Dividend Smoothing and Predictability

Long Chen∗
Olin Business School
Washington University in St. Louis

Zhi Da†
Mendoza College of Business
University of Notre Dame

Richard Priestley‡
Norwegian School of Management

November 2010§

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 in a finite sample. We further show that aggregate dividends are dramatically more smoothed in the postwar period than before. Therefore, the lack of dividend growth predictability in the postwar period, as widely documented in the literature, 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. Our take-away messages are that (i) dividend smoothing can severely affect dividend predictability in a finite sample, (ii) there is significant cash flow news in stock price variations, and (iii) when smoothed, dividends do not represent well the outlook of future cash flows.

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 School of Management, 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 Præsen, Paul Schultz, Sophes Shive, Charles Trichina, and seminar participants at 2009 Finance Research Association annual meeting, Hong Kong University, Hong Kong University of Science and Technology, Indiana University, Shanghai Advanced Institute of Finance, 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: while corporate managers have large discretion over payout options, such discretion should be irrelevant for stock prices. Rather, 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.¹

Although dividends might not be relevant for stock prices, they are critical for economic analysis. Understanding stock price variation has concerned economists for many years and has focused on determining whether investors’ revised forecasts regarding future cash flows or discount rates are the drivers of price variation. Empirically, the answer to this question is usually inferred from the predictability of cash flows relative to that of stock returns.² Since dividends are habitually seen as the cash flows to stockholders, “predictability of dividends and/or returns form, in many ways, the rational paradigm to interpret asset price variation.” (Bansal and Yaron (2007)).³

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 the period 1872-1945 but, consistent with the extant literature, 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.

¹While payout policy might be relevant to shareholders when the capital market is incomplete or imperfect (see the excellent survey by Allen and Michaely (2003) and the references therein), the intuition remains that the fundamental driver of a firm’s cash flow, namely, its earnings power, is of the first order importance for the firm’s valuation.

²The 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.

³For 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.
What has caused such a dramatic change of predictability? Dividend policy, being irrelevant, should not affect stock prices, but it affects dividend growth by definition (e.g., Marsh and Merton (1986)). Therefore, 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 the extent to which firms smooth dividends changes in the postwar period relative to the prewar period. This is done by fitting various dividend behavior models based on Lintner (1956) and Marsh and Merton (1987). The evidence is compelling: dividend payout at the aggregate level has become much more smoothed. For example, when running the standard Lintner’s (1956) model for the prewar period (1871-1945), the speed of adjustment to target is 0.37; the corresponding number for the postwar period (1946-2006) is 0.09. In other words, in the postwar period dividends adjust to their earnings target at a speed about one fourth of that in the prewar period. As another example, if we regress dividend change on its own lag, the coefficient is 0.061 for the prewar period and is statistically insignificant. The corresponding coefficient is 0.687 in the postwar period and statistically significant at the 1% level. 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 study its impact on predictive regressions. Throughout these simulations, we adopt a dividend policy such that it is sustainable and the dividend yield is always within a sensible range. Our simulated results show that, even though dividends are generated to be strongly predictable without smoothing, introducing dividend smoothing can eliminate this predictability at normal horizons through which we run regressions in a finite sample. Severe dividend smoothing also makes the dividend yield very persistent, a pattern evident in the data: the autoregressive coefficient for the log dividend yield is 0.557 in the prewar period, and 0.956 in the postwar period.
Despite a voluminous literature that relies on the relative extent of return and dividend growth predictability to interpret price variation, to the best of our knowledge, whether dividend smoothing plays a critical role in this interpretation has never been formally explored. 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.\(^4\) 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 portfolio that is least smoothed, but not so for the portfolio that is most smoothed. 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% [\(^4\)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.”]
(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 further outline the role of cash flow smoothing, we show that when earnings are smoothed, they also lose their predictability by the earnings yield. However, this lack of predictability is not driven primarily by the observed reduction in earnings volatility in the postwar sample. Additional results also show that the lack of dividend growth predictability is not restricted to major dividend payers (see DeAngelo, DeAngelo, and Skinner (2004), Skinner (2008), and Leary and Michaely (2010)).

Overall, once the impact of dividend smoothing is controlled for, regardless of whether we examine dividends, net payout, or earnings, we reach the consistent conclusion that cash flow news plays a significant role in driving price variation in the postwar period.

**Link to the literature** The relative dividend/return predictability forms the rational paradigm to interpret stock price variation (Bansal and Yaron (2007)). There is a large asset pricing literature on this issue (e.g., 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 (2009), Chen (2009), and Chen and Zhao (2009)).

In reality dividends are determined by corporate policies that can be arbitrary. 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. Fama and French (1988) note that dividends are more smoothed in the post war period. However, they do not investigate the impact this has on dividend predictability or return predictability.

---

5Chen (2009) also asks whether dividend smoothing has contributed to the lack of dividend predictability in the 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. We do both, and, in addition, we find that the net payout (earnings) yield can predict net payout (earnings) growth. Our combined evidence thus leads to a different conclusion.
Our message is that dividend growth might not be predictable by the dividend yield simply because dividends are smoothed; we provide several alternative methods, all of which highlight the important role of cash flow news in price variations. Notably, Lettau and Ludvigson (2005) point out another reason, different from dividend smoothing, of why dividend growth might not be predicted by dividend yield. Similarly, 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 (2009) show that dividend growth is predictable based on past values of dividend growth, but they do not find significant predictability using the dividend yield.

A separate literature tries to incorporate more forms of payout in addition to dividends when running predictive regressions of returns (e.g., Vuolteenaho (2000), Bansal, Khatchatrian, and Yaron (2005), 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)). Our paper is in spirit close to this literature. But our focus is to study the impact of dividend smoothing on predictability and use this to explain the dramatic reversal of dividend growth predictability. Our findings help connect this literature to the literature on dividend growth predictability.

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 Schimpf (2010) show that dividend growth is predictable in countries where dividends smoothing is much less pronounced.

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

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

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 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 interest is to understand why stock prices vary. Intuitively, stock price variation could reflect either a revised outlook for future cash flows or revisions to discount rates. The predictive regressions reveal which component is revised in prices.

This predictive regression approach is potentially problematic. On one hand, dividend policy is irrelevant in valuation (Miller and Modigliani (1961) and Marsh and Merton (1986)); on the other hand, dividend policy determines the path of dividend growth in a finite sample. Consequently, there is the potential for a large wedge between price variation and future dividend growth variation. This creates a problem since 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 the language of Miller and Modigliani (1961), the nature of the problem is that stock returns have little to do with how cash flows are distributed, but dividend growth has everything to do with it.

The counterargument is that, since dividend growth does seem to keep pace with price in the long run, dividend policy cannot be too wild. The difficulty comes down to identifying how long is the “long run.” For example, dividend growth is only unpredictable by the dividend yield in the
postwar period. If the postwar period has experienced a large change in dividend policy, how much has this contributed to the lack of predictability?

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

\[
\Delta d_{t+1} = \alpha_0 + \alpha_1 e_{t+1} + \alpha_2 d_t + u_{t+1},
\]

(2)

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

\[
\Delta d_{t+1} - \Delta d_t = \alpha_1 \Delta e_{t+1} + \alpha_2 \Delta d_t + \Delta u_{t+1},
\]

(3)

or

\[
\Delta d_{t+1} = \alpha_1 \Delta e_{t+1} + (1 + \alpha_2) \Delta d_t + \Delta u_{t+1}.
\]

(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

\[
\sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} = \alpha_0 + \alpha_1 \Delta e_{t+1} + (1 + \alpha_2) \Delta d_t + u_{t+1} + \rho \alpha_0 + \rho \alpha_1 \Delta e_{t+2} + \rho (1 + \alpha_2) \Delta d_{t+1}
\]

\[+ \rho u_{t+2} + \rho^2 \alpha_0 + \rho^2 \alpha_1 \Delta e_{t+3} + \rho^2 (1 + \alpha_2) \Delta d_{t+2} + \rho^2 u_{t+3} + ... \]

(5)

\[
= \text{constant} + \frac{(1 + \alpha_2)}{1 - (1 + \alpha_2) \rho} \Delta d_t + \frac{\alpha_1}{1 - (1 + \alpha_2) \rho} \sum_{j=0}^{\infty} \rho^j \Delta e_{t+1+j}
\]

\[+ \frac{1}{1 - (1 + \alpha_2) \rho} \sum_{j=0}^{\infty} \rho^j u_{t+1+j}. \]

(6)

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

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

(7)

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

(8)

\[
= \text{Discount rate component} - \text{Cash flow component}
\]

(9)

\[
= \text{Discount rate component} - [\text{Smoothing component} + \text{Earnings component}],
\]

(10)
where

\[
\text{Discount rate component} = E_t \left[ \sum_{j=0}^{\infty} \rho^j r_{t+1+j} \right], \quad (11)
\]

\[
\text{Cash flow component} = \text{Smoothing component} + \text{Earnings component}, \quad (12)
\]

\[
\text{Smoothing component} = \frac{(1 + \alpha_2)}{1 - (1 + \alpha_2) \rho} \Delta d_t, \quad (13)
\]

\[
\text{Earnings component} = \frac{\alpha_1}{1 - (1 + \alpha_2) \rho} E_t \left( \sum_{j=0}^{\infty} \rho^j \Delta e_{t+1+j} \right). \quad (14)
\]

The intuition is as follows. The variation of the dividend yield must reflect the variation of either
the discount rate component or the cash flow component. Within the cash flow component, 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.

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 the remainder of the paper: (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?

In this section, we examine whether there is a significant change of dividend policy from the prewar
to the postwar period.
### 3.1 Dividend policy models

Based on interviews with managers, Lintner (1956) proposes the following partial-adjustment model of dividend-setting behavior:

\[
\Delta D_t = \alpha_0 + \alpha_1 E_t + \alpha_2 D_{t-1} + u_t \tag{15}
\]

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 (15) is the first dividend policy model we will estimate.

If we take the first difference of equation (15), we obtain the second testable model:\(^6\)

\[
\Delta D_t = \beta_0 + \beta_1 \times \Delta E_t + \beta_2 \times \Delta D_{t-1} + \varepsilon_t. \tag{16}
\]

The advantage of equation (16) 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

\[
\Delta D_t = \gamma_0 + \gamma_1 E_t + \gamma_2 \times \Delta D_{t-1} + \upsilon_t. \tag{17}
\]

Equation (17) is the same as equation (15) 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 dividend payout depends on its own lag, and thus 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

---

\(^6\)For equation (15) to be fully consistent with equation (16), \(\beta_0\) 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.
written as
\[
\ln \left( \frac{D_{t+1}}{D_t} \right) + \frac{D_t}{P_{t-1}} = \lambda_0 + \lambda_1 \times \ln \left( \frac{P_t + D_t}{P_{t-1}} \right) + \lambda_2 \times \ln \left( \frac{D_t}{P_{t-1}} \right) + \varpi_{t+1},
\]
where \( \lambda_1 \) captures how much dividends respond to permanent earnings changes: 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.

### 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 mainly comes from Cowles (1939), which presumably covers all stocks traded on NYSE during the period; the 1926-2006 sample includes the S&P index firms. The Cowles data have been used by many studies (e.g., Campbell and Shiller (1988, 1998), Wilson and Jones (1987), Schwert (1989, 1990), Goetzmann (1995), Goetzmann and Jorion (1993), Lundblad (2007), and Chen (2009)).

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.

The reduction of dividend growth volatility is consistent with, but not necessarily a result of, dividend smoothing; it could 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. If dividends are more smoothed in the postwar period, the reduction of dividend volatility should be larger than the reduction of earnings volatility. To this end, we define the smoothness parameter as
\[
S = \frac{\sigma(\Delta d)}{\sigma(\Delta e)},
\]
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. The fact that the smoothness parameter has been cut by about half suggests that dividends are indeed much more smoothed in the postwar period. Another piece of supporting evidence is that, for the prewar period, the AR(1) coefficient for the dividend yield is 0.518; the corresponding number for the postwar period is 0.926. Interestingly, the AR(1) coefficient for the earnings yield is 0.621 in the prewar period and 0.832 in the postwar period. Therefore, dividend growth is less (more) persistent than earnings growth in the prewar (postwar) period, consistent with the dividend smoothing argument.

Panel B reports similar statistics for total payout yield \(=\text{(dividend+repurchase)/price}\) and net payout yield \(=\text{(dividend+repurchase-equity issuance)/price}\).\(^7\) Although these payout yields appear more persistent in the postwar period, the change from the full sample to the postwar sample is minor; we find a similar pattern for the smoothness parameter using total payout. 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. 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. Note that the close link between earnings growth and dividend growth in the early years is not restricted to the Great Depression period.

We estimate the four dividend behavior models and report the results in Table 2. 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, statistically significant at the 1% level in both cases; thus the postwar \(SA\) is only about one fourth of the prewar \(SA\). The final column of Panel A reports a Chow test and 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 which cases the null is clearly rejected. Similarly, in Panel B of Table 2

\(^7\)See Section 5.1 for data construction.
where we use first differences of the independent variables, the SA coefficient for the postwar period
is also about one fourth of that for the prewar period and the Chow test clearly rejects the null of
constant parameters in favor of a structural break. The reported F-tests also indicate that estimated
coefficients in the two samples are not equal to one another.

In Panel C of Table 2, we report estimates for the third model. The coefficient on the lagged
change in dividends is 0.061 for the prewar period, statistically insignificant from zero. In stark
contrast, the coefficient is 0.687 for the postwar period and statistically significant at the 1% level.
Therefore, dividend policy has evolved from little dependence on the lagged dividends in the prewar
period to heavy dependence in the postwar period. The lagged dividend payout has become the
dominant variable explaining current dividend payout. 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 coefficient that measures the response to permanent earnings change is 0.003,
statistically indifferent from zero, and the implied convergence coefficient is -0.061 indicating that
there is no convergence to the target. 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.8

To understand better the evolution of dividend policy, 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

---

8We have also tested the four models in log form and find very similar results. For brevity we do not report them.
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.

Figure 3 separately 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 where it subsequently remains. Combined, Figures 2 and 3 indicate that the drastically stronger pattern of dividend smoothing in the postwar period is not driven by extreme numbers in some particular years. Rather, it represents a genuine change of aggregate dividend behavior from the prewar to the postwar period.\(^9\)

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)).\(^10\) 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).\(^11\)

---

\(^9\) 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. Therefore, the increase of dividend smoothing appears to be a genuine pattern at the aggregate level that is not driven by certain firms.

\(^10\) 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. One can also argue that education from the academic world (e.g., Miller and Modigliani (1961)) has helped liberate managers’ decisions on dividend payout.

\(^11\) Consistent with this story, in untabulated results, we find that net payout, in contrast to dividends, is not more
Regardless of the interpretation, aggregate dividends are much more smoothed in the postwar period than earlier. This finding raises the possibility that the dividend smoothing in the postwar period might be the main reason why dividend growth is not predictable by the dividend yield. To thoroughly understand this issue we need to know how different degrees of dividend smoothing affect predictability. This question is relevant because, even though firms can follow arbitrary dividend policies, such policies might not be sustainable in the long run. Thus, it is not clear how sustainable dividend smoothing affects predictability in a finite sample. We explore this issue in the next section.

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 \( r_t \),

\[
\begin{align*}
dp_{t+1} &= a dp + \phi \times dp_t + \varepsilon_{t+1}^{dp} \\
g_{t+1} &= a_g + b_g \times dp_t + \varepsilon_{t+1}^{g}
\end{align*}
\]

(20)

\[
\begin{align*}
r_{t+1} &= a + b_r \times dp_t + \varepsilon_{t+1}^{r}.
\end{align*}
\]

(21)

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

\[
b_r \approx 1 - \rho \phi + b_g,
\]

(23)

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.

smoothed in the postwar period. The easiest way to see this is to consider the regression of the net payout growth on its own lag. The estimate for the whole sample where net payout is available (1928-2006) is 0.084. The corresponding estimate in the post war period is 0.085. This contrasts sharply with the increase in dependence of dividend growth on its own lag in the post war period.
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 Whitelaw (2006)). Second, equation (23) 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 (23), 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.

In sum, dividend smoothing is likely to make dividend growth less predictable, the dividend yield more persistent, and its impact on return predictability is not clear.

It might appear that we already know how dividend smoothing affects predictability. However, this is not obvious because even though a firm can adopt an extreme dividend policy in the short run, such a policy might not be sustainable in the long run. In other words, the impact of dividend smoothing is likely to be mitigated in a long-enough sample. Therefore, it is not clear 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, -0.0596, and -0.406 respectively, and are highly statistically significant. In comparison, the one-year return coefficients 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.\textsuperscript{13}

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)).

Therefore, summarizing the results from the historical data, we find that dividend growth is strongly predictable in the prewar period but is not predictable in the postwar period. Stock returns are less predictable than dividends in the prewar period and the predictability lacks statistical power in the postwar period. The empirical evidence documented above will serve as the benchmark case when we analyze the impact of dividend smoothing below.

Another important piece of evidence is that the dividend yield is much more persistent in the postwar period than in the prewar period. Regressing the log dividend yield on its own lag yields a coefficient of 0.557 for the prewar period and 0.956 for postwar. This finding is expected given our earlier finding that dividends are much more smoothed in the postwar period.

\textbf{4.1 Three cases of dividend smoothing}

We proceed to conduct simulations to study how dividend smoothing affects predictability. We examine three cases in sequence. In all cases, we ensure that the dividend policy is sustainable.

\textsuperscript{13}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.
4.1.1 First case

In this case, we first fit the Marsh and Merton (1987) dividend smoothing model (equation (18)) for the prewar and postwar periods separately, as shown in Panel D of Table 2. We then simulate dividend growth using the fitted equation (18). We also simulate returns under the null that returns are not predictable,

\[ r_{t+1} = a_r + \varepsilon_{t+1}^r, \]

(24)

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.

