

This file was downloaded from the institutional repository BI Brage - <http://brage.bibsys.no/bi> (Open Access)

***Dividend smoothing and predictability***

**Long Chen  
Cheung Kong Graduate School of Business**

**Zhi Da  
University of Notre Dame**

**Richard Priestley  
BI Norwegian Business School**

This is the authors' accepted and refereed manuscript to the article published in

***Management Science*, 58(2012)10:1834-1853**

DOI: <http://dx.doi.org/10.1287/mnsc.1120.1528>

The publisher, Informs, allows the author to retain rights to “post the working paper version (or the final version after peer review) on the institutional repository until one year after the paper has been published in the print edition of an INFORMS journal”. (Publisher’s general policy 2012)

# Dividend Smoothing and Predictability

Long Chen\*

Olin Business School and CKGSB  
Washington University in St. Louis

Zhi Da†

Mendoza College of Business  
University of Notre Dame

Richard Priestley‡

Norwegian School of Management

August 2011§

## Forthcoming, Management Science

### Abstract

The relative predictability of returns and dividends is a central issue since it forms the paradigm to interpret asset price variation. A little studied question is how dividend smoothing, as a choice of corporate policy, affects predictability. We show that, even if dividends are supposed to be predictable without smoothing, dividend smoothing can bury this predictability. Since aggregate dividends are dramatically more smoothed in the postwar period than before, the lack of dividend growth predictability in the postwar period does not necessarily mean that there is no cash flow news in stock price variations; rather, a more plausible interpretation is that dividends are smoothed. Using two alternative measures that are less subject to dividend smoothing – net payout and earnings – we reach the consistent conclusion that cash flow news plays a more important role than discount rate news in price variations in the postwar period.

JEL Classification: G12, E44

Key Words: Dividend-price ratio, earning-price ratio, dividend growth, earnings growth, return, predictability, dividend smoothing

---

\*Olin School of Business, Washington University in St. Louis. Tel: (314) 935-8374, and e-mail: lchen29@wustl.edu.

†239 Mendoza College of Business, University of Notre Dame, Notre Dame IN 46556. Tel: (574) 631-0354, and e-mail: zda@nd.edu.

‡Department of Financial Economics, BI Norwegian Business School, Nydalsveien 37, 0484 Oslo. Tel: (47) 46410515, and email: richard.priestley@bi.no.

§For helpful discussion and comments, we are grateful to Utpal Bhattacharya, Paul Burik, John Campbell, Chun Chang, Shane Corwin, Phil Dybvig, Wayne Ferson, Paul Gao, Tom George, Radhakrishnan Gopalan, Craig Holden, Ravi Jagannathan, Jennifer Juergens, Ralph Koijen (discussant at FRA), Tim Loughran, Bill MacDonald, Kumar Praveen, Paul Schultz, Sophes Shive, Charles Trzcinka, and seminar participants at 2009 Finance Research Association annual meeting, Beijing University, Hong Kong University, Hong Kong University of Science and Technology, Indiana University, Shanghai Advanced Institute of Finance, Tsinghua University, University of Notre Dame, University of Houston, University of Illinois at Chicago, Lancaster University, Tel Aviv University, and Washington University in St. Louis. The usual disclaimer applies.

# 1 Introduction

In their seminal paper, Miller and Modigliani (1961) argue forcefully that dividend policy is irrelevant: stock prices should be driven by “real” behavior – the earnings power of corporate assets and investment policy – and, crucially, not by how the earnings power is distributed.

Although dividends might not be relevant for stock prices, they are critical for economic analysis. To understand whether investors’ revised forecasts regarding future cash flows or discount rates are the drivers of price variation, economists usually compare the predictability of cash flows relative to that of stock returns.<sup>1</sup> “Predictability of dividends and/or returns form, in many ways, the rational paradigm to interpret asset price variation.” (Bansal and Yaron (2007)).<sup>2</sup>

The general conclusion of the extant literature is that in the postwar period the dividend-price ratio (i.e., dividend yield) can predict aggregate returns, but not dividend growth. This finding has led to the widely accepted view that almost all the variation in the dividend yield is driven by the variation in discount rates (Cochrane (1992, 2001, 2008) and Campbell and Ammer (1993)). However, Chen (2009) shows that dividend growth is strongly predictable by the dividend yield in 1872-1945 but this predictability completely disappears in the postwar period. This finding raises an interesting paradox since any conclusions regarding asset price variations based on the relative dividend growth/return predictability findings would be the *opposite* for the pre- and postwar periods.

What has caused such a dramatic change of predictability? How much of the inability of the dividend yield to predict dividend growth stems from the fact that over any period of time dividends can be arbitrary and delinked from asset prices? The answers to these questions are important since they shape our understanding of stock price movements.

We ask first whether firms smooth dividends more in the postwar period than the prewar period. We define dividend smoothing as the phenomenon that dividend payout is determined not only by current earnings (Lintner (1956)) or “permanent earnings” (Marsh and Merton (1987)), but also by past dividend payout. The evidence is compelling: dividend payout at the aggregate level has become much more smoothed. For example, applying Lintner’s (1956) model for 1871-1945, the speed of adjustment to target is 0.37; the corresponding number for 1946-2006 is 0.09. As another example, if we regress dividend change on its own lag, the coefficient on lagged dividend change is statistically insignificant at 0.061 for the prewar period; the corresponding coefficient is strongly significant at 0.687 in the postwar period. Dividend policy has evolved in such a way that its own lag has become its best predictor in the postwar period.

Having established the evidence of dividend smoothing, we then ask whether dividend smoothing affects predictability. Using simulation analysis, we start with the null hypothesis that dividends are predictable by the dividend yield. We then change the degree of dividend smoothing and adopt a dividend policy such that it is sustainable and the dividend yield is always within a sensible range.

---

<sup>1</sup>The idea is that, if cash flow growth rates and stock returns are predictable, the expected cash flow growth rates and the expected returns must be time-varying. Such variations must cause stock prices to change, and thus the relative predictability reveals which component is more important in driving price movements.

<sup>2</sup>For example, to explain the equity premium puzzle, Campbell and Cochrane (1999) focus on modeling the time-varying expected return while Bansal and Yaron (2004) model both expected return and dividend growth. As another example, see Ang and Liu (2004) for how to discount future cash flows using time-varying discount rates.

We find that introducing dividend smoothing can eliminate dividend growth predictability in a finite sample. Severe dividend smoothing also makes the dividend yield very persistent, a pattern evident in the data: its AR(1) coefficient is 0.557 in the prewar period, and 0.956 in the postwar period.

The combined evidence that (i) dividends are much more smoothed in the postwar period and (ii) dividend smoothing can severely affect predictability has the following implication: the lack of dividend growth predictability in the postwar period does not necessarily mean that aggregate stock price variations contain no cash flow news; rather, a more logical interpretation is that dividends are so smoothed that they do not reflect well future cash flows.

Since dividend smoothing makes the interpretation of the relative dividend/return predictability ineffective, we explore two alternative measures that are less subject to smoothing: net payout and earnings.<sup>3</sup> In both cases, we reach the same conclusion that is remarkably consistent for both the full and postwar samples. We find that the majority of the variation of the net payout (earnings) yield comes from net payout (earnings) growth, suggesting a role for cash flow news much larger than discount rate news. This conclusion contrasts with what we know through investigations of dividend growth predictability.

To further highlight the role of dividend smoothing in cash flow predictability, we sort firms into three portfolios based on how smooth a firm's dividend payout is. Smoothness is defined as the standard deviation of dividend growth divided by the standard deviation of earnings growth. Interestingly, in the postwar period dividend growth is predictable by the dividend yield for the least-smoothed portfolio, but not so for the most-smoothed portfolio. The evidence for the most smoothed portfolio suggests that, for the postwar period, more than 100% of the dividend yield variance is driven by discount rate news, a result that is widely accepted in the current literature. In stark contrast, the evidence for the least smoothed portfolio suggests that 70% (30%) of the variance is driven by cash flow (discount rate) news. Further confirming the evidence, we find that earnings growth is predictable for both portfolios in the postwar period. In this case, cash flow news, as measured by earnings growth is responsible for almost all the variation in the earnings yield.

To our best knowledge, this is the first paper that formally studies the role of dividend smoothing on predictability and the interpretation of price variation.<sup>4</sup> Given that dividends are widely regarded as the measure of cash flow to shareholders, and that dividends can be easily manipulated by firms, understanding the impact of dividend smoothing seems important. This study fills this void by building a bridge between corporate policy and asset pricing.<sup>5</sup>

---

<sup>3</sup>The benefit of using earnings as the meaningful measure of cash flows is summarized by Miller and Modigliani (1961): "We can follow the standard practice of the security analyst and think in terms of price per share, dividends per share, and the rate of growth of dividends per share; or we can think in terms of the total value of the enterprise, total earnings, and the rate of growth of total earnings. Our own preference happens to be for the second approach primarily because certain additional variables of interest — such as dividend policy, leverage, and size of firm — can be incorporated more easily and meaningfully into test equations in which the growth term is the growth of total earnings."

<sup>4</sup>There is a voluminous literature that relies on the relative extent of return and dividend growth predictability to interpret price variation. This literature includes, among others, Campbell and Shiller (1988, 1998), Cochrane (1992, 2001, 2008), Ang (2002), Goyal and Welch (2003), Lettau and Ludvigson (2005), Lettau and Nieuwerburgh (2008), Ang and Bekaert (2007), Binsbergen and Koijen (2010), Chen (2009), and Chen and Zhao (2009).

<sup>5</sup>Chen (2009) also asks whether dividend smoothing has contributed to the lack of dividend predictability in the

Our finding, through simulation, that dividend smoothing can affect predictability is not trivial. The general belief is that one cannot “hide cash flows” forever. Our contribution is to show that this belief does not necessarily translate into predictability. Dividends might not be predictable by the dividend yield even for long samples.

Many studies have used payout yield or earnings yield to predict returns.<sup>6</sup> This is the first study to show that (i) the relative equity return and cash flow predictability is stable for both the long sample and the postwar sample once one does not rely on dividends, and (ii) the lack of dividend predictability in the postwar period only applies to firms with strong dividend smoothing. These new pieces of evidence, together with our simulation results and the finding that dividends are much more smoothed in the postwar period, provide strong support to the main conclusion.

Differently, Lettau and Ludvigson (2005) point out that the comovement between dividends and prices can make dividends less predictable by dividend yield. Lacerda and Santa-Clara (2010) adjust the dividend price ratio by the past average rate of dividend growth in order to better predict returns. Binsbergen and Koijen (2010) show that dividend growth is predictable based on past values of dividend growth, but they do not find significant predictability using the dividend yield.<sup>7</sup>

The remainder of the paper is organized as follows. Section 2 provides a theoretical motivation on why dividend smoothing might affect predictability. Section 3 provides empirical evidence regarding the aggregate dividend behavior. Section 4 studies whether dividend smoothing affects predictability. The predictability of dividend growth, net payout growth, earnings growth, and returns is assessed in section 5. Section 6 concludes.

## 2 Theoretical motivation

Campbell and Shiller (1988) show that the log dividend yield, suppressing a constant, can be approximated as

$$d_t - p_t = E_t \left[ \sum_{j=0}^{\infty} \rho^j r_{t+1+j} \right] - E_t \left[ \sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} \right], \quad (1)$$

where  $d_t$  is log dividend,  $p_t$  is log price,  $r_{t+1+j}$  is log return, and  $\Delta d_{t+1+j}$  is log dividend growth. Equation (1) says that the log dividend yield is the difference between expected future returns and expected future dividend growth. It follows that the variation of the dividend yield must predict

---

postwar period. To answer this question, he examines whether the book-to-market ratio can predict the earnings return on equity and finds the answer is no. But he does not provide any evidence on increased dividend smoothing in the postwar period; nor does he investigate how dividend smoothing affects predictability.

