Market Capitalization, Corporate Payouts, and Expected Returns[*]

Bong-Gyu Jang[a], Bong-Soo Lee[b,c], and Hyun-Tak Lee[d,†]

(First draft: March 1, 2018; This draft: June 10, 2018)

[a] POSTECH, Republic of Korea

[b] Florida State University, United States

[c] UNIST, Republic of Korea

[d] Risk Management Institute, National University of Singapore, Singapore

Abstract

This paper examines the present value framework that links market capitalization to non-dividend cash flows (i.e., share repurchases and issuances) beyond the conventional price–dividend relationship. We show that total (dividend plus non-dividend) cash flows can account for a large fraction of both price and return variations by using tests of cross-equation restrictions on vector autoregression (VAR), impulse response functions, and variance decomposition of unexpected returns. These results come from the fact that total cash flow shocks exhibit a strong negative correlation with payout yield shocks. Non-dividend cash flows help to reconcile our findings with previous literature that prices move substantially relative to dividends.

JEL Classifications: C12, C32, G12

Key words: Asset pricing; VAR; Impulse response function; Variance decomposition

* We thank Sungjune Pyun, Hyeongsop Shim, Eun Jung Lee, Hyeng Keun Koo, Young Myung Ko, Soo Young Chang, Chi-Hyuck Jun, and participants at the 2015 KFA joint conference and 2017 KFMA conference for helpful comments and constructive criticism.

† Corresponding author.

E-mail Address: bonggyujang@postech.ac.kr (B.-G. Jang),
blee2@cob.fsu.edu (B.-S. Lee),
lht1107@gmail.com or rmileeh@nus.edu.sg (H.-T. Lee).

1. Introduction

Asset prices should equal expected discounted cash flows (Cochrane, 2011); prices today must reflect either expectation of future cash flows or expectation of future returns, or both. This central tenet has attracted a large body of literature to explore how much of the two parts account for price variation. For example, Cochrane (2008) shows that the variation in market price–dividend ratios corresponds *entirely* to the variation in expected returns. Empirically, the observation that price–dividend ratios fluctuate over time gives simple evidence of time-varying expected returns; therefore, returns are predictable by financial ratios such as price–dividend ratios and price–earnings ratios, especially over long horizons.

Despite abundant empirical evidence, return predictability is still one of the actively ongoing debates: "Why do prices move *so much*?" For example, Cochrane (2011) shows that the long-run return regression coefficient is too large, implying that even a small predictable change in expected returns can lead to large price variation. In contrast, the positive dividend growth forecast is statistically insignificant and even seems to move in the opposite way: high dividend yields should signal *low* (i.e., negative) future dividend growth. The rationale behind these results comes from the fact that price–dividend ratios are very persistent (Fama, 2014) along with the lack of dividend growth forecastability (Cochrane, 2008). Behaviorists often refer to such evidence as a strong indication for speculative bubbles since prices today seem to have nothing to do with future dividend streams (Shiller, 2014). In sum, dividends might have lost their fundamental role in explaining stock market fluctuations.

Motivated by a set of empirical studies mentioned above, we revisit the ongoing debate by advancing the conventional price–dividend relationship toward a new present value framework tied to total cash flows: dividend plus repurchase minus issuance of equity. Intuitively, firms are willing to convey their long-term potential to a financial market effectively through various payout forms since such payout policies can improve firms' market valuation. Therefore, investors can demand a premium for bearing the risk of policy shifts. Our work centers on an empirical underpinning for the comprehensive relationship between asset prices and payout activities.

Our main findings are as follows.

First, about two-thirds of current price variation arises from subsequent changes in expected total cash flows, and the remaining one-third results from variation in expected returns. To underscore these points, we carry out tests of cross-equation restrictions on VAR by restricting long-run expected returns not to move over time using a *constant-expected-real-return model*. The VAR approach shows that the log total payout ratio, defined as a linear combination of log prices and log total cash flows, can predict more total payout growth than returns. Such predictive cash flow growth thus displays by far *transitory* movements largely associated with long-run mean reversion inherent in total payout ratios.

We emphasize that even strong cash flow forecastability does *not* necessarily mean that returns are unpredictable. The remaining one-third of price variation can support the idea that returns are indeed predictable—expected returns vary over time but do not move too much. In essence, we must consider which of the two—total payout growth or returns—is *much* more predictable and *not* whether total payout growth (or returns) is forecastable. For example, Cochrane (2008) shows that the lack of dividend growth forecastability points to even stronger evidence than does the presence of return predictability, because return forecastability is the flip side of cash flow predictability in the context of the present value relationship, and vice versa. In this regard, we offer a great deal of present value structure to examine such relative significance explicitly, whereas most of the conventional studies (e.g., Boudoukh et al., 2007; Boudoukh et al., 2008) have focused on return predictability alone without the structure.

Second, to the best of our knowledge, our work is the first study to shed light on why price variation is closely linked to cash flow movements. This exposition clarifies that our work goes beyond Larrain and Yogo (2008), who show that prices can be justified by subsequent changes in the expected future total cash flows without scrutinizing the detailed reason. A concrete answer for the tight link is that payout yield shocks (i.e., shocks to total payout ratios) exhibit a *strong negative* correlation with total cash flow shocks (i.e., shocks to total payout growth). In brief, a rise in cash flows lowers the total payout ratios, but also raises the total payout growth.

Concretely, we perform impulse response functions to explain the previous VAR results. By exploiting the strong relationship between payout yields and cash flows, we create a payout yield shock with *no* move in current returns by using the new error identity—a payout yield shock, a total cash flow shock, and a return shock—which are all contemporaneously connected to one another. As a result, our shock identification gives a nice interpretation consistent with the cross-equation-restriction results, implying that most of the payout yield shocks signify expected cash flow news in favor of cash flow forecastability. Therefore, the rest of expected return news justifies why returns are *a little* predictable. In contrast, the convention that a dividend yield shock is nearly expected return news may lead to returns being a little *too* forecastable (Cochrane, 2008).