In summary, we have constructed a scenario in which the dividend policy is sustainable and the dividend yield is stationary in the long run. In addition, stock returns are not predictable. All the predictability is on the dividend side: Equation (18) implies dividend growth predictability when dividends are not too smoothed, and the correction of the dividend yield, when needed, is through the adjustment of dividends.

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, -0.679, and -0.798 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 in the small samples employed in the literature.

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. The null is that dividend growth is predictable without smoothing but return is not:

\[ g_{t+1} = a_g - 0.1 \times dp_t + \varepsilon^g_{t+1} \]  
\[ r_{t+1} = a_r + \varepsilon^r_{t+1}, \]  

(25)  
(26)

where \( g_{t+1} \) is dividend growth rate. The residuals \( \varepsilon^g_{t+1} \) and \( \varepsilon^r_{t+1} \) are chosen such that the historical variance-covariance matrix of dividend growth and returns in the prewar period is matched. Given the “true world” without smoothing, we assume that the actual dividend growth is governed by a smoothness parameter \( \lambda \):

\[ g_{t+1} = (1 - \lambda) \left( a_g - 0.1 \times dp_t + \varepsilon^g_{t+1} \right) + \lambda \times \left( g_{\text{ave}} + \varepsilon^\text{ave}_{t+1} \right), \]  

(27)

where \( g_{\text{ave}} \) is the historical average dividend growth rate and \( \varepsilon^\text{ave}_{t+1} \) is a shock to this target. The more smoothed the dividend policy, the higher \( \lambda \) is.

We simulate stock returns under the null of no predictability and simulate dividend growth according to equation (27). 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, not too far from it empirical counterpart of 0.956 in the postwar period, and the dividend growth coefficients become insignificant at the 5% level for one, three, and five-year horizons.

The impact of dividend smoothing on return predictability is unclear, confirming our prior. When $\lambda$ increases from zero to 0.5, the return coefficient increases and is more significant. Thus dividend smoothing makes return appear to be more predictable. However, when $\lambda$ increases more, returns become slightly less predictable. Overall, the $p$-values of the return coefficients are always lower than the case of “true world”.

In sum, dividend smoothing makes the dividend yield more persistent and dividend growth less predictable. When dividends are sufficiently smoothed, dividends are not predictable even though all the predictability should come from the dividend side. The impact of dividend smoothing on return predictability is not clear, though it seems that some degree of smoothing makes returns appear to be more predictable.

### 4.1.3 Third case

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

$$g_{t+1} = -1.315 - 0.447 \times dp_t + \varepsilon_{t+1}$$  \hspace{1cm} (28)

$$r_{t+1} = 0.142 + 0.025 \times dp_t + \varepsilon'_{t+1}.$$  \hspace{1cm} (29)

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?

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

\[
g_{t+1} = (1 - \lambda) \left( -1.315 - 0.447 \times dp_t + \varepsilon_{t+1}^g \right) + \lambda \times (g_{\text{ave}} + \varepsilon_{t+1}^{\text{ave}}).
\]

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 and returns follow similar processes as in the prewar world, and 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. The impact of dividend smoothing on return predictability is similar to that of the second case. As dividends become more smoothed, dividend yields become more persistent and returns may appear to be predictable although they are simulated to be close to random.

4.2 Summary and discussion

We have shown, through three cases of simulation, that, even if dividends are supposed to be predictable without smoothing, dividend smoothing can bury this predictability in a finite sample. Therefore, the lack of dividend growth predictability in a finite sample could simply mean severe dividend smoothing rather than anything else.

Earlier, we provided two pieces of evidence indicating that dividends are very smoothed in the postwar period. First, we fit various dividend behavior models, all of which show that dividends
have become much more smoothed in the postwar period. Second, the dividend yield has become much more persistent in the postwar period.

The empirical evidence, combined with the simulation exercise, provides 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 of 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.

It has been forcefully argued before that, since the dividend yield is never too low or too high, at least one of dividend growth or returns must be predictable by the dividend yield. Our simulations show that this is not necessarily the case in a finite sample. When dividends are severely smoothed, even if the predictability is supposed to be on the dividend side, it is possible to observe that neither dividend growth nor returns are predictable by the dividend yield at normal horizons in a finite sample. As a result, the dividend yield mainly just predicts itself (e.g., Goyal and Welch (2003, 2008)).

Cochrane (2008) argues that the lack of dividend growth predictability by the dividend yield – the dog that did not bark – implies strong return predictability. While this seems to be the case if one possesses a long sample, our simulations show that this does not necessarily hold in a finite sample. An alternative interpretation is simply that dividends are severely smoothed.

It seems obvious that, in a finite period, managers can smooth dividends to make them unpredictable. It is not so once one considers the fact that a large literature has interpreted the lack of dividend growth predictability without questioning the role of dividend smoothing. Our simulations suggests that dividend smoothing could play a critical role. As a result, previous exercises that infer the importance of discount rate and cash flow news in asset price variations by comparing the relative extent of return and dividend growth predictability by dividend yield deserve to be reconsidered.
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 allows 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

$$rp = \frac{P_{t+1}}{P_t} \times \left[ 1 - \min \left( \frac{n_{t+1}}{n_t}, 1 \right) \right].$$  \hspace{1cm} (30)

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

$$si = \frac{P_{t+1}}{P_t} \times \left[ \max \left( \frac{n_{t+1}}{n_t}, 1 \right) - 1 \right].$$  \hspace{1cm} (31)

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:

\[
E_t = B_t - B_{t-1} + RP_t - SI_t + D_t,
\]

(32)

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. 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%. We then aggregate the data to obtain the market portfolio. The final annual data cover the period 1928-2006.

Figure 4 plots the aggregate dividend yield, total payout (= dividend + repurchase), and net payout (= dividend + repurchase - issuance) yield. Repurchases are essentially non-existent until the end of the 1970s. From then on repurchases separate total payout from dividends. Between 2000 and 2006, the average dividend yield is 0.016, and the average total payout yield is 0.031. This pattern is consistent with the findings in Skinner (2008) that repurchases are roughly equal to dividends at the aggregate level. The net payout yield is comparable to that plotted in Figure 5 of Larrain and Yogo (2008). It is much less persistent compared to the other two yields. It can be negative when issuance is greater than the sum of dividend and repurchase and the two troughs took place at the end of 1929 and 2000 when the market boom ended.

5.2 Predictability and yield decomposition: dividend yield

Table 5 reports results from testing the ability of the dividend yield to predict returns and cash flows from one to five-year horizons. We run the following predictive regression:

\[
y_t = \alpha_0 + \alpha_1 \times x_{t-1} + \varepsilon_t,
\]

(33)

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
For each regression coefficient, we provide both the simulated $p$-values, the details of which are provided in the appendix. We boldface the simulated $p$-values that are smaller or equal to 10%.

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 postwar sample are never significant at any horizon.

We also report the decomposition of the variance of the dividend yield using the one-year coefficients. Based on the full sample estimates, about 64.04% (34.49%) of dividend yield variance is due to dividend growth (returns). In stark contrast, based on the postwar sample, about -11.92% (103.84%) of 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 variation in the dividend yield is driven by cash flow news.

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

$$
\begin{align*}
v_t &= [\theta \times d_t - (\theta - 1) \times i_t] - p_t \\
&= E_t \left[ \sum_{j=0}^{\infty} \rho^j r_{t+1+j} \right] - E_t \left[ \sum_{j=0}^{\infty} \rho^j [\theta \times \Delta d_{t+1+j} - (\theta - 1) \times \Delta i_{t+1+j}] \right],
\end{align*}
$$

(34)

(35)

If a vector of $[\Delta d_t, r_t, d_p_t]$ follows a first-order VAR, then equation (23) indicates that

$$\frac{b_v}{1-\rho_\phi} + \frac{-b_g}{1-\rho_\phi} \approx 1.
$$

What this says is that 100% of dividend yield variance can be approximately decomposed into the return component and cash flow component.
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. 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. Equation (35) says that a higher net payout yield implies that either future expected returns will be higher, or future expected net payout growth will be lower.

Table 6 reports the predictability of returns and the growth rates in total payout (dividends + repurchases), issuances, and net payout (total payout - issuances) by the net payout yield. We note that the net payout yield predicts returns significantly in both the full sample (1928-06) and the postwar sample (46-06). In contrast, recall from Panel A of Table 5 that return predictability lacks statistical power even in the postwar period when the dividend yield is used. The finding that 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 corresponding coefficient in the postwar period is -0.015 and is insignificant, as are all the coefficients in the postwar period, irrespective of horizon. Recall from Table 5 that the dividend yield predicts dividend growth with the wrong sign in the postwar period. Therefore, when predicting cash flows with the net payout yield, measuring cash flows by adding repurchases to dividends helps predictability only marginally since no coefficient is significant in the postwar period.

The fourth and eighth columns of Table 6 report the predictability of the growth in issuances by the net payout ratio. In the full sample (column 4) there is a distinct lack of predictability, however, in the post war period (column 8) predictability is evident at every horizon. More interestingly, the fifth and ninth columns report the results that measure cash flows as the growth in net payout. In this case, 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 68.19% 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

$$\Delta d_t = \ln (E_t \times DE_t) - \ln (E_{t-1} \times DE_{t-1}) = \Delta e_t + \Delta de_t,$$

(36)

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

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

(37)

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 $ep_{0t-1}$ 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 $ep_{0t-1}$ ensures that, regardless of the fiscal year end, the price we use is way
ahead of earnings information. It is conservative and is likely to yield smaller predictive power.\footnote{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 reports the ability of the earnings yield to predict earnings growth, returns, and payout ratio. In contrast to the dividend yield coefficients in Table 5, 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 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: the coefficients in the postwar period are -0.651 at one-year horizon with an adjusted $R^2$ of 32%, and around -0.850 at the remaining horizons with an adjusted $R^2$ of just over 40%. 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 not statistically significant. Finally, the earnings yield hardly predicts the payout ratio.

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 consistent. In contrast to the case of the dividend yield, discount rates only play a small role in the yield decomposition, suggesting a larger role for cash flows. The results are stable for the full sample as well as for the postwar sample.
5.5 Portfolio tests

5.5.1 Smooth versus flexible dividend portfolios

If it is true that dividend smoothing has played a critical role in preventing us from finding cash flow predictability when dividend growth is used to measure cash flows, then we have the following additional testable hypotheses: for the postwar period, dividend growth should not be predictable for firms that have the most smoothed dividend payout; dividend growth should be relatively more predictable for firms that have the least smoothed dividend payout. In contrast, earnings growth should be predictable for all firms regardless of how much dividends are smoothed.

To test these hypotheses, we sort firms into three portfolios according to the smoothness parameter $S \left( = \frac{\sigma(\Delta d)}{\sigma(\Delta e)} \right)$ and then repeat the predictive regressions. 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. Again, similar to the case of the dividend yield, returns of both smooth and flexible portfolios are predictable by the earnings yield when treated separately, but not
when aggregated together (see Table 7). 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.

In summary, the discrepancy of dividend growth predictability between the smoothed and flexible dividend portfolios provides an interesting contrast to the consistency of earnings growth predictability between the same set of portfolios. The evidence strongly supports our hypothesis that dividend smoothing plays a crucial role in the lack of dividend growth predictability.

5.5.2 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.\textsuperscript{16}

Table 9 reports the results. Panel A shows that dividend growth is not predictable with the dividend yield for either the smooth or flexible earnings portfolios. However, for both portfolios, returns are predictable with the dividend yield, just as in the case of when we formed portfolios on smooth and flexible dividends. In Panel B, we examine predictability of earnings growth and returns using the earnings yield. 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.

\textsuperscript{16}This is the only table where we use earnings data directly from COMPUSTAT rather than from the clean surplus formula, hence the sample starts in 1951.
5.5.3 Stable versus volatile earnings growth portfolios

The volatility of both dividend and earnings growth has decreased in the postwar period. Such a reduction, however, might not be responsible for the lack of dividend growth predictability. Conceptually, even if dividends are not very volatile, a persistent shock to future dividend growth can still cause large variations in dividend yield (e.g., Bansal and Yaron (2004)). Therefore, our dividend smoothing measure is based on the volatility of dividend growth relative to earnings growth. Consistent with this view, Chen (2009) also 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. The firms with the lowest (highest) standard deviations comprise the stable (volatile) earnings growth portfolio.

Panel A of Table 10 reports the results using the dividend yield as the predictor. For both the stable and volatile earnings portfolios, dividend growth is predictable at some horizons for the full sample, but is not predictable for the postwar sample. Disaggregating stocks into portfolios by the volatility of earnings growth also leads to stock return predictability. Panel B reports predictability results using the earnings yield. For both the stable and volatile earnings portfolios, earnings growth is predictable for both the full and the postwar sample and for different horizons, though the predictability is stronger for the volatile portfolio. The decomposition, reported in the final row of Panel B, shows that cash flows are the major factor in the earnings yield, both in the full sample and the postwar sample.

It is clear 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.
5.5.4 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. Are these firms primarily responsible for the lack of dividend growth predictability? 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.”

Predictability results for these portfolios are reported in Table 11. In Panel A, the dividend growth coefficients for “dinosaurs” are all of the wrong sign in the postwar period; in comparison, the same coefficients for “non-dinosaurs” are all negative, though still insignificant. Payout “dinosaurs” thus appear to be more responsible for the lack of dividend growth predictability, but the issue remains even for “non-dinosaurs.” In Panel B, for both “dinosaurs” and “non-dinosaurs,” earnings growth is predictable in both the full sample and the postwar sample. We thus 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 Further robustness checks

We report the following robustness checks in the online appendix of this paper. First, 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.

Second, 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 current stock price is the sum of discounted expected future cash flows; its variation must reflect the revisions to expected future cash flows or to discount rates. The crucial question is “by how much of each” (Cochrane (2008)).

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. 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 severely smoothed 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.

Our take-away messages are that (i) dividend smoothing can severely affect dividend predictability in a finite sample, (ii) there is significant cash flow news in stock price variations, and (iii) when smoothed, dividends do not represent well the outlook of future cash flows.
References


Ang, A., 2002, Characterizing the ability of dividend yields to predict future dividends in log-linear present value models, working paper, Columbia University.


37


Stambaugh, R., 1986, Bias in regression with lagged stochastic regressors, working paper, University of Chicago.


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:

$$y_{i}^{t} = \xi_{i} + \alpha_{i} \times x_{t-1} + \varepsilon_{it}, \quad (A1)$$

where $y_{i}^{t}$, $i = 1, 2, \ldots, 5$, is the cumulative summation of $y_{t}$ from 1 to horizon $i$. Also suppose $y_{t}^{1}$ and $x_{t}$ follow AR(1) processes:

$$y_{t}^{1} = \beta_{0} + \beta_{1} \times y_{t-1}^{1} + \omega_{t}, \quad (A2)$$
$$x_{t} = \gamma_{0} + \gamma_{1} \times x_{t-1} + v_{t}, \quad (A3)$$

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 $\alpha_{\text{sim},i}$. We repeat the exercise 10,000 times to obtain the time series of $\alpha_{\text{sim},i}$. We finally compare the estimated $\alpha_{i}$ with the time series of $\alpha_{\text{sim},i}$ 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.
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.
Figure 2. Rolling-Window Regressions for the Lintner Model All panels correspond to variants of Lintner’s (1956) model (Equations (15)-(17)). 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. 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.
Figure 4. Aggregate Dividend Yield with (dotted line) and without Repurchase (solid line) and Net Payout Yield (dashed line).
Table 1: Summary Statistics

In Panel A, we summarize the S&P index annual data obtained from Robert Shiller’s website. $\Delta d$ 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 here includes shares repurchase (or $D = \text{dividend} + \text{repurchase}$). The net payout here is total payout minus equity issuance ($D-I$). This sample covers annual data during 1928-2006.

<table>
<thead>
<tr>
<th>Panel A: S&amp;P</th>
<th>$\Delta d$</th>
<th>$\Delta e$</th>
<th>$\frac{E}{P}$</th>
<th>$\frac{D}{P}$</th>
<th>$S = \frac{\sigma(\Delta d)}{\sigma(\Delta e)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1872-2006</td>
<td>Mean</td>
<td>0.034</td>
<td>0.039</td>
<td>0.045</td>
<td>0.075</td>
</tr>
<tr>
<td></td>
<td>(sd)</td>
<td>(0.12)</td>
<td>(0.25)</td>
<td>(0.02)</td>
<td>(0.03)</td>
</tr>
<tr>
<td></td>
<td>AR(1)</td>
<td>0.256</td>
<td>0.024</td>
<td>0.781</td>
<td>0.740</td>
</tr>
<tr>
<td>1872-1945</td>
<td>Mean</td>
<td>0.013</td>
<td>0.012</td>
<td>0.053</td>
<td>0.077</td>
</tr>
<tr>
<td></td>
<td>(sd)</td>
<td>(0.16)</td>
<td>(0.29)</td>
<td>(0.14)</td>
<td>(0.03)</td>
</tr>
<tr>
<td></td>
<td>AR(1)</td>
<td>0.204</td>
<td>-0.017</td>
<td>0.518</td>
<td>0.621</td>
</tr>
<tr>
<td>1946-2006</td>
<td>Mean</td>
<td>0.059</td>
<td>0.073</td>
<td>0.036</td>
<td>0.073</td>
</tr>
<tr>
<td></td>
<td>(sd)</td>
<td>(0.05)</td>
<td>(0.18)</td>
<td>(0.01)</td>
<td>(0.03)</td>
</tr>
<tr>
<td></td>
<td>AR(1)</td>
<td>0.473</td>
<td>0.089</td>
<td>0.926</td>
<td>0.832</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: CRSP (D = dividend + repurchase)</th>
<th>$\Delta d$</th>
<th>$\Delta e$</th>
<th>$\frac{D}{P}$</th>
<th>$\frac{E}{P}$</th>
<th>$\frac{D-I}{P}$</th>
<th>$S = \frac{\sigma(\Delta d)}{\sigma(\Delta e)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1928-2006</td>
<td>Mean</td>
<td>0.054</td>
<td>0.066</td>
<td>0.045</td>
<td>0.072</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>(sd)</td>
<td>(0.15)</td>
<td>(0.53)</td>
<td>(0.01)</td>
<td>(0.04)</td>
<td>(0.02)</td>
</tr>
<tr>
<td></td>
<td>AR(1)</td>
<td>0.115</td>
<td>-0.124</td>
<td>0.637</td>
<td>0.588</td>
<td>0.666</td>
</tr>
<tr>
<td>1946-2006</td>
<td>Mean</td>
<td>0.069</td>
<td>0.082</td>
<td>0.042</td>
<td>0.074</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>(sd)</td>
<td>(0.10)</td>
<td>(0.48)</td>
<td>(0.01)</td>
<td>(0.04)</td>
<td>(0.02)</td>
</tr>
<tr>
<td></td>
<td>AR(1)</td>
<td>-0.081</td>
<td>-0.139</td>
<td>0.765</td>
<td>0.734</td>
<td>0.723</td>
</tr>
</tbody>
</table>
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: $\Delta D_t = \alpha_0 + \alpha_1 E_t + \alpha_2 D_{t-1} + u_t$

<table>
<thead>
<tr>
<th></th>
<th>$c$</th>
<th>$E_t$</th>
<th>$D_{t-1}$</th>
<th>$R^2$</th>
<th>SA</th>
<th>TPR</th>
<th>Chow 1945</th>
</tr>
</thead>
<tbody>
<tr>
<td>1872-2006</td>
<td>0.035</td>
<td>0.052</td>
<td>−0.079</td>
<td>0.73</td>
<td>0.08</td>
<td>0.08</td>
<td>2.656 [0.05]</td>
</tr>
<tr>
<td></td>
<td>(0.42)</td>
<td>(10.99)</td>
<td>(5.87)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1872-1945</td>
<td>0.005</td>
<td>0.248</td>
<td>−0.373</td>
<td>0.60</td>
<td>0.37</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.32)</td>
<td>(10.22)</td>
<td>(8.93)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1946-2006</td>
<td>0.120</td>
<td>0.054</td>
<td>−0.090</td>
<td>0.68</td>
<td>0.09</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.74)</td>
<td>(7.69)</td>
<td>(4.25)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F$ - $T$est</td>
<td>766.43</td>
<td>175.08</td>
<td>[0.00]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel B: $\Delta D_t = \beta_0 + \beta_1 \times \Delta E_t + \beta_2 \times \Delta D_{t-1} + \varepsilon_t$

<table>
<thead>
<tr>
<th></th>
<th>$c$</th>
<th>$\Delta E_t$</th>
<th>$\Delta D_{t-1}$</th>
<th>$R^2$</th>
<th>SA</th>
<th>TPR</th>
<th>Chow 1945</th>
</tr>
</thead>
<tbody>
<tr>
<td>1872-2006</td>
<td>0.025</td>
<td>0.037</td>
<td>0.825</td>
<td>0.81</td>
<td>0.17</td>
<td>0.22</td>
<td>3.677 [0.01]</td>
</tr>
<tr>
<td></td>
<td>(1.38)</td>
<td>(7.39)</td>
<td>(17.25)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1872-1945</td>
<td>0.001</td>
<td>0.237</td>
<td>0.284</td>
<td>0.35</td>
<td>0.72</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.20)</td>
<td>(6.09)</td>
<td>(2.94)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1946-2006</td>
<td>0.062</td>
<td>0.036</td>
<td>0.808</td>
<td>0.79</td>
<td>0.19</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.47)</td>
<td>(5.67)</td>
<td>(10.91)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F$ - $T$est</td>
<td>773.78</td>
<td>50.15</td>
<td>[0.00]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel C: $\Delta D_t = \gamma_0 + \gamma_1 E_t + \gamma_2 \times \Delta D_{t-1} + \nu_t$

<table>
<thead>
<tr>
<th></th>
<th>$c$</th>
<th>$E_t$</th>
<th>$\Delta D_{t-1}$</th>
<th>$R^2$</th>
<th>Chow 1945</th>
</tr>
</thead>
<tbody>
<tr>
<td>1872-2006</td>
<td>−0.012</td>
<td>0.011</td>
<td>0.652</td>
<td>0.78</td>
<td>1.311 [0.27]</td>
</tr>
<tr>
<td></td>
<td>(0.57)</td>
<td>(5.29)</td>
<td>(8.46)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1872-1945</td>
<td>−0.056</td>
<td>0.093</td>
<td>0.061</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.90)</td>
<td>(3.45)</td>
<td>(0.53)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1946-2006</td>
<td>−0.025</td>
<td>0.011</td>
<td>0.657</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.47)</td>
<td>(3.29)</td>
<td>(4.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F$ - $T$est</td>
<td>618.87</td>
<td>30.39</td>
<td>[0.00]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel D: $\ln \left[ \frac{D_{t+1}}{D_t} \right] + \frac{D_t}{D_{t-1}} = \lambda_0 + \lambda_1 \times \ln \left[ \frac{P_t + D_t}{P_{t-1}} \right] + \lambda_2 \times \ln \left[ \frac{D_t}{P_{t-1}} \right] + \omega_{t+1}$

<table>
<thead>
<tr>
<th></th>
<th>$\lambda_0$</th>
<th>$\ln \frac{P_t + D_t}{P_{t-1}}$</th>
<th>$\ln \frac{D_t}{P_{t-1}}$</th>
<th>$R^2$</th>
<th>Chow 1945</th>
</tr>
</thead>
<tbody>
<tr>
<td>1872-2006</td>
<td>−0.026</td>
<td>0.461</td>
<td>−0.021</td>
<td>0.38</td>
<td>21.24 [0.00]</td>
</tr>
<tr>
<td></td>
<td>(0.33)</td>
<td>(6.21)</td>
<td>(0.89)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1872-1945</td>
<td>−0.565</td>
<td>0.673</td>
<td>−0.198</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.60)</td>
<td>(9.01)</td>
<td>(2.72)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1946-2006</td>
<td>0.299</td>
<td>0.003</td>
<td>0.061</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.49)</td>
<td>(0.06)</td>
<td>(2.87)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F$ - $T$est</td>
<td>176.49</td>
<td>246.24</td>
<td>[0.00]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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, \( dg_1^t \) is the annual dividend growth, \( dg_5^t \) is the five-year dividend growth, \( r_1^t \) is annual return, and \( r_5^t \) 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. See Appendix for more details. 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.

<table>
<thead>
<tr>
<th>Panel A: Actual Data</th>
<th>( dg_1^t )</th>
<th>( dg_3^t )</th>
<th>( dg_5^t )</th>
<th>( r_1^t )</th>
<th>( r_3^t )</th>
<th>( r_5^t )</th>
<th>( AR(1) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1872-1945</td>
<td>-0.448</td>
<td>-0.596</td>
<td>-0.406</td>
<td>0.024</td>
<td>0.303</td>
<td>0.636</td>
<td>0.557</td>
</tr>
<tr>
<td></td>
<td>[0.00]</td>
<td>[0.00]</td>
<td>[0.07]</td>
<td>[0.27]</td>
<td>[0.07]</td>
<td>[0.01]</td>
<td>[0.00]</td>
</tr>
<tr>
<td>1946-2006</td>
<td>0.026</td>
<td>0.076</td>
<td>0.088</td>
<td>0.101</td>
<td>0.289</td>
<td>0.505</td>
<td>0.956</td>
</tr>
<tr>
<td></td>
<td>[0.35]</td>
<td>[0.25]</td>
<td>[0.26]</td>
<td>[0.17]</td>
<td>[0.19]</td>
<td>[0.15]</td>
<td>[0.00]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Simulated Data</th>
<th>( dg_1^t )</th>
<th>( dg_3^t )</th>
<th>( dg_5^t )</th>
<th>( r_1^t )</th>
<th>( r_3^t )</th>
<th>( r_5^t )</th>
<th>( AR(1) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1872-1945</td>
<td>-0.460</td>
<td>-0.679</td>
<td>-0.798</td>
<td>0.032</td>
<td>0.086</td>
<td>0.135</td>
<td>0.565</td>
</tr>
<tr>
<td></td>
<td>[0.00]</td>
<td>[0.00]</td>
<td>[0.00]</td>
<td>[0.37]</td>
<td>[0.34]</td>
<td>[0.32]</td>
<td>[0.00]</td>
</tr>
<tr>
<td>1946-2006</td>
<td>-0.033</td>
<td>-0.096</td>
<td>-0.156</td>
<td>0.020</td>
<td>0.058</td>
<td>0.094</td>
<td>0.983</td>
</tr>
<tr>
<td></td>
<td>[0.43]</td>
<td>[0.42]</td>
<td>[0.41]</td>
<td>[0.17]</td>
<td>[0.18]</td>
<td>[0.19]</td>
<td>[0.00]</td>
</tr>
</tbody>
</table>
Table 4: Dividend Smoothing and Predictability by 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:

\[ g_{t+1} = a_g - 0.1 \times dp_t + \varepsilon_g^{t+1} \]

\[ r_{t+1} = a_r + \varepsilon_r^{t+1} \]

where \( g_{t+1} \) is dividend growth rate and \( r_{t+1} \) stock return. The residuals \( \varepsilon_g^{t+1} \) and \( \varepsilon_r^{t+1} \) 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 \( \lambda \):

\[ g_{t+1} = (1 - \lambda) \left( a_g - 0.1 \times dp_t + \varepsilon_g^{t+1} \right) + \lambda \times (g_{ave} + \varepsilon_{ave}^{t+1}) \]

where \( g_{ave} \) is historical average dividend growth rate and \( \varepsilon_{ave}^{t+1} \) 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:

\[ g_{t+1} = (1 - \lambda) \left( -1.315 - 0.447 \times dp_t + \varepsilon_g^{t+1} \right) + \lambda \times (g_{ave} + \varepsilon_{ave}^{t+1}) \]

\[ r_{t+1} = 0.142 + 0.025 \times dp_t + \varepsilon_r^{t+1} \]

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.

<table>
<thead>
<tr>
<th>( \lambda )</th>
<th>( dg_1^t )</th>
<th>( dg_3^t )</th>
<th>( dg_5^t )</th>
<th>( r_1^t )</th>
<th>( r_3^t )</th>
<th>( r_5^t )</th>
<th>AR(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-0.124</td>
<td>-0.325</td>
<td>-0.477</td>
<td>0.052</td>
<td>0.146</td>
<td>0.227</td>
<td>0.903</td>
</tr>
<tr>
<td>[0.00]</td>
<td>[0.00]</td>
<td>[0.00]</td>
<td>[0.19]</td>
<td>[0.18]</td>
<td>[0.17]</td>
<td>[0.00]</td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>-0.097</td>
<td>-0.272</td>
<td>-0.423</td>
<td>0.053</td>
<td>0.150</td>
<td>0.234</td>
<td>0.946</td>
</tr>
<tr>
<td>[0.01]</td>
<td>[0.01]</td>
<td>[0.01]</td>
<td>[0.11]</td>
<td>[0.11]</td>
<td>[0.12]</td>
<td>[0.00]</td>
<td></td>
</tr>
<tr>
<td>0.95</td>
<td>-0.081</td>
<td>-0.239</td>
<td>-0.389</td>
<td>0.037</td>
<td>0.105</td>
<td>0.164</td>
<td>0.973</td>
</tr>
<tr>
<td>[0.09]</td>
<td>[0.10]</td>
<td>[0.10]</td>
<td>[0.12]</td>
<td>[0.13]</td>
<td>[0.14]</td>
<td>[0.00]</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-0.079</td>
<td>-0.233</td>
<td>-0.381</td>
<td>0.033</td>
<td>0.095</td>
<td>0.148</td>
<td>0.976</td>
</tr>
<tr>
<td>[0.15]</td>
<td>[0.15]</td>
<td>[0.16]</td>
<td>[0.13]</td>
<td>[0.14]</td>
<td>[0.15]</td>
<td>[0.00]</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \lambda )</th>
<th>( dg_1^t )</th>
<th>( dg_3^t )</th>
<th>( dg_5^t )</th>
<th>( r_1^t )</th>
<th>( r_3^t )</th>
<th>( r_5^t )</th>
<th>AR(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-0.456</td>
<td>-0.809</td>
<td>-0.882</td>
<td>0.058</td>
<td>0.126</td>
<td>0.170</td>
<td>0.533</td>
</tr>
<tr>
<td>[0.00]</td>
<td>[0.00]</td>
<td>[0.00]</td>
<td>[0.29]</td>
<td>[0.29]</td>
<td>[0.28]</td>
<td>[0.00]</td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>-0.236</td>
<td>-0.544</td>
<td>-0.708</td>
<td>0.062</td>
<td>0.150</td>
<td>0.214</td>
<td>0.769</td>
</tr>
<tr>
<td>[0.00]</td>
<td>[0.00]</td>
<td>[0.00]</td>
<td>[0.24]</td>
<td>[0.24]</td>
<td>[0.24]</td>
<td>[0.00]</td>
<td></td>
</tr>
<tr>
<td>0.95</td>
<td>-0.104</td>
<td>-0.305</td>
<td>-0.493</td>
<td>0.051</td>
<td>0.139</td>
<td>0.212</td>
<td>0.955</td>
</tr>
<tr>
<td>[0.03]</td>
<td>[0.03]</td>
<td>[0.04]</td>
<td>[0.08]</td>
<td>[0.10]</td>
<td>[0.11]</td>
<td>[0.00]</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-0.096</td>
<td>-0.284</td>
<td>-0.462</td>
<td>0.039</td>
<td>0.105</td>
<td>0.159</td>
<td>0.970</td>
</tr>
<tr>
<td>[0.14]</td>
<td>[0.14]</td>
<td>[0.15]</td>
<td>[0.10]</td>
<td>[0.12]</td>
<td>[0.14]</td>
<td>[0.00]</td>
<td></td>
</tr>
</tbody>
</table>
Table 5: Predictability by 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. See Appendix for more details. 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.

<table>
<thead>
<tr>
<th>Horizon</th>
<th>$g$</th>
<th>$R^2$</th>
<th>$r$</th>
<th>$R^2$</th>
<th>$g$</th>
<th>$R^2$</th>
<th>$r$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.087</td>
<td>8.7</td>
<td>0.049</td>
<td>1.2</td>
<td>0.012</td>
<td>-0.9</td>
<td>0.103</td>
<td>6.9</td>
</tr>
<tr>
<td></td>
<td>[0.01]</td>
<td></td>
<td>[0.28]</td>
<td></td>
<td>[0.67]</td>
<td></td>
<td>[0.20]</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-0.114</td>
<td>5.6</td>
<td>0.135</td>
<td>3.2</td>
<td>0.03</td>
<td>0.7</td>
<td>0.207</td>
<td>15.9</td>
</tr>
<tr>
<td></td>
<td>[0.07]</td>
<td></td>
<td>[0.22]</td>
<td></td>
<td>[0.79]</td>
<td></td>
<td>[0.17]</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-0.121</td>
<td>3.5</td>
<td>0.199</td>
<td>5.3</td>
<td>0.033</td>
<td>-0.2</td>
<td>0.274</td>
<td>22.8</td>
</tr>
<tr>
<td></td>
<td>[0.14]</td>
<td></td>
<td>[0.24]</td>
<td></td>
<td>[0.76]</td>
<td></td>
<td>[0.20]</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-0.097</td>
<td>1.1</td>
<td>0.277</td>
<td>9.9</td>
<td>0.042</td>
<td>-0.1</td>
<td>0.349</td>
<td>29.3</td>
</tr>
<tr>
<td></td>
<td>[0.27]</td>
<td></td>
<td>[0.22]</td>
<td></td>
<td>[0.78]</td>
<td></td>
<td>[0.20]</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-0.067</td>
<td>-0.2</td>
<td>0.349</td>
<td>15.7</td>
<td>0.036</td>
<td>-0.6</td>
<td>0.455</td>
<td>36.2</td>
</tr>
<tr>
<td></td>
<td>[0.38]</td>
<td></td>
<td>[0.21]</td>
<td></td>
<td>[0.74]</td>
<td></td>
<td>[0.17]</td>
<td></td>
</tr>
<tr>
<td>Decomposition</td>
<td>64.04%</td>
<td>34.49%</td>
<td>-11.92%</td>
<td>103.84%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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. See Appendix for more details. We boldface the $p$-value if it is lower than or equal to 0.10. The estimated value for $\theta$ is 1.6933.