<sup>6</sup>A partial list includes Vuolteenaho (2000), Bansal, Khatchatrian, and Yaron (2005), Robertson and Wright (2006), Boudoukh, Michaely, Richardson, and Roberts (2007), Bansal and Yaron (2007), Ang and Bekaert (2007), Sadka (2007), Larrain and Yogo (2008), Hansen, Heaton, and Li (2008), and Pontiff and Woodgate (2008).

<sup>7</sup>Our conclusion that dividend smoothing might have contributed to the lack of dividend growth predictability is consistent with the conclusion by Mankiw and Miron (1986) that interest rate smoothing by the Federal Reserve might have led to the lack of interest rate predictability. In the same vein, both Engsted and Pedersen (2009) and Rangvid, Schmeling, and Schrimpf (2010) show that dividend growth is predictable in countries where dividends smoothing is much less pronounced.

the revisions to the two expectation components. This identity has inspired economists to examine whether expected returns or expected dividend growth is more predictable by the dividend yield. In doing so, the key objective is to understand why stock prices vary.

This predictive regression approach is potentially problematic. The rationale for running predictive regressions is to understand whether price variation contains news about future cash flows. However, if dividends do not vary according to the outlook of future cash flows, then it deems the exercise of predictive regressions futile in a finite sample.

To understand the issue, consider the Lintner (1956) partial adjustment model in log form as an illustration:

$$\Delta d_{t+1} = \alpha_0 + \alpha_1 e_{t+1} + \alpha_2 d_t + u_{t+1}, \quad (2)$$

where  $e_{t+1}$  is earnings and  $u_{t+1}$  is an error term. Rewrite (2) in terms of differences:

$$\Delta d_{t+1} - \Delta d_t = \alpha_1 \Delta e_{t+1} + \alpha_2 \Delta d_t + \Delta u_{t+1}, \quad (3)$$

or

$$\Delta d_{t+1} = \alpha_1 \Delta e_{t+1} + (1 + \alpha_2) \Delta d_t + \Delta u_{t+1}. \quad (4)$$

Dividends are most smoothed if  $\alpha_1 = 0$  and  $\alpha_2 = 0$ , in which case dividends grow at a constant rate plus some noise.

The summation of dividend growth is

$$\begin{aligned} \sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} &= \text{constant} + \frac{(1 + \alpha_2)}{1 - (1 + \alpha_2) \rho} \Delta d_t + \frac{\alpha_1}{1 - (1 + \alpha_2) \rho} \sum_{j=0}^{\infty} \rho^j \Delta e_{t+1+j} \\ &+ \frac{1}{1 - (1 + \alpha_2) \rho} \sum_{j=0}^{\infty} \rho^j u_{t+1+j}. \end{aligned} \quad (5)$$

Suppressing the constant, the dividend yield can then be written as

$$d_t - p_t = E_t \left[ \sum_{j=0}^{\infty} \rho^j r_{t+1+j} \right] - E_t \left[ \sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} \right] \quad (6)$$

$$= E_t \left[ \sum_{j=0}^{\infty} \rho^j r_{t+1+j} \right] - \left[ \frac{(1 + \alpha_2)}{1 - (1 + \alpha_2) \rho} \Delta d_t + \frac{\alpha_1}{1 - (1 + \alpha_2) \rho} E_t \left( \sum_{j=0}^{\infty} \rho^j \Delta e_{t+1+j} \right) \right] \quad (7)$$

$$= \text{Discount rate component} - [\text{Smoothing component} + \text{Earnings component}], \quad (8)$$

where

$$\text{Discount rate component} = E_t \left[ \sum_{j=0}^{\infty} \rho^j r_{t+1+j} \right], \quad (9)$$

$$\text{Smoothing component} = \frac{(1 + \alpha_2)}{1 - (1 + \alpha_2)\rho} \Delta d_t, \quad (10)$$

$$\text{Earnings component} = \frac{\alpha_1}{1 - (1 + \alpha_2)\rho} E_t \left( \sum_{j=0}^{\infty} \rho^j \Delta e_{t+1+j} \right). \quad (11)$$

The intuition is as follows. The smoothing component is deterministic as it is known at time  $t$ . Given  $\Delta d_t$ , one knows precisely its contribution to future dividend payout as a result of dividend smoothing. If dividends are very smoothed (i.e., both  $\alpha_1$  and  $\alpha_2$  are close to zero), the variation of dividend growth is not informative of future cash flows. The earnings component is important because its variation represents cash flow news.<sup>8</sup>

The above theoretical discussion indicates that dividend smoothing could defeat the purpose of predictive regressions using dividend growth. If so, it could explain two puzzling findings: first, Chen (2009) finds that dividend growth is strongly predictable during the prewar period but is not predictable in the postwar period; second, only discount rate news appears to be important in asset price variations.

Based on this discussion, we ask three questions in sequence: (i) are dividends more smoothed in the postwar period? (ii) does dividend smoothing affect predictability? and (iii) do alternative cash flow measures that are less smoothed address the issue?

### 3 Are dividends more smoothed in the postwar period?

#### 3.1 Dividend policy models

Lintner (1956) proposes the following partial-adjustment model of dividend-setting behavior:

$$\Delta D_t = \alpha_0 + \alpha_1 E_t + \alpha_2 D_{t-1} + u_t \quad (12)$$

where  $\Delta D_t$  is the change of the level of dividends,  $E_t$  is earnings and  $u_t$  is an error term. In this equation  $-\alpha_1/\alpha_2$  is the target payout ratio (TPR) and  $-\alpha_2$  is the speed of adjustment (SA) to the target. Equation (12) is the first dividend policy model we will estimate.

This model posits that over the long term firms aim at paying a constant fraction of earnings in the form of dividends. Under this policy, a positive earnings shock would imply additional dividend payout but firms often respond by increasing the dividend only by a portion of the dividend hike implied by the target payout ratio. This portion is also known as the speed of adjustment and reflects

---

<sup>8</sup>One could argue that dividend smoothing in this illustrative example may not be sustainable in the long run in some states of the world. For this reason, in our formal simulation exercise later on, we impose constraints to ensure that the dividend policy is always sustainable.

the intention of firms to avoid having to cut dividends in response to negative shocks to earnings.

If we take the first difference of equation (12), we obtain the second testable model:<sup>9</sup>

$$\Delta D_t = \beta_0 + \beta_1 \times \Delta E_t + \beta_2 \times \Delta D_{t-1} + \varepsilon_t. \quad (13)$$

The advantage of equation (13) is that the variables on the right hand side are not persistent. In this equation  $1 - \beta_2$  is the speed of adjustment and thus  $\beta_2$  measures the degree of smoothness.

In a third variation of the dividend policy model, we estimate

$$\Delta D_t = \gamma_0 + \gamma_1 E_t + \gamma_2 \times \Delta D_{t-1} + v_t. \quad (14)$$

Equation (14) is the same as equation (12) except that the lagged change of dividends is used as the regressor. Since this deviates from the Linter's model, our focus is on interpreting the persistence parameter  $\gamma_2$ . The higher  $\gamma_2$  is, the more smoothed is the dividend payout.

One drawback of the variants of Lintner's model is that they do not specify whether the dividend-smoothing behavior can be sustained. Addressing this issue, Marsh and Merton (1987) develop a model in which dividend payouts not only respond to permanent earnings in the short run, but converge to a steady-state target ratio in the long run. This is an error-correction model and can be written as

$$\ln \left[ \frac{D_{t+1}}{D_t} \right] + \frac{D_t}{P_{t-1}} = \lambda_0 + \lambda_1 \times \ln \left[ \frac{P_t + D_t}{P_{t-1}} \right] + \lambda_2 \times \ln \left[ \frac{D_t}{P_{t-1}} \right] + \varpi_{t+1}, \quad (15)$$

where  $\lambda_1$  captures how much dividends respond to permanent earnings changes. The implicit assumption is that price changes adequately capture information regarding changes in permanent earnings. Accordingly, a higher  $\lambda_1$  means less dividend smoothing;  $\lambda_2$  is supposed to be negative and  $-\lambda_2$  captures the speed of convergence to the long-term target: a higher  $-\lambda_2$  (in magnitude) also implies less dividend smoothing.

We will estimate these four versions of dividend policy models. The goal is to examine whether we can draw consistent conclusions without relying on a particular statistical specification.

### 3.2 Evidence on dividend smoothing

We use the annual S&P index data, obtained from Robert Shiller's website, to conduct the dividend policy tests. The data cover 1871-2006. The 1871-1925 sample presumably covers all stocks traded on NYSE during the period (Schwert (1990)); the 1926-2006 sample includes the S&P index firms.

Table 1 reports the summary statistics of the sample. We call 1872-1945 the prewar period and 1946-2006 the postwar period. In Panel A, the average log dividend growth in the prewar period is 1.3% with a standard deviation of 16%; the corresponding postwar growth rate is 5.9% with a standard deviation of 5%. Therefore, the average dividend growth rate has largely increased while the volatility has largely decreased.

---

<sup>9</sup>For equation (12) to be fully consistent with equation (13),  $\beta_0$  should be zero. In the empirical tests, we find that whether  $\beta_0$  is zero or not makes little difference on other estimated parameters. In light of this, we estimate all the models with a constant.

Table 1: : **Summary Statistics**

In Panel A, we summarize the annual S&P index.  $\Delta d$  is the log dividend growth rate;  $\Delta e$  is the log earnings growth rate;  $\frac{D}{P}$  is the dividend yield;  $\frac{E}{P}$  is the earnings yield;  $\frac{D}{E}$  is the payout ratio; and  $S$  is the standard deviation of dividend growth divided by the standard deviation of earnings growth, which is a measure of dividend smoothing. The data cover 1872-2006. In Panel B, we use data constructed from merging CRSP, COMPUSTAT, and Moody's book equity. The total payout includes dividend and repurchase. The net payout is total payout minus equity issuance. This sample covers 1928-2006.

		Panel A: S&P					
		$\Delta d$	$\Delta e$	$\frac{D}{P}$	$\frac{E}{P}$	$\frac{D}{E}$	$S = \frac{\sigma(\Delta d)}{\sigma(\Delta e)}$
1872-2006	<i>Mean</i> ( <i>sd</i> )	0.034 (0.12)	0.039 (0.25)	0.045 (0.02)	0.075 (0.03)	0.618 (0.20)	0.500
	AR(1)	0.256	0.024	0.781	0.740	0.632	
1872-1945	<i>Mean</i> ( <i>sd</i> )	0.013 (0.16)	0.012 (0.29)	0.053 (0.14)	0.077 (0.03)	0.719 (0.21)	0.545
	AR(1)	0.204	-0.017	0.518	0.621	0.440	
1946-2006	<i>Mean</i> ( <i>sd</i> )	0.059 (0.05)	0.073 (0.18)	0.036 (0.01)	0.073 (0.03)	0.497 (0.09)	0.295
	AR(1)	0.473	0.089	0.926	0.832	0.649	
		Panel B: CRSP (D = dividend + repurchase)					
		$\Delta d$	$\Delta e$	$\frac{D}{P}$	$\frac{E}{P}$	$\frac{D-I}{P}$	$S = \frac{\sigma(\Delta d)}{\sigma(\Delta e)}$
1928-2006	<i>Mean</i> ( <i>sd</i> )	0.054 (0.15)	0.066 (0.53)	0.045 (0.01)	0.072 (0.04)	0.022 (0.02)	0.283
	AR(1)	0.115	-0.124	0.637	0.588	0.666	
1946-2006	<i>Mean</i> ( <i>sd</i> )	0.069 (0.10)	0.082 (0.48)	0.042 (0.01)	0.074 (0.04)	0.017 (0.02)	0.216
	AR(1)	-0.081	-0.139	0.765	0.734	0.723	

The reduction of dividend growth volatility is consistent with dividend smoothing; it could also be due to the volatility reduction of the aggregate economy. We are thus more interested in the reduction of dividend volatility relative to the reduction of earnings volatility. To this end, we define the smoothness parameter as

$$S = \frac{\sigma(\Delta d)}{\sigma(\Delta e)}, \quad (16)$$

where  $\sigma(\Delta d)$  is the volatility of dividend growth and  $\sigma(\Delta e)$  is the volatility of earnings growth (see also Leary and Michaely (2010)). The smoothness parameter is 0.545 in the prewar period but only 0.295 in the postwar period, suggesting that dividends are indeed much more smoothed in the postwar period. Another piece of supporting evidence is that, for the prewar (postwar) period, the dividend yield AR(1) coefficient is 0.518 (0.926). Interestingly, the AR(1) coefficient for the earnings yield is 0.621 (0.832) in the prewar (postwar) period. Therefore, dividend growth is less (more) persistent than earnings growth in the prewar (postwar) period. Panel B reports similar statistics for total payout yield (= (dividend+repurchase)/price) and net payout yield (= (dividend+repurchase-equity issuance)/price).<sup>10</sup> The results in Panel B suggest that smoothing is much less likely a problem for payouts other than dividends.

Figure 1 plots the dividend growth and earnings growth during 1872-2006. Both growth rates are volatile and trace each other quite well in the first period leading up to the end of 1940s.

<sup>10</sup>See Section 5.1 for data construction.

Subsequently, dividend growth becomes much less volatile than earnings, less dependent on earnings and more dependent on its own lag, confirming the evidence in Table 1.



**Figure 1. Dividend and Earnings Growth Rates** Annual dividend growth (DG) and earnings growth (EG) rates during 1872-2006. Data are downloaded from Robert Shiller's website.

We next estimate the four dividend behavior models. Panel A of Table 2 reports the estimates from the standard Lintner model where we find that the speed of adjustment coefficient,  $SA$ , is 0.373 in the prewar period and only 0.090 in the postwar period. The final column of Panel A reports a Chow test that indicates a significant structural break around 1945. We also report two F-tests of the null hypothesis that the estimated coefficients are the same in each sample; in both cases the null is clearly rejected. We find similar evidence in Panel B where we use first differences of the independent variables.

In Panel C of Table 2, we report estimates for the third model. The coefficient on the lagged change in dividends is statistically insignificant at 0.061 for the prewar period. In stark contrast, the coefficient is highly significant at 0.687 for the postwar period. Therefore, dividend policy has evolved from little dependence on the lagged dividends in the prewar period to heavy dependence in the postwar period. This finding is consistent with the survey by Brav, Graham, Harvey, and Michaely (2005), in which the managers acknowledge the importance of maintaining the level of dividends but show little willingness to change dividends beyond that.

In Panel D, which reports the Marsh and Merton (1987) model, the coefficient that measures the response to permanent earnings change is 0.673 during 1872-1945 and the implied convergence coefficient is -0.198, both highly significant. These coefficients say that aggregate dividends respond strongly to permanent earnings changes and converge to a long-term target. In contrast, in the postwar period, the response coefficient is 0.003, statistically insignificant, and the implied convergence coefficient is 0.061 indicating no convergence. The Chow test indicates a strong structural break around 1945. Therefore, the overwhelming statistical evidence is that dividends are much more smoothed in the postwar period than in the prewar period.<sup>11</sup>

Figure 2 plots the rolling-regression coefficients and their  $t$ -statistics for the three Lintner dividend models, with a rolling window of 30 years. In the first panel for the standard Lintner model, we observe a relatively stable speed-of-adjustment coefficient, around 0.3, between 1872 and the mid 1940s; this coefficient then quickly drops and approaches zero toward the end of the sample. We find a qualitatively similar pattern in the second panel for the second model. In the third panel, the coefficient on the lagged dividend change fluctuates around zero from 1872 until the early 1940s; it then quickly jumps up and approaches 0.7 towards the end of the sample.

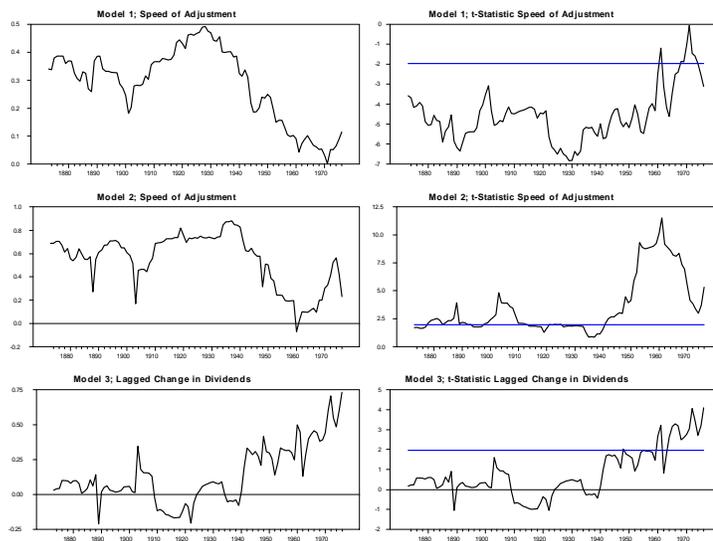
---

<sup>11</sup>We have also tested the four models in log form and find very similar results. For brevity we do not report them.

Table 2: : Dividend Policy Models Using Actual Dividends and Earnings

Denote  $D_t$  the level of dividends,  $E_t$  the level of earnings, and  $\Delta$  the change operator. Four dividend behavior models are estimated. The first is the original Lintner (1956) model and the second is estimated using the first differences. For these two models the speed of adjustment ( $SA$ ) and the target payout ratio ( $TPR$ ) are implied. The focus of the third models is the coefficient on the lagged  $\Delta D_t$ , which measures persistence (smoothness). The fourth is the Marsh and Merton (1987) model, in which  $\lambda_1$  measures response to permanent earnings change and  $-\lambda_2$  measures speed of convergence to long-term target. Newey-West  $t$ -values are provided below each coefficient controlling for heteroskedasticity and autocorrelation. We also report the Chow test for structural break around 1945. The full sample is the S&P 500 annual data covering 1872-2006.

Panel A: $\Delta D_t = \alpha_0 + \alpha_1 E_t + \alpha_2 D_{t-1} + u_t$							
	$c$	$E_t$	$D_{t-1}$	$\bar{R}^2$	SA	TPR	Chow 1945
1872-2006	0.035 (1.42)	0.052 (10.99)	-0.079 (-5.87)	0.73	0.08	0.08	2.656 [0.05]
1872-1945	0.005 (0.32)	0.248 (10.22)	-0.373 (-8.93)	0.60	0.37	0.18	
1946-2006	0.120 (1.74)	0.054 (7.69)	-0.090 (-4.25)	0.68	0.09	0.05	
$F - Test$		766.43 [0.00]	175.08 [0.00]				
Panel B: $\Delta D_t = \beta_0 + \beta_1 \times \Delta E_t + \beta_2 \times \Delta D_{t-1} + \varepsilon_t$							
	$c$	$\Delta E_t$	$\Delta D_{t-1}$	$\bar{R}^2$	SA	TPR	Chow 1945
1872-2006	0.025 (1.38)	0.037 (7.30)	0.825 (17.25)	0.81	0.17	0.22	3.677 [0.01]
1872-1945	0.001 (0.20)	0.237 (6.09)	0.284 (2.94)	0.35	0.72	0.33	
1946-2006	0.062 (1.47)	0.036 (5.07)	0.808 (10.91)	0.79	0.19	0.19	
$F - Test$		773.78 [0.00]	50.15 [0.00]				
Panel C: $\Delta D_t = \gamma_0 + \gamma_1 E_t + \gamma_2 \times \Delta D_{t-1} + v_t$							
	$c$	$E_t$	$\Delta D_{t-1}$	$\bar{R}^2$			Chow 1945
1872-2006	-0.012 (-0.57)	0.011 (5.29)	0.652 (8.46)	0.78			1.311 [0.27]
1872-1945	-0.056 (-2.90)	0.093 (3.45)	0.061 (0.53)	0.15			
1946-2006	-0.025 (-0.47)	0.011 (3.29)	0.687 (4.04)	0.75			
$F - Test$		618.87 [0.00]	30.39 [0.00]				
Panel D: $\ln \left[ \frac{D_{t+1}}{D_t} \right] + \frac{D_t}{P_{t-1}} = \lambda_0 + \lambda_1 \times \ln \left[ \frac{P_t + D_t}{P_{t-1}} \right] + \lambda_2 \times \ln \left[ \frac{D_t}{P_{t-1}} \right] + \varpi_{t+1}$							
	$\lambda_0$	$\ln \left[ \frac{P_t + D_t}{P_{t-1}} \right]$	$\ln \left[ \frac{D_t}{P_{t-1}} \right]$	$\bar{R}^2$			Chow 1945
1872-2006	-0.026 (-0.33)	0.461 (6.21)	-0.021 (-0.89)	0.38			21.24 [0.00]
1872-1945	-0.565 (-2.60)	0.673 (9.01)	-0.198 (-2.72)	0.62			
1946-2006	0.299 (3.99)	0.003 (0.06)	0.061 (2.87)	0.18			
$F - Test$		176.49 [0.00]	246.24 [0.00]				



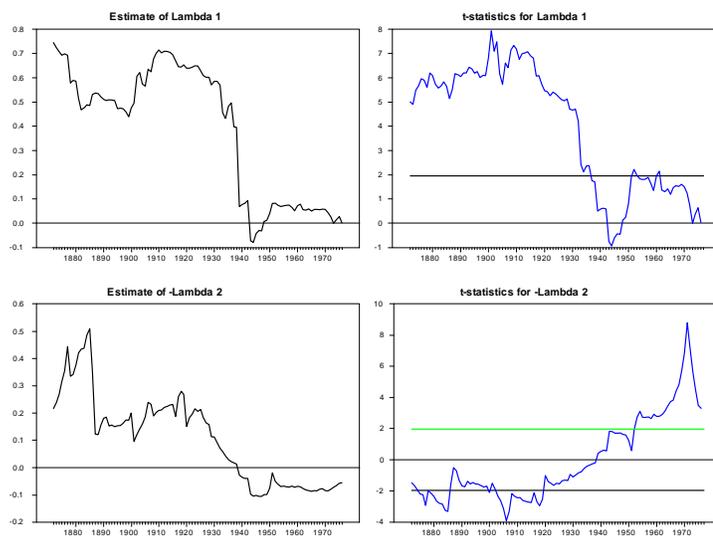
1

**Figure 2. Rolling-Window Regressions for the Lintner Model** All panels correspond to variants of Lintner's (1956) model (Equations (12)-(14)). The length of rolling window is 30 years. The first two panels plot the rolling speed-of-adjustment coefficients and their Newey-West t-statistics. The third panel plots the coefficient on the lagged dividend change and its t-statistic.

Figure 3 plots the rolling parameters for the Marsh and Merton (1987) model. The response to permanent earnings parameter,  $\lambda_1$ , is between 0.4 and 0.75 from 1872 to the end of 1940s; it then quickly drops to close to zero and subsequently remains so. The convergence to the long-run target parameter,  $-\lambda_2$ , is between 0.1 and 0.5 from 1872 to the end of 1940s; it then quickly drops to be lower than zero and remains so. Figures 2 and 3 indicate that the drastically stronger pattern of dividend smoothing in the postwar period represents a genuine change of aggregate dividend behavior.<sup>12</sup> Fama and French (1988) also note that dividends are more smoothed in the postwar period. We reinforce their findings by extending the data back from 1926 to 1872 and forward from 1986 to 2006.