Third, return variation (or unexpected returns) is attributed to roughly four-fifths of the changing expectation of future cash flows (so-called "news about total cash flows"), whereas changing expectations of future returns (so-called "news about discount rates") account for the remainder. This finding seems surprising because conventional studies have found that a large fraction of return variation comes from news about discount rates (Campbell, 1991; Campbell and

Ammer, 1993). This convention also leaves an unanswered question: “Which economic forces create such persistent return variation?” The answer is that the expectation of non-dividend cash flows reflected in prices today exerts the underlying economic forces. To support this, we use Chave’s (2009) variance decomposition of unexpected returns.

Further, we proceed to decompose news about total cash flows into a total cash flow shock and news about future cash flows, and find a new result contrary to the conventional result. Most importantly, we find that the total cash flow shock explains roughly half of the return variance, implying that equity returns respond *largely* to current cash flow announcements. This finding also underlies that payout yield and total cash flow shocks are strongly correlated. In this sense, the convention that dividend yield and dividend shocks are *not* strongly correlated suggests that a dividend shock does not nearly account for the return variance.

Consequently, it turns out that non-dividend cash flows are essentially business-cycle indicators, which help to reconcile our results with the conventional price–dividend implications: “How can non-dividend cash flows *attenuate* the discount part (or *strengthen* the cash flow part)?” In short, this transition can arise because total payout ratios are much less persistent than price–dividend ratios in favor of cash flow forecastability.

Our analysis reveals that the first-order autocorrelation of the total payout ratios is about 0.68, whereas that of the price–dividend ratios is about 0.93. Noteworthy are autocorrelations that allow short-run forecasting power to build up over time toward long-run forecasting capability because they determine how fast the ratios revert over time. As for the price–dividend ratios, the mean-reverting speed is so slow (i.e., persistent) that even a small variation in expected returns can cumulate over long horizons, amounting to a big price variation due to the lack of dividend growth forecastability. In contrast, relatively small autocorrelations (e.g., 0.68 for the total payout ratio) help to revert rapidly, so a small return variation does not naturally lead to a huge price variation because of strong cash flow forecastability.

Therefore, even reliable short-horizon predictability does not ensure strong evidence for long-term forecastability. For example, Boudoukh et al. (2007) show that various payout yields including repurchases and issuances yield higher R^2s for forecasting returns than do the conventional dividend yields. Cochrane (2008) refers to this finding as clear evidence of long-horizon return predictability because the payout yields are more stationary than the dividend yields. The present value logic says that his contention might be correct, only when cash flow growth is still *less* predictable than returns.

The rest of the paper is organized as follows. In Section 1.1, we review the literature on our work. In Section 2, we derive our present value framework. In Section 3, we present the data used in this paper. In Section 4, we introduce the VAR approach, impulse response functions, and variance decomposition and then study our main findings. In Section 5, we present the concluding remarks.

1.1. Literature review

Our motivation comes from some of the empirical literature (Wright, 2004; Robertson and Wright, 2006; Boudoukh et al., 2007). For example, Boudoukh et al. (2007) call the net payout yields compatible with our total payout ratios. Wright (2004) and Robertson and Wright (2006) define our total payout ratios as cash flow yields. All of them show commonly that total payout ratios have strong and stable predictive abilities for future returns, but do not consider cash flow forecastability (i.e., the other side of return predictability) with great care.

These conventional studies call into question the role of dividends in driving asset prices. They point out that measurement errors exist due to the narrow definition of cash flows as dividends only. In addition to paying out a substantial amount of their earnings as cash dividends, firms occasionally reduce the number of shares outstanding through repurchases and rely on external financing through issuances of equity.

Apart from the measurement errors, the statistical outcomes of return predictability might be spurious because of using the highly persistent price–dividend ratios as a regressor since this spurious relation may lead to a bias in the estimated regression coefficients (Stambaugh, 1999). Moreover, the price–dividend ratios have inconsistent results for in-sample and out-of-sample performance (Goyal and Welch, 2003; Welch and Goyal, 2008). Therefore, Allen and Michaely (2003) advocate that non-dividend cash flows should be incorporated to redefine all cash distributions to shareholders. As such, our present value framework provides a unified empirical framework that can be related to a capital structure in the area of corporate finance.

Another debate over whether corporate payouts are relevant to price valuation is of central importance. In short, Gordon (1959) claims that cash flows such as dividends have an important effect on prices. However, Miller and Modigliani (1961) show that the critical factor may be corporate earnings power and investment policy, *not* the cash distribution method.

Specifically, this payout irrelevance proposition assumes perfect capital markets, rational behavior, and fixed investment opportunities. The first assumption implies that the amount of debt does not affect prices, so new debt issuances do not generate any additional value (i.e., 100% equity financing). The second assumption is widely used in financial economics, meaning that people would rather have more assets than less. Of paramount importance is the third assumption that expected returns are constant over time.

At a technical level, our present value framework relaxes the third assumption to allow for time-varying expected returns. We highlight that even weak evidence for time variation in expected returns could demonstrate vast economic significance of payout relevance to price valuation. For example, suppose that cash distributions perfectly offset each other (Miller and Modigliani, 1961). In this case, they no longer affect market valuation, so shareholders would not need to be rewarded, thereby no time variation in expected returns afterward.

Among non-dividend cash flows, open-market repurchases have gradually become a popular form of payout, and have been often referred to as a substitute for cash dividends (Fama and French, 2001; Grullon and Michaely, 2002; