predictor of dividend growth due to an omitted variable bias. Within the context of the seminal model of Lintner (1956), we show how omitting the earnings yield conceals predictability by biasing the dividend yield coefficient towards zero. We provide empirical evidence in support of this implication of the Lintner model. In particular, consistent with Ang and Bekaert (2007), we find that dividend growth becomes strongly predictable when combining the dividend and earnings yields in a joint specification. In a simulation study, we further show how omitting the earnings yield leads to a substantial reduction in power. Our empirical results are robust in pre- and post-war U.S. data, in recessions and expansions, when controlling for additional predictors, and in international data.

[17] Our results are not directly comparable with those of Rangvid et al. (2014): they use an unbalanced panel of 50 countries, while we use a balanced panel of 14 countries.

[18] This means that we relax the assumption of a constant payout smoothing coefficient across countries.

References

Ang, A., Bekaert, G., 2007. Stock return predictability: is it there? Rev. Financ. Stud. 20, 651–707.
Ball, R., Watts, R., 1972. Some time series properties of accounting income. J. Finance 27, 663–681.
Bekaert, G., Harvey, C., Lundblad, C., Siegel, S., 2011. What segments equity markets? Rev. Financ. Stud. 24, 3841–3890.
Binsbergen, J.H.v., Koijen, R.S.J., 2010. Predictive regressions: a present-value approach. J. Finance 65, 1439–1471.
Campbell, J.Y., Shiller, R.J., 1988. The dividend-price ratio and expectations of future dividends and discount factors. Rev. Financ. Stud. 1, 195–228.
Chen, L., 2009. On the reversal of return and dividend growth predictability: a tale of two periods. J. Financ. Econ. 92, 128–151.
Chen, L., Da, Z., Priestley, R., 2012. Dividend smoothing and predictability. Manage. Sci. 58, 1834–1853.
Cochrane, J.H., 2008. The dog that did not bark: a defense of return predictability. Rev. Financ. Stud. 21, 1533–1575.
Cochrane, J.H., 2011. Discount rates. J. Finance 66, 1047–1108.
Dangl, T., Halling, M., 2012. Predictive regressions with time-varying coefficients. J. Financ. Econ. 106, 157–181.
Engsted, T., Pedersen, T.Q., 2010. The dividend-price ratio does predict dividend growth: international evidence. J. Empir. Finance 17, 585–605.
Garrett, I., Priestley, R., 2000. Dividend behavior and dividend signalling. J. Financ. Quant. Anal. 35, 173–189.
Garrett, I., Priestley, R., 2012. Dividend growth, cash flow, and discount rate news. J. Financ. Quant. Anal. 47, 1003–1028.
Golez, B., 2014. Expected returns and dividend growth rates implied by derivative markets. Rev. Financ. Stud. 27, 790–822.
Golez, B., Koudijs, P., 2014. Four Centuries of Return Predictability. NBERWorking Paper No. 20814
Goyal, A., Welch, I., 2008. A comprehensive look at the empirical performance of equity premium prediction. Rev. Financ. Stud. 21, 1455–1508.
Henkel, S.J., Martin, J.S., Nardari, F., 2011. Time-varying short-horizon predictability. J. Financ. Econ. 99, 560–580.
Hjalmarsson, E., 2010. Predicting global stock returns. J. Financ. Quant. Anal. 45, 49–80.
Hodrick, R.J., 1992. Dividend yields and expected stock returns: alternative procedures for inference and measurement. Rev. Financ. Stud. 5, 357–386.
Kothari, S.P., 2001. Capital markets research in accounting. J. Account. Econ. 31, 105–231.
Kothari, S.P., Shanken, J., 1997. Book-to-market, dividend yield, and expected market returns: a time-series analysis. J. Financ. Econ. 44, 169–203.
Lettau, M., Ludvigson, S.C., 2001. Consumption, aggregate wealth, and expected returns. J. Finance 56, 815–849.
Lettau, M., Ludvigson, S.C., 2005. Expected returns and expected dividend growth. J. Financ. Econ. 76, 583–626.
Lintner, J., 1956. Distribution of incomes of corporations among dividends, retained earnings, and taxes. Am. Econ. Rev. 46, 97–113.
Maio, P., Santa-Clara, P., 2015. Dividend yields, dividend growth, and return predictability in the cross-section of stocks. J. Financ. Quant. Anal. 50, 33–60.
Marsh, T.A., Merton, R.C., 1987. Dividend behavior for the aggregate stock market. J. Bus. 60, 1–40.
Menzly, L., Santos, T., Veronesi, P., 2004. Understanding predictability. J. Polit. Economy 112, 1–47.
Nelson, C.R., Kim, M.J., 1993. Predictable stock returns: the role of small sample bias. J. Finance 48, 641–661.
Politis, D.N., White, H., 2004. Automatic block-length selection for the dependent bootstrap. Econ. Rev. 23, 53–70.
Pukthuanthong, K., Roll, R., 2009. Global market integration: an alternative measure and its application. J. Financ. Econ. 94, 214–232.
Rangvid, J., Schmeling, M., Schrimpf, A., 2014. Dividend predictability around the world. J. Financ. Quant. Anal. 49, 1255–1277.
Rapach, D.E., Strauss, J.K., Zhou, G., 2010. Out-of-sample equity premium prediction: combination forecasts and links to the real economy. Rev. Financ. Stud. 23, 821–862.
Rapach, D.E., Zhou, G., 2013. Forecasting stock returns. In: Elliott, G., Timmermann, A. (Eds.), Handbook of Economic Forecasting 2.
Sims, C.A., Stock, J.H., Watson, M.W., 1990. Inference in linear time-series models with some unit roots. Econometrica 58 (1), 113–144.
Stambaugh, R., 1999. Predictive regressions. Journal of Financial Economics 54, 375–421.
Thompson, S.B., 2011. Simple formulas for standard errors that cluster by both firm and time. J. Financ. Econ. 99, 1–10.
Valkanov, R., 2003. Long-horizon regressions: theoretical results and applications. J. Financ. Econ. 68, 201–232.
Zhu, X., 2015. Tug-of-wars: time-varying predictability of stock returns and dividend growth. Rev. Finance 19, 2317–2358.