---

<sup>12</sup>We note that the dramatically increased dividend smoothing in the postwar period is unlikely to be driven by the changing composition of the S&P index firms. For example, the S&P index contains only 90 stocks from 1926 to 1957, and 500 firms after that; in comparison, the CRSP market portfolio already contains more than 500 firms in 1926, and more than 1000 firms in 1957 (e.g., Chen (2009)). Yet, we find the same change of dividend smoothing from prewar to postwar period if we use the CRSP market portfolio.



1

**Figure 3. Rolling-Window Regressions for the Marsh-Merton Model** The length of rolling window is 30 years. The first panel plots the response-to-permanent-earnings coefficient ( $\lambda_1$ ) and its Newey-West t-statistic. A higher coefficient means less dividend smoothing. The second panel plots the implied convergence-to-target coefficient ( $-\lambda_2$ ) and its Newey-West t-statistic. A higher coefficient means less dividend smoothing.

Why are dividends so much more smoothed in the postwar period? While there seems to be no authoritative studies on this issue, we can identify two potential explanations. The first is a more liberal attitude from investors toward dividend payout (Graham and Dodd (6th edition, 2008)) and a reluctance to accept dividend cuts (Lintner (1956)).<sup>13</sup> This combination suggests that managers will try to (i) pay low dividends when they can (Graham and Dodd (2008)) and (ii) smooth dividends since they are sticky once increased.

A second story is that equity financing has become cheaper, a trend that makes dividend smoothing less costly since managers can use equity repurchase and issuance to adjust payout and funds. In this story what managers target is not dividends, but net payout (i.e., dividends plus repurchase minus equity issuance).<sup>14</sup>

Regardless of the interpretation, aggregate dividends are much more smoothed in the postwar period than earlier. The next natural question is the role of dividend smoothing on predictability.

## 4 How does dividend smoothing affect predictability?

Consider a VAR consisting of the log dividend yield ( $dp_t$ ), the dividend growth rate ( $g_t$ ), and returns ( $r_t$ ),

$$dp_{t+1} = a_{dp} + \phi \times dp_t + \varepsilon_{t+1}^{dp} \quad (17)$$

$$g_{t+1} = a_g + b_g \times dp_t + \varepsilon_{t+1}^g \quad (18)$$

$$r_{t+1} = a_r + b_r \times dp_t + \varepsilon_{t+1}^r \quad (19)$$

One does not have to estimate all three equations. Cochrane (2008) shows that the VAR coefficients are linked:

$$b_r \approx 1 - \rho\phi + b_g, \quad (20)$$

where  $\rho$  is a linearization parameter ( $\approx 0.96$  for annual data).

Theoretically,  $b_g$  is expected to be negative if dividend growth is predictable – a higher dividend yield means that dividends will grow slower. With an increasing degree of dividend smoothing,  $b_g$  is expected to be smaller in magnitude. The reason is that when dividend growth is smoothed, it does not adequately reflect the outlook of future cash flows; the latter drives the variation of the dividend yield.

Dividend smoothing also makes the dividend yield more persistent, i.e.,  $\phi$  becomes larger. A more persistent dividend yield has two effects on predictability. First, it biases  $b_g$  to be more negative and  $b_r$  to be more positive in estimation (e.g., Stambaugh (1999) and Boudoukh, Richardson, and

---

<sup>13</sup>Written more than 50 years ago, Graham and Dodd (2008) point out that “in recent years there has been a definite tendency toward greater liberty in dividend payments.” This increased payout liberty, as they discuss, is partly due to the implementation (in 1936) and cancelation (in 1938) of a penalty tax on retained earnings. That is, a policy meant to force dividend payout backfired and caused a more liberal attitude toward dividend payout.

<sup>14</sup>Consistent with this story, in untabulated results, we find that net payout, in contrast to dividends, is not more smoothed in the postwar period.

Whitelaw (2008)). Second, equation (20) says that, holding all else constant, a higher  $\phi$  makes either  $b_r$  or  $b_g$  smaller in magnitude, i.e., less predictable.

How does dividend smoothing affect return predictability? From equation (20), since it makes  $b_g$  smaller in magnitude but  $\phi$  bigger, the net effect on  $b_r$  is not clear. In addition, a higher  $\phi$  biases returns to appear to be more predictable.

It might appear that we already know how dividend smoothing affects predictability. However, the impact of dividend smoothing is likely to be mitigated in a long sample. Therefore, it is unclear how a sustainable dividend policy, with different degrees of smoothing, affects predictability in a finite sample. This issue has been largely neglected in the literature.

**The benchmark** Before we investigate how a sustainable dividend policy affects predictability, we report in Panel A of Table 3 the regressions of dividend growth and returns on the lagged dividend yield, for prewar and postwar periods separately. Following Kendall (1954), Stambaugh (1999), and Boudoukh, Richardson, and Whitelaw (2006), we simulate the  $p$ -values that consider the contemporaneous correlation between the independent and dependent variables, the persistence of the independent variable, and the overlapping nature of the variables when conducting long-horizon tests. The details of the simulation are provided in the appendix. We boldface the simulated  $p$ -values that are smaller or equal to 10%.

Dividend growth is strongly predictable during the prewar period: the one, three, and five-year coefficients are -0.448, -.596, and -0.406 respectively, and are highly statistically significant.<sup>15</sup> In comparison, the one-year return coefficient is 0.024 and is insignificant; the three and five-year return coefficients are 0.303 and 0.636 and are significant. Overall, during the prewar period dividend growth is strongly predictable and returns are less predictable, especially at the short horizon.<sup>16</sup>

Dividend growth is not predictable in the postwar period: the one, three, and five-year coefficients are all insignificant with the wrong sign: 0.026, 0.076, and 0.088 respectively. Stock returns appear to be more predictable at the one-year horizon in the postwar period than in the prewar period, but none of the coefficients for the postwar period are significant. The fact that stock return predictability (by dividend yield) lacks statistical power is well documented (e.g., Stambaugh (1999) and Cochrane (2008)).

Another important piece of evidence is that the dividend yield is much more persistent in the postwar period than in the prewar period (Table 1). The empirical evidence documented above will serve as the benchmark case when we analyze the impact of dividend smoothing below.

---

<sup>15</sup>The lack of monotonicity in the coefficients is related to the Great Depression period. In particular, the dividend-price ratio was very low reflecting high equity valuation before the 1929 stock market crash, and the dividend growth collapsed for a few years after the crash, opposite to the prediction of a low dividend-price ratio. If we remove a few years surrounding 1929 from our sample, the coefficients will be monotonically decreasing as the horizon increases.

<sup>16</sup>Chen (2009) shows that, for 1872-1945, returns are not predictable beyond the five-year horizon. In contrast, dividends are much more predictable at 15-year and 20-year horizons.

Table 3: : Predictability by Dividend Yield in the S&P Sample: Empirical and Simulation Evidence (First case)

We examine the S&P 500 annual data covering 1872-2006. In Panel A, we regress cumulative log dividend growth or returns, from one to five years, on the lagged log dividend yield, for 1872-1945 and 1946-2006 separately. For example,  $dg_t^1$  is the annual dividend growth,  $dg_t^5$  is the five-year dividend growth,  $r_t^1$  is annual return, and  $r_t^5$  is the five-year return. We provide the simulated  $p$ -values below each coefficients. The simulation considers the biases caused by the persistence of the variables, the contemporaneous correlation between the dependent and independent variables, and the overlapping small sample. We boldface the  $p$ -value if it is lower than or equal to 0.10. In Panel B, we regress simulated cumulative log dividend growth or returns, from one to five years, on the lagged simulated log dividend yield. We first fit the Marsh and Merton (1987) dividend smoothing model for 1872-1945 and 1946-2006 separately. We then simulate dividend growth using the fitted model and simulate returns under the null of no predictability. We match the historical means and standard deviations of dividend growth and return and the covariance between them. We back out the stock price from the simulated total return and dividend, and then calculate the dividend yield. We also set the maximum and minimum log dividend yields to be -1 and -10 and adjust dividends (when needed) to ensure that the dividend policy is sustainable. We report the regression coefficients and the associated  $p$ -values and the AR(1) coefficient for the log dividend yield.

Panel A: Actual Data							
	$dg_t^1$	$dg_t^3$	$dg_t^5$	$r_t^1$	$r_t^3$	$r_t^5$	AR(1)
1872-1945	-0.448 <b>[0.00]</b>	-0.596 <b>[0.00]</b>	-0.406 <b>[0.07]</b>	0.024 [0.27]	0.303 <b>[0.07]</b>	0.636 <b>[0.01]</b>	0.557 <b>[0.00]</b>
1946-2006	0.026 [0.35]	0.076 [0.25]	0.088 [0.26]	0.101 [0.17]	0.289 [0.19]	0.505 [0.15]	0.956 <b>[0.00]</b>
Panel B: Simulated Data							
	$dg_t^1$	$dg_t^3$	$dg_t^5$	$r_t^1$	$r_t^3$	$r_t^5$	AR(1)
1872-1945	-0.460 <b>[0.00]</b>	-0.679 <b>[0.00]</b>	-0.798 <b>[0.00]</b>	0.032 [0.37]	0.086 [0.34]	0.135 [0.32]	0.565 <b>[0.00]</b>
1946-2006	-0.033 [0.43]	-0.096 [0.42]	-0.156 [0.41]	0.020 [0.17]	0.058 [0.18]	0.094 [0.19]	0.983 <b>[0.00]</b>

## 4.1 Three cases of dividend smoothing

### 4.1.1 First case

We simulate three cases of dividend smoothing. In the first case, we first fit the Marsh and Merton (1987) dividend smoothing model (equation (15)) for the prewar and postwar periods separately, as shown in Panel D of Table 2. We then simulate dividend growth using the fitted equation (15). We also simulate returns under the null that returns are not predictable,

$$r_{t+1} = a_r + \varepsilon_{t+1}^r, \quad (21)$$

where  $a_r$  is a constant and  $\varepsilon_{t+1}^r$  the residual. We match the historical means and standard deviations of dividend growth and returns and the covariance between them. We back out stock prices from the simulated total return and dividend series, and then calculate the dividend yield. We also set the maximum and minimum log dividend yields to be -1 and -10 and once these points are reached we adjust the dividends to bring the dividend yield within the acceptable range. In this way, we ensure that the dividend policy is sustainable.

We perform 10,000 simulations, each time matching the sample size of the postwar data. For each simulation, we regress dividend growth and returns on the lagged dividend yield, for one, three and five years. Panel B of Table 3 reports the results for the dividend smoothing model that fits the prewar data. Similar to the actual data, dividend growth is strongly predictable: the coefficients are -0.460, -0.679, and -0.798 for one, three, and five-year horizons (the historical counterparts are -0.448, -.0596, and -0.406 respectively). Compared to the actual data, the return coefficients are small and insignificant. Panel B of Table 3 also reports the results for the dividend smoothing model that fits the postwar data. With highly smoothed dividends, dividend growth is not predictable at either one, three, or five-year horizons. Stock returns are not predictable at any horizon.

Therefore, when dividends are not highly smoothed and when the predictability is on the dividend side, dividend growth predictability can be easily detected, as in the prewar case. In contrast, when dividends are highly smoothed, even though the null is that dividends are predictable, dividend growth might not be predictable.

Regressing the simulated log dividend yield on its own lag yields a coefficient of 0.565 for 1872-1945 and 0.983 for 1946-2006. These numbers are close to their empirical counterparts and support the earlier finding that dividends are much more smoothed in the postwar period.