<table>
<thead>
<tr>
<th>Horizon</th>
<th>$r$</th>
<th>$\Delta d$</th>
<th>$\Delta i$</th>
<th>$\theta \Delta d - (\theta - 1) \Delta i$</th>
<th>$r$</th>
<th>$\Delta d$</th>
<th>$\Delta i$</th>
<th>$\theta \Delta d - (\theta - 1) \Delta i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.054</td>
<td>-0.044</td>
<td>0.150</td>
<td>-0.179</td>
<td>0.086</td>
<td>-0.015</td>
<td>0.272</td>
<td>-0.215</td>
</tr>
<tr>
<td></td>
<td>[0.04]</td>
<td>[0.03]</td>
<td>[0.29]</td>
<td>[0.13]</td>
<td>[0.04]</td>
<td>[0.34]</td>
<td>[0.06]</td>
<td>[0.10]</td>
</tr>
<tr>
<td>2</td>
<td>0.106</td>
<td>-0.032</td>
<td>0.322</td>
<td>-0.277</td>
<td>0.177</td>
<td>0.007</td>
<td>0.479</td>
<td>-0.320</td>
</tr>
<tr>
<td></td>
<td>[0.05]</td>
<td>[0.23]</td>
<td>[0.13]</td>
<td>[0.04]</td>
<td>[0.01]</td>
<td>[0.64]</td>
<td>[0.02]</td>
<td>[0.05]</td>
</tr>
<tr>
<td>3</td>
<td>0.133</td>
<td>-0.021</td>
<td>0.416</td>
<td>-0.323</td>
<td>0.219</td>
<td>0.006</td>
<td>0.529</td>
<td>-0.357</td>
</tr>
<tr>
<td></td>
<td>[0.08]</td>
<td>[0.36]</td>
<td>[0.14]</td>
<td>[0.07]</td>
<td>[0.02]</td>
<td>[0.62]</td>
<td>[0.04]</td>
<td>[0.09]</td>
</tr>
<tr>
<td>4</td>
<td>0.148</td>
<td>0.003</td>
<td>0.518</td>
<td>-0.354</td>
<td>0.231</td>
<td>0.021</td>
<td>0.543</td>
<td>-0.341</td>
</tr>
<tr>
<td></td>
<td>[0.11]</td>
<td>[0.50]</td>
<td>[0.13]</td>
<td>[0.09]</td>
<td>[0.04]</td>
<td>[0.69]</td>
<td>[0.07]</td>
<td>[0.17]</td>
</tr>
<tr>
<td>5</td>
<td>0.157</td>
<td>0.015</td>
<td>0.604</td>
<td>-0.393</td>
<td>0.301</td>
<td>0.004</td>
<td>0.668</td>
<td>-0.456</td>
</tr>
<tr>
<td></td>
<td>[0.15]</td>
<td>[0.54]</td>
<td>[0.13]</td>
<td>[0.10]</td>
<td>[0.02]</td>
<td>[0.60]</td>
<td>[0.05]</td>
<td>[0.11]</td>
</tr>
<tr>
<td>Decomposition</td>
<td>21.78%</td>
<td>30.19%</td>
<td>42.24%</td>
<td>72.43%</td>
<td>27.80%</td>
<td>8.40%</td>
<td>60.79%</td>
<td>69.19%</td>
</tr>
</tbody>
</table>
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 \((ep_{t-1}^0 \text{ or } ep_{t-1}^0)\). 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. See Appendix for more details. 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.

<table>
<thead>
<tr>
<th>Horizon</th>
<th>(eg)</th>
<th>(R^2)</th>
<th>(r)</th>
<th>(R^2)</th>
<th>(1 - \rho)(de))</th>
<th>(R^2)</th>
<th>(eg)</th>
<th>(R^2)</th>
<th>(r)</th>
<th>(R^2)</th>
<th>(1 - \rho)(de))</th>
<th>(R^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.721</td>
<td>36.2</td>
<td>0.037</td>
<td>2.6</td>
<td>-0.006</td>
<td>1.8</td>
<td>-0.651</td>
<td>32.1</td>
<td>0.027</td>
<td>1.9</td>
<td>-0.006</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>[0.00]</td>
<td></td>
<td></td>
<td></td>
<td>[0.10]</td>
<td></td>
<td>[0.02]</td>
<td></td>
<td></td>
<td>[0.12]</td>
<td>[0.45]</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-0.855</td>
<td>40.2</td>
<td>0.035</td>
<td>0.4</td>
<td>-0.006</td>
<td>-0.2</td>
<td>-0.827</td>
<td>40.7</td>
<td>0.038</td>
<td>2.7</td>
<td>-0.004</td>
<td>-0.9</td>
</tr>
<tr>
<td></td>
<td>[0.00]</td>
<td></td>
<td></td>
<td></td>
<td>[0.19]</td>
<td></td>
<td>[0.00]</td>
<td></td>
<td></td>
<td>[0.13]</td>
<td>[0.56]</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-0.839</td>
<td>38.2</td>
<td>0.032</td>
<td>-0.2</td>
<td>-0.007</td>
<td>-0.4</td>
<td>-0.849</td>
<td>42.7</td>
<td>0.030</td>
<td>-0.1</td>
<td>-0.003</td>
<td>-1.6</td>
</tr>
<tr>
<td></td>
<td>[0.00]</td>
<td></td>
<td></td>
<td></td>
<td>[0.26]</td>
<td></td>
<td>[0.02]</td>
<td></td>
<td></td>
<td>[0.24]</td>
<td>[0.58]</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-0.898</td>
<td>40.4</td>
<td>0.085</td>
<td>5.6</td>
<td>-0.006</td>
<td>-0.8</td>
<td>-0.880</td>
<td>45.3</td>
<td>0.037</td>
<td>0.2</td>
<td>0.000</td>
<td>-1.8</td>
</tr>
<tr>
<td></td>
<td>[0.00]</td>
<td></td>
<td></td>
<td></td>
<td>[0.08]</td>
<td></td>
<td>[0.02]</td>
<td></td>
<td></td>
<td>[0.23]</td>
<td>[0.64]</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-0.846</td>
<td>37.6</td>
<td>0.071</td>
<td>3.85</td>
<td>-0.008</td>
<td>-0.7</td>
<td>-0.850</td>
<td>43.7</td>
<td>0.022</td>
<td>-1.30</td>
<td>0.000</td>
<td>-1.8</td>
</tr>
<tr>
<td></td>
<td>[0.01]</td>
<td></td>
<td></td>
<td></td>
<td>[0.44]</td>
<td></td>
<td>[0.04]</td>
<td></td>
<td></td>
<td>[0.35]</td>
<td>[0.59]</td>
<td></td>
</tr>
</tbody>
</table>

| Decomposition | 95.41% | 4.93% | 0.89% | 96.62% | 4.07% | 0.96% |
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. See Appendix for more details. 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.

<table>
<thead>
<tr>
<th>Horizon</th>
<th>Smooth portfolios</th>
<th>Flexible portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Predictability by dividend yield</td>
<td></td>
<td></td>
</tr>
<tr>
<td>g</td>
<td>r</td>
<td>g</td>
</tr>
<tr>
<td>1</td>
<td>0.142</td>
<td>0.153</td>
</tr>
<tr>
<td>[0.01]</td>
<td>[0.02]</td>
<td>[0.71]</td>
</tr>
<tr>
<td>3</td>
<td>-0.145</td>
<td>0.476</td>
</tr>
<tr>
<td>[0.14]</td>
<td>[0.00]</td>
<td>[0.86]</td>
</tr>
<tr>
<td>5</td>
<td>-0.022</td>
<td>0.683</td>
</tr>
<tr>
<td>[0.44]</td>
<td>[0.01]</td>
<td>[0.86]</td>
</tr>
<tr>
<td>Decomposition</td>
<td>49.31%</td>
<td>53.30%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Horizon</th>
<th>eg</th>
<th>r</th>
<th>eg</th>
<th>r</th>
<th>eg</th>
<th>r</th>
<th>eg</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel B: Predictability by earnings yield</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-0.779</td>
<td>0.055</td>
<td>-0.714</td>
<td>0.028</td>
<td>-0.769</td>
<td>0.043</td>
<td>-0.449</td>
<td>0.106</td>
</tr>
<tr>
<td>[0.01]</td>
<td>[0.01]</td>
<td>[0.01]</td>
<td>[0.08]</td>
<td>[0.00]</td>
<td>[0.010]</td>
<td>[0.05]</td>
<td>[0.00]</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-0.613</td>
<td>0.049</td>
<td>-0.848</td>
<td>0.027</td>
<td>-0.844</td>
<td>0.081</td>
<td>-0.570</td>
<td>0.152</td>
</tr>
<tr>
<td>[0.04]</td>
<td>[0.08]</td>
<td>[0.01]</td>
<td>[0.23]</td>
<td>[0.00]</td>
<td>[0.11]</td>
<td>[0.06]</td>
<td>[0.02]</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-0.856</td>
<td>0.086</td>
<td>-0.851</td>
<td>0.017</td>
<td>-0.737</td>
<td>0.151</td>
<td>-0.518</td>
<td>0.254</td>
</tr>
<tr>
<td>[0.01]</td>
<td>[0.03]</td>
<td>[0.03]</td>
<td>[0.37]</td>
<td>[0.00]</td>
<td>[0.04]</td>
<td>[0.17]</td>
<td>[0.00]</td>
<td></td>
</tr>
<tr>
<td>Decomposition</td>
<td>96.18%</td>
<td>6.81%</td>
<td>98.12%</td>
<td>3.97%</td>
<td>95.75%</td>
<td>5.43%</td>
<td>75.45%</td>
<td>17.94%</td>
</tr>
</tbody>
</table>
Table 9: Smooth- versus Flexible-Earnings Portfolios

We sort firms into three portfolios according to the earnings smoothness measure. The smoothness measure is computed as the ratio between 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). The firms with the lowest (highest) ratios consist the smooth-(flexible-) earnings 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. See Appendix for more details. 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. The sample period starts 1951 as COMPUSTAT data are required to compute the earnings smoothness measure.