### 4.1.2 Second case

In this case, we start with a “true world” without dividend smoothing that defines how dividends (before paying out) are generated. The null is that dividend growth is predictable without smoothing but return is not:

$$g_{t+1} = a_g - 0.1 \times dp_t + \varepsilon_{t+1}^g \quad (22)$$

$$r_{t+1} = a_r + \varepsilon_{t+1}^r, \quad (23)$$

where  $g_{t+1}$  is dividend growth rate. The coefficient of -0.1 is chosen based on equation 23 with a persistent dividend yield and an unpredictable return (see Cochrane (2008) for a similar choice). Given the “true world” without smoothing, we assume that the actual dividend growth is governed by a smoothness parameter  $\lambda$ :

$$g_{t+1} = (1 - \lambda) (a_g - 0.1 \times dp_t + \varepsilon_{t+1}^g) + \lambda \times (g_{ave} + \varepsilon_{t+1}^{ave}), \quad (24)$$

where  $g_{ave}$  is the historical average dividend growth rate and  $\varepsilon_{t+1}^{ave}$  is a shock to this target. The more smoothed the dividend policy, the higher  $\lambda$  is. The residuals  $\varepsilon_{t+1}^g$  and  $\varepsilon_{t+1}^r$  are chosen such that the historical variance-covariance matrix of dividend growth and returns in the prewar period is matched.

We simulate stock returns under the null of no predictability and simulate dividend growth according to equation (24). As in the first case, we back out new stock prices from the simulated total return and dividend series. We ensure that prices are always higher than dividends by adjusting dividends. In addition, whenever the dividend yield reaches an upper or lower limit, we adjust the dividends to pull the dividend yield back. In sum, our null is that stock returns are unpredictable, dividends are predictable but are also smoothed, and the dividend policy is sustainable.

We report the results in Panel A of Table 4. In the scenario of the “true world”, dividend growth is strongly predictable at all horizons and stock returns have insignificant but positive coefficients at all horizons. With increasing  $\lambda$ , the dividend yield becomes more and more persistent, as shown by the AR(1) coefficients, and the dividend growth coefficient steadily goes down. When  $\lambda$  is equal to 0.95, the AR(1) coefficient of dividend yield is 0.973, and the dividend growth coefficients become insignificant at the 5% level for one, three, and five-year horizons.

The point that dividends might not be predictable by the dividend yield when dividends are highly smoothed is not trivial. For example, a firm might find it difficult to pursue a constant dividend policy forever while keeping the dividend yield stationary. Indeed, in simulations we have to adjust dividends whenever the dividend yield hits some boundaries. The point of the simulations is to show that these adjustments do not necessarily translate into predictability. Therefore, the simulation cases, even when  $\lambda = 1$ , provide new information.

### 4.1.3 Third case

In this case, we first use the prewar data to obtain the following estimated equations:

$$g_{t+1} = -1.315 - 0.448 \times dp_t + \varepsilon_{t+1}^g \quad (25)$$

$$r_{t+1} = 0.142 + 0.024 \times dp_t + \varepsilon_{t+1}^r. \quad (26)$$

This set of equations show strong dividend growth predictability but little return predictability. We ask the following question: if the “true world” without smoothing in the postwar world is actually the same as the prewar world, except that dividends are smoothed, what kind of dividend growth predictability should we expect?

Table 4: : **Dividend Smoothing and Predictability by the Dividend Yield: Simulation Evidence (Second and Third Cases)**

Panel A reports the second case of simulation. We simulate dividend growth rates, returns and dividend yields under the null that dividend growth is predictable without smoothing but return is not:

$$\begin{aligned} g_{t+1} &= a_g - 0.1 \times dp_t + \varepsilon_{t+1}^g \\ r_{t+1} &= a_r + \varepsilon_{t+1}^r, \end{aligned}$$

where  $g_{t+1}$  is dividend growth rate and  $r_{t+1}$  stock return. The residuals  $\varepsilon_{t+1}^g$  and  $\varepsilon_{t+1}^r$  are chosen such that the historical variance-covariance matrix of dividend growth and return in the prewar period is matched. We assume that the actual dividend growth is governed by a smoothness parameter  $\lambda$ :

$$g_{t+1} = (1 - \lambda) (a_g - 0.1 \times dp_t + \varepsilon_{t+1}^g) + \lambda \times (g_{ave} + \varepsilon_{t+1}^{ave}),$$

where  $g_{ave}$  is historical average dividend growth rate and  $\varepsilon_{t+1}^{ave}$  is a shock to this target. The more smoothed the dividend policy, the higher  $\lambda$  is. We back out new prices from the simulated total returns and dividends.

Panel B reports the third case of simulation. We simulate dividend growth rates and returns from the fitted equations:

$$\begin{aligned} g_{t+1} &= (1 - \lambda) (-1.315 - 0.448 \times dp_t + \varepsilon_{t+1}^g) + \lambda \times (g_{ave} + \varepsilon_{t+1}^{ave}) \\ r_{t+1} &= 0.142 + 0.024 \times dp_t + \varepsilon_{t+1}^r. \end{aligned}$$

In both panels the more smoothed the dividend policy, the higher  $\lambda$  is. We match the standard deviations of dividend growth and return and the covariance between them. We back out new prices from the simulated total returns and dividends. We also set the maximum and minimum log dividend yields to be -1 and -10 and adjust dividends (when needed) to ensure that the dividend policy is sustainable. We regress simulated cumulative log dividend growth or returns, from one to five years, on the lagged simulated log dividend yield. We report the regression coefficients and the associated  $p$ -values and the AR(1) coefficient for the log dividend yield. We boldface the  $p$ -value if it is lower than or equal to 0.10.

	$dg_t^1$	$dg_t^3$	$dg_t^5$	$r_t^1$	$r_t^3$	$r_t^5$	AR(1)
Panel A: Second Case							
$\lambda = 0$	-0.124 <b>[0.00]</b>	-0.325 <b>[0.00]</b>	-0.477 <b>[0.00]</b>	0.052 [0.19]	0.146 [0.18]	0.227 [0.17]	0.903 <b>[0.00]</b>
$\lambda = 0.5$	-0.097 <b>[0.01]</b>	-0.272 <b>[0.01]</b>	-0.423 <b>[0.01]</b>	0.053 [0.11]	0.150 [0.11]	0.234 [0.12]	0.946 <b>[0.00]</b>
$\lambda = 0.95$	-0.081 <b>[0.09]</b>	-0.239 <b>[0.10]</b>	-0.389 <b>[0.10]</b>	0.037 [0.12]	0.105 [0.13]	0.164 [0.14]	0.973 <b>[0.00]</b>
$\lambda = 1$	-0.079 [0.15]	-0.233 [0.15]	-0.381 [0.16]	0.033 [0.13]	0.095 [0.14]	0.148 [0.15]	0.976 [0.00]
Panel B: Third Case							
$\lambda = 0$	-0.456 <b>[0.00]</b>	-0.809 <b>[0.00]</b>	-0.882 <b>[0.00]</b>	0.058 [0.29]	0.126 [0.29]	0.170 [0.28]	0.533 <b>[0.00]</b>
$\lambda = 0.5$	-0.236 <b>[0.00]</b>	-0.544 <b>[0.00]</b>	-0.708 <b>[0.00]</b>	0.062 [0.24]	0.150 [0.24]	0.214 [0.24]	0.769 <b>[0.00]</b>
$\lambda = 0.95$	-0.104 <b>[0.03]</b>	-0.305 <b>[0.03]</b>	-0.493 <b>[0.04]</b>	0.051 [0.08]	0.139 [0.10]	0.212 [0.11]	0.955 <b>[0.00]</b>
$\lambda = 1$	-0.096 [0.14]	-0.284 [0.14]	-0.462 [0.15]	0.039 [0.10]	0.105 [0.12]	0.159 [0.14]	0.970 <b>[0.00]</b>

To answer this question, we simulate dividend growth according to different degrees of smoothness:

$$g_{t+1} = (1 - \lambda) (-1.315 - 0.448 \times dp_t + \varepsilon_{t+1}^g) + \lambda \times (g_{ave} + \varepsilon_{t+1}^{ave}).$$

As before, stock prices are backed out from these simulations. Dividends are adjusted to ensure that the dividend yield is within the range identified earlier.

The results are reported in Panel B of Table 4. If dividends are not smoothed, then dividends are strongly predictable at all horizons, returns are not predictable, and the AR(1) coefficient of the log dividend yield is only 0.533. When  $\lambda$  is equal to 0.5, the one-year dividend growth coefficient drops to -0.236, about half of the corresponding number without smoothing; the AR(1) coefficient of the log dividend yield jumps to 0.769. This pattern continues as  $\lambda$  increases. In the extreme case of  $\lambda$  being equal to one, dividends are still supposed to be predictable since by construction we adjust dividends when the dividend yield reaches boundaries. The simulated AR(1) coefficient is 0.970 (compared to 0.956 for the postwar data); the dividend growth coefficient is insignificant at one to five-year horizons.

Thus far, we have shown that dividends might not be predictable by the dividend yield at a sample size similar to the postwar data. How about longer horizons? We find that the answer comes down to the particular assumptions on dividend policy. In particular, if the assumption is that the firm only pulls the dividend yield back within the boundaries, then dividends might not be predictable even at a horizon of 100 years. In one experiment, when we assume that the firm, whenever it reaches the boundary, pulls the dividend yield back to the starting point of the simulation, then dividends become predictable at the 100-year horizon. It seems difficult to pin down a general conclusion. What we can conclude is that, depending on the particular policies, dividends might not be predictable by the dividend yield at long horizons even though dividend yield does not “explode” in the long run.

Our empirical evidence and the simulation exercise provide a reasonable interpretation on why dividend growth is strongly predictable in the prewar period, but not so in the postwar period. In particular, the lack of dividend growth predictability in the postwar period does not necessarily imply that the variation in stock prices contains no information regarding future cash flows; rather, it might only mean that dividends are severely smoothed. As such, dividends are a poor measure of future cash flows, and it becomes pointless to infer cash flow predictability from dividend predictability.

## 5 Predictability and yield decomposition using alternative cash flow measures

A main purpose of running predictive regressions using the dividend yield is to conduct dividend yield decomposition, through which price variation can be understood. Since dividend smoothing makes this exercise ineffective, it is natural to conduct alternative yield decompositions using alternative cash flow measures. We explore this issue in this section.

To test return and cash flow predictability, we use CRSP data. Besides producing the more widely used market portfolio, the data also allow us to separately consider repurchases and equity

issuances.

## 5.1 Data construction

We follow Bansal, Dittmar, Lundblad (2005) and Larrian and Yogo (2008) to separately consider equity repurchase and issuance. In particular, denote  $n_t$  the number of shares (after adjusting for splits, stock dividends, etc. using the CRSP share adjustment factor) and  $P_t$  stock price. Then repurchases are defined as

$$rp = \frac{P_{t+1}}{P_t} \times \left[ 1 - \min \left( \frac{n_{t+1}}{n_t}, 1 \right) \right]. \quad (27)$$

When there is a repurchase,  $\frac{n_{t+1}}{n_t} < 1$  and  $\left[ 1 - \min \left( \frac{n_{t+1}}{n_t}, 1 \right) \right]$  is the proportional repurchase;  $rp$  then captures the repurchase return. Similarly, stock issue returns are defined as

$$si = \frac{P_{t+1}}{P_t} \times \left[ \max \left( \frac{n_{t+1}}{n_t}, 1 \right) - 1 \right]. \quad (28)$$

We calculate dividends, repurchases, and issues in dollars for each firm month, and sum them across months to get the annual numbers for each firm. We then merge this annual data with the COMPUSTAT annual tape. The COMPUSTAT data are used to calculate book equity following Cohen, Polk, and Vuolteenaho (2003). For earlier years when book equity is not available we use the book equity data from Davis, Fama, and French (2000). Earnings for each firm year are then obtained through the clean surplus formula:

$$E_t = B_t - B_{t-1} + RP_t - SI_t + D_t, \quad (29)$$

where  $E_t$  is earning in year  $t$ ,  $B_t$  is book equity,  $RP$  is repurchase,  $SI$  is share issuance, and  $D_t$  is dividend. The equation says that earnings are equal to the change of book equity plus repurchases and minus net issues; retained earnings plus dividends gives total earnings.<sup>17</sup> We then aggregate the data to obtain the market portfolio. The final annual data cover 1928-2006.

## 5.2 Predictability and yield decomposition: dividend yield

Table 5 reports results from running the following predictive regression:

$$y_t = \alpha_0 + \alpha_1 \times x_{t-1} + \varepsilon_t, \quad (30)$$

where  $y_t$  is either the cumulative log dividend growth ( $\Delta d_t$ ) or log returns ( $r_t$ ); and  $x_{t-1}$  is the log dividend yield. The regressions are run for the full sample (1928-2006) and the postwar sample (1946-2006). For each regression coefficient, we provide the simulated  $p$ -values (see appendix). We boldface the simulated  $p$ -values that are smaller or equal to 10%.

<sup>17</sup>We take a number of steps to remove outliers. First, we treat the earnings data as missing if they are more negative than the market capitalization of stocks. Second, we winsorize  $RP$  (repurchases) and  $SI$  (share issuances) at 99.9%.

Table 5: : **Predictability by the Dividend Yield**

We regress cumulative log dividend growth ( $g$ ) or returns ( $r$ ), from one to five years, on the lagged log dividend yield. We provide the simulated  $p$ -value for each coefficients. The simulation considers the biases caused by the persistence of the variables, the contemporaneous correlation between the dependent and independent variables, and the overlapping small sample. We boldface the  $p$ -value if it is lower than or equal to 0.10. The sample is constructed using the merged dataset of CRSP and COMPUSTAT.

Horizon	1928-2006				1946-2006			
	$g$	$\bar{R}^2$	$r$	$\bar{R}^2$	$g$	$\bar{R}^2$	$r$	$\bar{R}^2$
1	-0.087 <b>[0.01]</b>	8.7	0.049 [0.28]	1.2	0.012 [0.67]	-0.9	0.103 [0.20]	6.9
2	-0.114 <b>[0.07]</b>	5.6	0.135 [0.22]	3.2	0.03 [0.79]	0.7	0.207 [0.17]	15.9
3	-0.121 [0.14]	3.5	0.199 [0.24]	5.3	0.033 [0.76]	-0.2	0.274 [0.20]	22.8
4	-0.097 [0.27]	1.1	0.277 [0.22]	9.9	0.042 [0.78]	-0.1	0.349 [0.20]	29.3
5	-0.067 [0.38]	-0.2	0.349 [0.21]	15.7	0.036 [0.74]	-0.6	0.455 [0.17]	36.2
Decomposition	64.04%		34.49%		-11.92%		103.84%	

For the full sample, the one-year coefficient on the lagged dividend yield is -0.087 with a  $p$ -value of 0.01 and an adjusted  $R^2$  of 9%. At the two-year horizon dividend growth is predictable but the adjusted  $R^2$  falls to 6%. At longer horizons the dividend yield coefficient becomes statistically insignificant. For the postwar period, the coefficient has the wrong (positive) sign from one to five-year horizons and is insignificant. In comparison, the estimated coefficients on the return predictability regression for the full sample and the post war sample are never significant at any horizon.

We also report the decomposition of the variance of the dividend yield using the one-year coefficients.<sup>18</sup> Based on the full sample estimates, about 64.04% (34.49%) of the dividend yield variance is due to dividend growth (returns). In stark contrast, based on the postwar sample, about -11.92% (103.84%) of the dividend yield variance is due to dividend growth (returns). The lack of dividend growth predictability, which is a postwar phenomenon, leads to the conclusion that almost all the dividend yield variation is driven by discount rates.

<sup>18</sup>If a vector of  $[\Delta d_t \ r_t \ dp_t]$  follows a first-order VAR, then equation (20) indicates that

$$\frac{b_r}{1 - \rho\phi} + \frac{-b_g}{1 - \rho\phi} \approx 1.$$

What this says is that 100% of the dividend yield variance can be approximately decomposed into the return component and cash flow component.

Table 6: : **Predictability by Net Payout Yield**

We regress cumulative log return, total payout (= dividend + repurchase) growth ( $\Delta d$ ), issuance growth ( $\Delta i$ ), or total net payout growth ( $\theta\Delta d - (\theta - 1)\Delta i$ ) from one to five years, on the lagged log net payout yield ( $v$ ). We provide the simulated  $p$ -value for each coefficients. The simulation considers the biases caused by the persistence of the variables, the contemporaneous correlation between the dependent and independent variables, and the overlapping small sample. We boldface the  $p$ -value if it is lower than or equal to 0.10. The estimated value for  $\theta$  is 1.6933.

Horizon	1928-2006				1946-2006			
	$r$	$\Delta d$	$\Delta i$	$\theta\Delta d - (\theta - 1)\Delta i$	$r$	$\Delta d$	$\Delta i$	$\theta\Delta d - (\theta - 1)\Delta i$
1	0.054 <b>[0.04]</b>	-0.044 <b>[0.03]</b>	0.150 [0.29]	-0.179 [0.13]	0.086 <b>[0.04]</b>	-0.015 [0.34]	0.272 <b>[0.06]</b>	-0.215 <b>[0.10]</b>
2	0.106 <b>[0.05]</b>	-0.032 [0.23]	0.322 [0.13]	-0.277 <b>[0.04]</b>	0.177 <b>[0.01]</b>	0.007 [0.64]	0.479 <b>[0.02]</b>	-0.320 <b>[0.05]</b>
3	0.133 <b>[0.08]</b>	-0.021 [0.36]	0.416 [0.14]	-0.323 <b>[0.07]</b>	0.219 <b>[0.02]</b>	0.006 [0.62]	0.529 <b>[0.04]</b>	-0.357 <b>[0.09]</b>
4	0.148 [0.11]	0.003 [0.50]	0.518 [0.13]	-0.354 <b>[0.09]</b>	0.231 <b>[0.04]</b>	0.021 [0.69]	0.543 <b>[0.07]</b>	-0.341 [0.17]
5	0.157 [0.15]	0.015 [0.54]	0.604 [0.13]	-0.393 <b>[0.10]</b>	0.301 <b>[0.02]</b>	0.004 [0.60]	0.668 <b>[0.05]</b>	-0.456 [0.11]
Decomposition	21.78%	30.19%	42.24%	72.43%	27.80%	8.40%	60.79%	69.19%

### 5.3 Predictability and yield decomposition: net payout yield

We next consider the case of the net payout yield. As shown in the appendix of Larrain and Yogo (2008), the Campbell-Shiller decomposition of the net payout yield,  $v_t$ , (see equation (1)) is

$$v_t = [\theta \times d_t - (\theta - 1) \times i_t] - p_t \quad (31)$$

$$= E_t \left[ \sum_{j=0}^{\infty} \rho^j r_{t+1+j} \right] - E_t \left[ \sum_{j=0}^{\infty} \rho^j [\theta \times \Delta d_{t+1+j} - (\theta - 1) \times \Delta i_{t+1+j}] \right], \quad (32)$$

where  $d_t$  is the logarithm of total payout,  $i_t$  is the logarithm of equity issuance, and thus  $[\theta \times d_t - (\theta - 1) \times i_t]$  is essentially the log net payout; the log-linearization parameter  $\theta$  is greater than one.<sup>19</sup>

Table 6 reports the predictability results. We note that the net payout yield predicts returns significantly in both the full sample and the postwar sample. The finding that the net payout yield can help predict returns is consistent with Boudoukh, Michaely, Richardson, and Roberts (2007), Larrain and Yogo (2008), and Pontiff and Woodgate (2008).

The net payout yield can predict total payout growth (dividends plus repurchases) only at the one-year horizon in the full sample with a coefficient of -0.044 (p-value 0.03); the postwar coefficients are all insignificant irrespective of horizon. Therefore, measuring cash flows by adding repurchases to dividends helps cash flow predictability only marginally.

The fifth and ninth columns of Table 6 report the results that measure cash flows as the growth

<sup>19</sup>The reason why total payout and equity issuance need to be log-linearized separately is that net payout (=total payout - equity issuance) can be negative.

in net payout. The net payout yield predicts net payout growth significantly for both the full sample and the postwar sample, and from short to long horizons.

Using one-year coefficients, for the full sample about 21.78% of the net payout yield variance is due to returns, 30.19% due to total payout, and 42.24% due to issuance; combined, 72.34% is due to net payout. For the postwar sample, 27.80% of net payout variance is due to returns and 69.19% is due to net payout. Therefore, in contrast to the case of dividend yield, discount rate news explains less than 50% of the net payout yield variance even in the postwar sample, suggesting that a large portion of price variation is due to cash flows. Importantly, despite dividend smoothing, the role of discount rate news remains very stable for both the full sample and the postwar sample. Equity issuance plays a crucial role in this stability, as shown in the coefficients.

#### 5.4 Predictability and yield decomposition: earnings yield

We can also understand price variation through the earnings yield. Denote dividends  $D_t$  and earnings  $E_t$ , then the payout ratio is  $DE_t = \frac{D_t}{E_t}$ , and dividend growth is

$$\begin{aligned}\Delta d_t &= \ln(E_t \times DE_t) - \ln(E_{t-1} \times DE_{t-1}) \\ &= \Delta e_t + \Delta de_t,\end{aligned}\tag{33}$$

where  $\Delta de_t$  is the growth rate of the payout ratio. Equation (1) can be rewritten as

$$\begin{aligned}e_t - p_t &= E_t \left[ \sum_{j=0}^{\infty} \rho^j r_{t+1+j} \right] - E_t \left[ de_t + \sum_{j=0}^{\infty} \rho^j (\Delta e_{t+1+j} + \Delta de_{t+1+j}) \right] \\ &= E_t \left[ \sum_{j=0}^{\infty} \rho^j r_{t+1+j} \right] - E_t \left[ \sum_{j=0}^{\infty} \rho^j (\Delta e_{t+1+j} + (1 - \rho)de_{t+1+j}) \right].\end{aligned}\tag{34}$$

We can use the earnings yield to predict returns, earnings growth, and payout ratio, and decompose the earnings yield accordingly.

Compared to dividends, predictability involving earnings requires additional care. In particular, when we use the log earnings yield ( $ep_{t-1}$ ) to predict return, we use the return from April of year  $t$  to April of year  $t + 1$ . This lag is to ensure that earnings become public information before we count future returns. When predicting log earnings growth rates ( $eg_t$ ), we use  $ep0_{t-1}$  which uses price at the beginning of year  $t - 1$ . When we predict returns and the payout rates we use price at the end of year  $t - 1$ . The use of  $ep0_{t-1}$  ensures that, regardless of the fiscal year end, the price we use is way ahead of earnings information.<sup>20</sup>

Table 7 reports the results. The earnings yield coefficients are always significant when predicting earnings growth. In particular, for the full sample, the  $ep0_{t-1}$  coefficient at the one-year horizon is -0.721 with a  $p$ -value of 0.00 and an adjusted  $R^2$  of 36%. At horizons greater than one year

<sup>20</sup>When the aggregate earning is negative, we set the earnings yield to be 0.0001, which translates to a log earnings yield of  $-9.21$ . Negative earnings occur only during 1933 following the great depression. Omitting this observation does not alter our results in any significant way.