<table>
<thead>
<tr>
<th>Horizon</th>
<th>Smooth ( g )</th>
<th>Smooth ( r )</th>
<th>Flexible ( g )</th>
<th>Flexible ( r )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.001</td>
<td>0.312</td>
<td>−0.086</td>
<td>0.188</td>
</tr>
<tr>
<td></td>
<td>[0.65]</td>
<td>[0.00]</td>
<td>[0.25]</td>
<td>[0.03]</td>
</tr>
<tr>
<td>3</td>
<td>0.104</td>
<td>0.673</td>
<td>0.029</td>
<td>0.474</td>
</tr>
<tr>
<td></td>
<td>[0.88]</td>
<td>[0.00]</td>
<td>[0.74]</td>
<td>[0.02]</td>
</tr>
<tr>
<td>5</td>
<td>0.191</td>
<td>0.899</td>
<td>−0.009</td>
<td>0.630</td>
</tr>
<tr>
<td></td>
<td>[0.91]</td>
<td>[0.00]</td>
<td>[0.64]</td>
<td>[0.03]</td>
</tr>
<tr>
<td>Decomposition</td>
<td>−0.37%</td>
<td>98.76%</td>
<td>30.98%</td>
<td>67.97%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Horizon</th>
<th>Smooth ( g )</th>
<th>Smooth ( r )</th>
<th>Flexible ( g )</th>
<th>Flexible ( r )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>−0.042</td>
<td>0.208</td>
<td>−0.375</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>[0.40]</td>
<td>[0.02]</td>
<td>[0.01]</td>
<td>[0.08]</td>
</tr>
<tr>
<td>3</td>
<td>−0.090</td>
<td>0.418</td>
<td>−0.901</td>
<td>0.020</td>
</tr>
<tr>
<td></td>
<td>[0.36]</td>
<td>[0.02]</td>
<td>[0.01]</td>
<td>[0.24]</td>
</tr>
<tr>
<td>5</td>
<td>−0.095</td>
<td>0.606</td>
<td>−0.906</td>
<td>0.080</td>
</tr>
<tr>
<td></td>
<td>[0.43]</td>
<td>[0.01]</td>
<td>[0.03]</td>
<td>[0.09]</td>
</tr>
<tr>
<td>Decomposition</td>
<td>14.40%</td>
<td>71.64%</td>
<td>85.96%</td>
<td>5.00%</td>
</tr>
</tbody>
</table>

54
Table 10: Stable versus Volatile Earnings Growth Portfolios

We sort firms into three portfolios according to the standard deviation of annual earnings growth. The firms with the lowest (highest) standard deviations consist the stable-(volatile-) earnings growth portfolio. For the stable and volatile 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. See Appendix for more details. 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.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Predictability by dividend yield</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horizon</td>
<td>(g)</td>
<td>(r)</td>
<td>(g)</td>
<td>(r)</td>
<td>(g)</td>
<td>(r)</td>
</tr>
<tr>
<td>1</td>
<td>−0.173</td>
<td>0.164</td>
<td>0.012</td>
<td>0.248</td>
<td>−0.322</td>
<td>0.062</td>
</tr>
<tr>
<td></td>
<td>[0.00]</td>
<td>[0.02]</td>
<td>[0.77]</td>
<td>[0.00]</td>
<td>[0.00]</td>
<td>[0.14]</td>
</tr>
<tr>
<td>3</td>
<td>−0.168</td>
<td>0.495</td>
<td>0.098</td>
<td>0.587</td>
<td>−0.469</td>
<td>0.297</td>
</tr>
<tr>
<td></td>
<td>[0.13]</td>
<td>[0.00]</td>
<td>[0.91]</td>
<td>[0.00]</td>
<td>[0.00]</td>
<td>[0.05]</td>
</tr>
<tr>
<td>5</td>
<td>−0.053</td>
<td>0.692</td>
<td>0.119</td>
<td>0.852</td>
<td>−0.361</td>
<td>0.483</td>
</tr>
<tr>
<td></td>
<td>[0.46]</td>
<td>[0.00]</td>
<td>[0.87]</td>
<td>[0.00]</td>
<td>[0.11]</td>
<td>[0.04]</td>
</tr>
<tr>
<td>Decomposition</td>
<td>52.66%</td>
<td>50.08%</td>
<td>−5.30%</td>
<td>108.25%</td>
<td>83.85%</td>
<td>16.11%</td>
</tr>
<tr>
<td><strong>Panel B: Predictability by earnings yield</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horizon</td>
<td>(eg)</td>
<td>(r)</td>
<td>(eg)</td>
<td>(r)</td>
<td>(eg)</td>
<td>(r)</td>
</tr>
<tr>
<td>1</td>
<td>−0.790</td>
<td>0.072</td>
<td>−0.196</td>
<td>0.092</td>
<td>−0.651</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>[0.00]</td>
<td>[0.01]</td>
<td>[0.08]</td>
<td>[0.02]</td>
<td>[0.00]</td>
<td>[0.45]</td>
</tr>
<tr>
<td>3</td>
<td>−0.477</td>
<td>0.065</td>
<td>−0.405</td>
<td>0.194</td>
<td>−0.855</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>[0.15]</td>
<td>[0.09]</td>
<td>[0.07]</td>
<td>[0.04]</td>
<td>[0.00]</td>
<td>[0.43]</td>
</tr>
<tr>
<td>5</td>
<td>−0.839</td>
<td>0.142</td>
<td>−0.444</td>
<td>0.263</td>
<td>−0.914</td>
<td>0.048</td>
</tr>
<tr>
<td></td>
<td>[0.01]</td>
<td>[0.02]</td>
<td>[0.16]</td>
<td>[0.05]</td>
<td>[0.00]</td>
<td>[0.17]</td>
</tr>
<tr>
<td>Decomposition</td>
<td>93.54%</td>
<td>8.55%</td>
<td>72.27%</td>
<td>34.10%</td>
<td>93.62%</td>
<td>0.47%</td>
</tr>
</tbody>
</table>
Following Skinner (2008), we focus on a small group of firms that consistently make both dividend payments and repurchases and call them payout “dinosaurs.” They are firms making dividend payments for more than 15 years and making repurchases for more than 10 years as recorded by COMPUSTAT. For both the payout “dinosaurs” and “non-dinosaurs” 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. See Appendix for more details. 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.

### Table 11: Payout “Dinosaurs” versus “Non-dinosaurs”

Following Skinner (2008), we focus on a small group of firms that consistently make both dividend payments and repurchases and call them payout “dinosaurs.” They are firms making dividend payments for more than 15 years and making repurchases for more than 10 years as recorded by COMPUSTAT. For both the payout “dinosaurs” and “non-dinosaurs” 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. See Appendix for more details. 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.

<table>
<thead>
<tr>
<th>Horizon</th>
<th>“Dinosaurs”</th>
<th></th>
<th>“Non-dinosaurs”</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>g</td>
<td>r</td>
<td>g</td>
<td>r</td>
<td>g</td>
</tr>
<tr>
<td>1</td>
<td>-0.194</td>
<td>0.168</td>
<td>0.000</td>
<td>0.232</td>
</tr>
<tr>
<td></td>
<td>[0.00]</td>
<td>[0.00]</td>
<td>[0.72]</td>
<td>[0.00]</td>
</tr>
<tr>
<td>3</td>
<td>-0.154</td>
<td>0.499</td>
<td>0.109</td>
<td>0.547</td>
</tr>
<tr>
<td></td>
<td>[0.17]</td>
<td>[0.00]</td>
<td>[0.95]</td>
<td>[0.00]</td>
</tr>
<tr>
<td>5</td>
<td>0.000</td>
<td>0.718</td>
<td>0.144</td>
<td>0.804</td>
</tr>
<tr>
<td></td>
<td>[0.60]</td>
<td>[0.00]</td>
<td>[0.92]</td>
<td>[0.00]</td>
</tr>
<tr>
<td>Decomposition</td>
<td>54.00%</td>
<td>46.89%</td>
<td>-0.15%</td>
<td>99.84%</td>
</tr>
</tbody>
</table>

Panel A: Predictability by dividend yield

<table>
<thead>
<tr>
<th>Horizon</th>
<th>“Dinosaurs”</th>
<th></th>
<th>“Non-dinosaurs”</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>eg</td>
<td>r</td>
<td>eg</td>
<td>r</td>
<td>eg</td>
</tr>
<tr>
<td>1</td>
<td>-0.804</td>
<td>0.051</td>
<td>-0.716</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td>[0.00]</td>
<td>[0.01]</td>
<td>[0.02]</td>
<td>[0.11]</td>
</tr>
<tr>
<td>3</td>
<td>-0.630</td>
<td>0.053</td>
<td>-0.840</td>
<td>0.040</td>
</tr>
<tr>
<td></td>
<td>[0.04]</td>
<td>[0.09]</td>
<td>[0.01]</td>
<td>[0.17]</td>
</tr>
<tr>
<td>5</td>
<td>-0.888</td>
<td>0.096</td>
<td>-0.846</td>
<td>0.033</td>
</tr>
<tr>
<td></td>
<td>[0.00]</td>
<td>[0.03]</td>
<td>[0.03]</td>
<td>[0.27]</td>
</tr>
<tr>
<td>Decomposition</td>
<td>96.58%</td>
<td>6.07%</td>
<td>97.27%</td>
<td>3.86%</td>
</tr>
</tbody>
</table>