Table 7: : **Predictability by Earnings Yield**

We regress cumulative log earnings growth ( $eg$ ), return ( $r$ ), and payout ratio ( $(1 - \rho)de$ ), from one to five years, on the lagged log earnings yield ( $ep_{t-1}$  or  $ep_{t-1}^0$ ). When predicting earnings growth, we use  $ep_{t-1}^0$ , in which the price is from the beginning (rather than the end) of the year. We provide the simulated  $p$ -value for each coefficient. The simulation considers the biases caused by the persistence of the variables, the contemporaneous correlation between the dependent and independent variables, and the overlapping small sample. We boldface the  $p$ -value if it is lower than or equal to 0.10. The sample is constructed using the merged dataset of CRSP, COMPUSTAT, and Moody's book equity.

Horizon	1928-2006						1946-2006					
	$eg$	$\bar{R}^2$	$r$	$\bar{R}^2$	$(1 - \rho)de$	$\bar{R}^2$	$eg$	$\bar{R}^2$	$r$	$\bar{R}^2$	$(1 - \rho)de$	$\bar{R}^2$
1	-0.721 <b>[0.00]</b>	36.2	0.037 <b>[0.10]</b>	2.6	-0.006 0.64	1.8	-0.651 <b>[0.02]</b>	32.1	0.027 [0.12]	1.9	-0.006 [0.45]	1.5
2	-0.855 <b>[0.00]</b>	40.2	0.035 [0.19]	0.4	-0.006 [0.69]	-0.2	-0.827 <b>[0.00]</b>	40.7	0.038 [0.13]	2.7	-0.004 [0.56]	-0.9
3	-0.839 <b>[0.00]</b>	38.2	0.032 [0.26]	-0.2	-0.007 [0.56]	-0.4	-0.849 <b>[0.02]</b>	42.7	0.030 [0.24]	-0.1	-0.003 [0.58]	-1.6
4	-0.898 <b>[0.00]</b>	40.4	0.085 <b>[0.08]</b>	5.6	-0.006 [0.55]	-0.8	-0.880 <b>[0.02]</b>	45.3	0.037 [0.23]	0.2	0.000 [0.64]	-1.8
5	-0.846 <b>[0.01]</b>	37.6	0.071 [0.15]	3.85	-0.008 [0.44]	-0.7	-0.850 <b>[0.04]</b>	43.7	0.022 [0.35]	-1.30	0.000 [0.59]	-1.8
Decomposition	95.41%		4.93%		0.89%		96.62%		4.07%		0.96%	

the estimated coefficients are around -0.85 and are always statistically significant. Remarkably, considering the earlier results regarding dividend growth predictability, the results are as strong in the postwar sample as in the full sample. In all cases the coefficients are statistically significant.

The predictive power of the earnings yield for returns is much weaker than that for earnings growth, consistent with Lamont (1998) and Goyal and Welch (2008). With the exception of the four-year horizon, the remaining estimates are insignificant.

For the full sample, 95.41% of the earnings yield variance is due to earnings growth, 4.93% due to discount rates, and 0.89% due to payout ratio. For the postwar sample, 96.62% of the earnings yield variance is due to earnings growth, 4.07% due to discount rates, and 0.96% due to payout ratio. Overall, the payout ratio plays a very minor role in driving earnings yield variance and hence we will ignore it for the remaining analysis in the paper.

In summary, to get around the issue of dividend smoothing, one can decompose either the net payout yield or the earnings yield. When doing so, the results are very stable and suggest a large role of cash flow news in stock price variation.

## 5.5 Portfolio tests

### 5.5.1 Smooth versus flexible dividend portfolios

We have the following additional testable hypotheses on dividend smoothing: for the postwar period, dividend growth should (not) be predictable for the least (most) smoothed firms. In contrast, earnings growth should be predictable for all firms regardless of how much dividends are smoothed.

Table 8: : **Smooth- versus Flexible-dividend Portfolios**

We sort firms into three portfolios according to the ratio of the standard deviation of dividend growth to the standard deviation of earnings growth. The firms with the lowest (highest) ratios consist the smooth-(flexible-) dividend portfolio. For the smooth and flexible portfolios respectively, we regress cumulative log dividend growth ( $dg$ ) or log earnings growth ( $eg$ ) and returns ( $r$ ) on the lagged log dividend yield ( $dp$ ). We provide the simulated  $p$ -value for each coefficient. The simulation considers the biases caused by the persistence of the variables, the contemporaneous correlation between the dependent and independent variables, and the overlapping small sample. We boldface the  $p$ -value if it is lower than or equal to 0.10. The sample is constructed using the merged dataset of CRSP, COMPUSTAT, and Moody's book equity. We provide dividend yield decomposition in Panel A and earnings yield decomposition in Panel B.

Smooth portfolios					Flexible portfolios			
1928-2006		1946-2006			1928-2006		1946-2006	
Panel A: Predictability by dividend yield								
Horizon	$g$	$r$	$g$	$r$	$g$	$r$	$g$	$r$
1	-0.142 [ <b>0.01</b> ]	0.153 [ <b>0.02</b> ]	0.014 [0.71]	0.211 [ <b>0.02</b> ]	-0.548 [ <b>0.00</b> ]	0.111 [ <b>0.05</b> ]	-0.432 [ <b>0.06</b> ]	0.182 [ <b>0.00</b> ]
3	-0.145 [0.14]	0.476 [ <b>0.00</b> ]	0.077 [0.86]	0.476 [ <b>0.02</b> ]	-0.436 [ <b>0.04</b> ]	0.319 [ <b>0.01</b> ]	-0.285 [0.31]	0.303 [ <b>0.00</b> ]
5	-0.022 [0.44]	0.683 [ <b>0.01</b> ]	0.109 [0.86]	0.715 [ <b>0.01</b> ]	-0.370 [0.17]	0.482 [ <b>0.00</b> ]	-0.475 [ <b>0.08</b> ]	0.450 [ <b>0.00</b> ]
Decomposition	49.31%	53.30%	-7.40%	110.02%	84.09%	16.98%	71.21%	30.00%
Panel B: Predictability by earnings yield								
Horizon	$eg$	$r$	$eg$	$r$	$eg$	$r$	$eg$	$r$
1	-0.779 [ <b>0.01</b> ]	0.055 [ <b>0.01</b> ]	-0.714 [ <b>0.01</b> ]	0.028 [ <b>0.08</b> ]	-0.769 [ <b>0.00</b> ]	0.043 [ <b>0.10</b> ]	-0.449 [ <b>0.05</b> ]	0.106 [ <b>0.00</b> ]
3	-0.613 [ <b>0.04</b> ]	0.049 [ <b>0.08</b> ]	-0.848 [ <b>0.01</b> ]	0.027 [0.23]	-0.844 [ <b>0.00</b> ]	0.081 [0.11]	-0.570 [ <b>0.06</b> ]	0.152 [ <b>0.02</b> ]
5	-0.856 [ <b>0.01</b> ]	0.086 [ <b>0.03</b> ]	-0.851 [ <b>0.03</b> ]	0.017 [0.37]	-0.737 [ <b>0.00</b> ]	0.151 [ <b>0.04</b> ]	-0.518 [0.17]	0.254 [ <b>0.00</b> ]
Decomposition	96.18%	6.81%	98.12%	3.97%	95.75%	5.43%	75.45%	17.94%

To test these hypotheses, we sort firms into three portfolios according to the smoothness parameter  $S$  ( $= \frac{\sigma(\Delta d)}{\sigma(\Delta e)}$ ). Panel A of Table 8 reports the results using the dividend yield. For the smoothed portfolio in the postwar period, the dividend growth coefficients all have the wrong sign and are statistically insignificant. In contrast, for the flexible dividend portfolio, dividend growth is predictable in both the full and the postwar samples, and for different horizons. If we take the flexible dividend portfolio as indicative of the case with little dividend smoothing, then for the postwar period 71.21% of the dividend yield variance is due to cash flows (dividend growth) and only 30.00% is due to discount rates.

Interestingly, for both portfolios, return coefficients become significant for both the full and postwar samples. This is in contrast to Table 3 and Table 5 where returns are never significant. Therefore, grouping firms by dividend smoothing not only helps interpreting dividend growth predictability, but also helps in recovering return predictability.

Panel B of Table 8 reports the results using earnings yield. Remarkably, for both the smooth

and flexible dividend portfolios, across short and long horizons, and for both the full and postwar samples, the earnings growth is highly predictable. Returns are also predictable with the earnings yield for both smooth and flexible portfolios and at all horizons except long horizons for the smooth portfolio in the post war period. Examining the final row of Panel B reveals that only a small portion of earnings yield variance is due to discount rates even in the postwar period.

## **5.6 Further robustness checks**

We conduct the following further robustness checks without reporting the results in tables.

### **5.6.1 Smooth versus flexible earnings portfolios**

If dividend smoothing kills dividend growth predictability, then earnings smoothing should kill earnings growth predictability. Therefore, one way to validate our argument is to examine portfolios according to earnings smoothing. The smoothness measure is computed as the ratio between the standard deviation of the firm's earnings (scaled by total asset) and the standard deviation of the firm's operating cash flow (scaled by total asset). We sort firms into three portfolios according to the earnings smoothness measure; the firms with the lowest (highest) ratios comprise the smooth (flexible) earnings portfolio. The sample is constructed using the merged dataset of CRSP and COMPUSTAT. The sample period starts in 1951 as COMPUSTAT data are required to compute the earnings smoothness measure. We find that there is no evidence that the smooth earnings portfolio is predictable with the earnings yield. In contrast, the earnings growth of the portfolio of firms with flexible earnings is highly predictable with the earnings yield. For both portfolios, returns are predictable with the earnings yield. The evidence is consistent our claim that cash flow smoothing, whether it is dividends or earnings, can kill cash flow predictability.

### **5.6.2 Stable versus volatile earnings growth portfolios**

The volatility of both dividend and earnings growth has decreased in the postwar period. Chen (2009) finds that dividend volatility per se does not explain the lack of dividend growth predictability. To further verify this point, we sort firms into three portfolios according to the standard deviation of annual earnings growth. We find that separating firms by earnings volatility leads to the same conclusions as in the case of the aggregate portfolio: dividend growth is not predictable in the postwar period, but earnings growth is predictable. The evidence further strengthens our hypothesis that it is cash flow smoothing, rather than volatility per se, that contributes to the lack of predictability.

### **5.6.3 Payout “dinosaurs” versus “non-dinosaurs”**

DeAngelo, DeAngelo, and Skinner (2004) and Skinner (2008) show that earnings and dividends have become increasingly concentrated in a small set of firms. Following Skinner (2008), we form a portfolio based on a small group of firms that consistently make both dividend payments and repurchases and call them payout “dinosaurs.” These are firms making dividend payments for more

than 15 years and making repurchases for more than 10 years. The portfolio using the rest of the firms are payout “non-dinosaurs.” For both the “dinosaurs” and “non-dinosaurs,” we uncover the same patterns as in the case of the aggregate portfolio.

An intriguing finding is that returns are strongly predictable by dividend yield in the postwar period for both “dinosaurs” and “non-dinosaurs.” In fact, returns are strongly predictable by dividend yield for all portfolios regardless of whether we separate firms by “dinosaurs,” by dividend smoothing, by earnings smoothing, or even by earnings volatility. Such evidence contrasts with the standard finding that return predictability by the dividend yield lacks statistical power. It is clear that cash flow patterns and payout behavior affect return predictability. We leave a thorough study on this issue for future work.

#### **5.6.4 Other robustness tests**

The earnings data in most tests are calculated using the clean surplus formula. This approach helps to increase our sample length and allows more firms and thus represents the market better. For robustness, we construct the following alternative: starting from 1950 (the starting year of COMPUSTAT data) we only include those firm years with earnings data available from COMPUSTAT; before 1950 we still use the clean surplus formula to calculate earnings. We find that the main conclusions remain.

In addition, since we have used the S&P index portfolio earlier to establish the results regarding dividend policy, it is useful to also examine the predictability using S&P index firms. We thus construct a market portfolio as earlier but with only CRSP firms belonging to the S&P index. We find that our conclusions are robust to the case of S&P index firms.

## **6 Conclusion**

A central issue for financial economists is to understand stock price variations. The answer to this question is usually obtained by comparing the relative predictability of cash flows and returns by the dividend yield. In this regard, the usual finding is that, at the aggregate level, returns are predictable by the dividend yield but dividend growth is not. This leads to the somewhat uncomfortable conclusion that there is little cash flow news in stock price variations.

Chen (2009) shows that dividend growth is strongly predictable in the prewar period, but this predictability completely disappears in the postwar period. It is difficult to imagine that financial markets have evolved in such a way that a lot of cash flow news is incorporated in price variations in the prewar period but little is incorporated in the postwar period. Rather, it is natural to suspect that the dramatic change of cash flow predictability has more to do with the cash flow measures than with the way investors evaluate securities.

To verify this conjecture, we first document a significant change of dividend policy at the aggregate level from the prewar to the postwar period. In the postwar period, dividends are much more smoothed and respond much more to their past levels rather than to the outlook of future cash

flows.

Our simulated results provide two conclusions regarding dividend smoothing. First, even if dividends are supposed to be strongly predictable without smoothing, dividend smoothing can bury this predictability in a finite sample. Second, dividend smoothing leads to a persistent dividend yield, a phenomenon that can be verified in the data.

The finding that dividends are dramatically more smoothed in the postwar period, combined with the finding from the simulations that dividend smoothing can kill predictability, provides a reasonable interpretation on why dividend growth is predictable in the prewar period but not so in the postwar period.

We proceed to show how one can interpret price variation by using measures that are less subject to dividend smoothing: net payout and earnings. In both cases, we find remarkably consistent results for both the full and the postwar sample: the majority of the variation of the net payout (earnings) yield comes from net payout (earnings) growth, suggesting a role of cash flow news much larger than discount rate news.

We further sort firms according to the degree of dividend smoothness. For the most smoothed portfolio, dividend growth is not predictable in the postwar period; for the least smoothed portfolio, dividend growth is predictable. In contrast, for both portfolios, earnings growth is predictable in the full sample as well as the postwar sample. Therefore, the lack of cash flow predictability has more to do with dividend smoothness than with cash flow per se.

## References

- Ang, A., 2002, Characterizing the ability of dividend yields to predict future dividends in log-linear present value models, working paper, Columbia University.
- Ang, A. and G. Bekaert, 2007, Stock return predictability: Is it there? *Review of Financial Studies*, 20 (3): 651-707.
- Ang, A. and J. Liu, 2004, How to discount cashflows with time-varying expected returns, *Journal of Finance*, 59 (6): 2745-2783.
- Bansal, R., R. Dittmar, and C. Lundblad, 2005, Consumption, dividends, and the cross section of equity returns, *Journal of Finance*, 60: 1639-1672.
- Bansal, R., V. Khatchatrian, and A. Yaron, 2005, Interpretable asset markets?, *European Economic Review*, 49 (3): 531-560.
- Bansal, R. and A. Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance*, 59: 1481-1509.
- Bansal, R. and A. Yaron, 2007, The asset pricing-macro nexus and return-cash flow predictability, working paper, Duke University.
- Binsbergen, J.H. and R. Koijen, 2009, Predictive regressions: A present-value approach, *Journal of Finance*, 65: 1439-1471.
- Boudoukh, J., M. Richardson, and R.F. Whitelaw, 2008, The myth of long-horizon predictability, *Review of Financial Studies*, 21: 1577-1605.
- Boudoukh, J., R. Michaely, M. Richardson, and M. Roberts, 2007, On the importance of measuring payout yield: Implications for empirical asset pricing,” *Journal of Finance*, 62: 877-915.
- Brav, A., J. Graham, C. Harvey, and R. Michaely, 2005, Payout policy in the 21st century, *Journal of Financial Economics*, 77 (3): 483-527.
- Campbell, J. and J. Ammer, 1993, What moves the stock and bond market? A variance decomposition of long term asset returns, *Journal of Finance*, 48: 3-37.
- Campbell, J. and J. Cochrane, 1999, By force of habit: A consumption-based explanation of stock market behavior, *Journal of Political Economy*, 107: 205-251.
- Campbell, J. and R. Shiller, 1988, The dividend yield and expectations of future dividends and discount factors, *Review of Financial Studies*, 1: 195-228.
- Campbell, J. and R. Shiller, 1998, Valuation ratios and the long-run stock market outlook: An update, working paper, Harvard University.

- Chen, L. 2009, On the reversal of return and dividend predictability: A tale of two periods, *Journal of Financial Economics*, 92: 128-151.
- Chen, L. and X. Zhao, 2009, Return decomposition, *Review of Financial Studies*, 22(12): 5213-5249.
- Cochrane, J. H., 1992, Explaining the variance of price-dividend ratio, *Review of Financial Studies*, 5: 243-280.
- Cochrane, J. H., 2001, Asset pricing, Princeton University Press, Princeton, New Jersey.
- Cochrane, J. H., 2008, The dog that did not bark: A defense of return predictability, *Review of Financial Studies*, 21(4): 1533-1575.
- Cohen, R. B., C. Polk, and T. Vuolteenaho, 2003, The value spread, *Journal of Finance*, 58: 609-641.
- Davis, J. L., E. F. Fama, and K. R. French, 2000, Characteristics, covariances, and average returns: 1929 to 1997, *Journal of Finance*, 40: 389-406.
- DeAngelo, H., L. DeAngelo, and D. J. Skinner, 2004, Are dividends disappearing? Dividend concentration and the consolidation of earnings, *Journal of Financial Economics*, 72: 425-456.
- Engsted, T. and T. Q. Pedersen, 2009, The dividend-price ratio does predict dividend growth: International evidence, working paper, University of Aarhus.
- Fama, E. F. and K. R. French, 1988, Dividend yields and expected stock returns, *Journal of Financial Economics*, 22: 3-25.
- Goyal, A. and I. Welch, 2003, Predicting the equity premium with dividend ratios, *Management Science*, 49: 639-654.
- Goyal, A. and I. Welch, 2008, A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies*, 21: 1455-1508.
- Graham, B. and D. Dodd, 2008. Security analysis, sixth edition, McGraw-Hill.
- Hansen, L. P., J. C. Heaton, and N. Li, 2008, Consumption strikes back? Measuring long-run risk, *Journal of Political Economy*, 116 (2): 260-302.
- Kendall, M., 1954, Note on bias in the estimation of autocorrelation, *Biometrics*, 41: 403-404.
- Lacerda, F. and P. Santa-Clara, 2010, Forecasting dividend growth to better predict returns, *Working Paper*, University of Chicago.
- Lamont, O., 1998, Earnings and expected returns, *Journal of Finance*, 53 (5): 1563-1587.
- Larrain, B. and M. Yogo, 2008, Does firm value move too much to be justified by subsequent changes in cash flow? *Journal of Financial Economics*, 87(1): 200-226.

- Leary, M. and R. Michaely, 2010, Determinants of dividend smoothing: empirical evidence, working paper, Washington University in St. Louis and Cornell University.
- Lettau, M. and S. C. Ludvigson, 2005, Expected returns and expected dividend growth, *Journal of Financial Economics*, 76: 583-626.
- Lettau, M. and S. V. Nieuwerburgh, 2008, Reconciling the return predictability evidence, *Review of Financial Studies*, 21(4): 1607-1652.
- Lintner, J., 1956, Distribution of incomes of corporations among dividends, retained earnings and taxes, *American Economic Review*, 46: 97-113.
- Mankiw, N. G. and J. A. Miron, 1986, The changing behavior of the term structure of interest rates, *Quarterly Journal of Economics*, 101: 211-228.
- Marsh, T. A., and R. Merton, 1987, Dividend behavior for the aggregate stock market, *Journal of Business*, 60: 1-40.
- Miller, M. H. and Modigliani, F., 1961, Dividend policy, growth, and the valuation of shares, *Journal of Business*, 34: 411-433.
- Newey, W. and K. West, 1987, A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica*, 55: 703-708.
- Pastor, L. and R. F. Stambaugh, 2009, Predictive systems: Living with imperfect Predictors, forthcoming, *Journal of Finance*.
- Pontiff, J. and A. Woodgate, 2008, Share issuance and cross-sectional returns, *Journal of Finance*, 63: 921-945.
- Rangvid, J., M. Schmeling, and A. Schrimpf, 2009, Cash flow-predictability: Still going strong, *Working Paper*, Copenhagen Business School.
- Robertson, D. and S. Wright, 2006, Dividends, total cash flow to shareholders, and predictive return regressions, *The Review of Economics and Statistics*, 88(1): 91-99.
- Sadka, G., 2007, Understanding stock-price volatility: The role of earnings, *Journal of Accounting Research*, 48: 199-228.
- Schwert, G. W. 1990, Indexes of U.S. stock prices from 1802 to 1987, *Journal of Business*, 63: 399-426.
- Skinner, D., 2008, The evolving relation between earnings, dividends, and stock repurchases, *Journal of Financial Economics*, 87: 582-609. 375-421.
- Stambaugh, R., 1986, Bias in regression with lagged stochastic regressors, working paper, University of Chicago.

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

Vuolteenaho, T., 2000, Understanding the aggregate book-to-market ratio and its implication to current equity-premium expectations, working paper, Harvard University.

## Appendix

The power of predictability tests is frequently questioned because of the persistence of the independent variable and its contemporaneous correlation with the dependent variables (e.g., Kendall (1954), Stambaugh (1986, 1999), and Pastor and Stambaugh (2009)), and the overlapping nature of the dependent variable when conducting long-horizon tests (e.g., Boudoukh, Richardson, and Whitelaw (2006)), compounded with small sample size. We describe below the procedure through which we simulate  $p$ -value for each predictive coefficient to take care of the above problems.

Suppose we will run the following predictive regressions:

$$y_t^i = \xi_i + \alpha_i \times x_{t-1} + \varepsilon_{it}, \quad (A1)$$

where  $y_t^i$ ,  $i = 1, 2, \dots, 5$ , is the cumulative summation of  $y_t$  from 1 to horizon  $i$ . Also suppose  $y_t^1$  and  $x_t$  follow  $AR(1)$  processes:

$$y_t^1 = \beta_0 + \beta_1 \times y_{t-1}^1 + \omega_t, \quad A2 \quad (35)$$

$$x_t = \gamma_0 + \gamma_1 \times x_{t-1} + v_t, \quad A3 \quad (36)$$

and the correlation  $\text{corr}(\omega_t, v_t) = \rho$ . In addition, the sample size is  $T$ .

To simulate the  $p$ -value for the predictive coefficient  $\alpha_1$ , we first conduct OLS regressions for equations A1-A3 and obtain estimates for the coefficients and the residuals. We then jointly simulate time series for  $y_t^1$  and  $x_t$  with size  $T$ . To preserve the distribution properties of the historical data, we draw from the residuals of the historical data when conducting the simulations. The null is that  $y_t^1$  is not predictable by  $x_{t-1}$ . Long-horizon simulates of  $y_t^i$  are subsequently constructed by summing the simulated  $y_t^1$ . We regress the simulated  $y_t^i$  on the simulated  $x_{t-1}$ , obtaining the simulated  $\alpha_i$ , which we call  $\alpha_{sim,i}$ . We repeat the exercise 10,000 times to obtain the time series of  $\alpha_{sim,i}$ . We finally compare the estimated  $\alpha_i$  with the time series of  $\alpha_{sim,i}$  to obtain the  $p$ -value for the estimated  $\alpha_i$ .

The above simulations take into consideration the autocorrelation of the variables, the contemporaneous correlation between the variables, the small sample size, and the overlapping data construction. We report the simulated  $p$ -values in the paper.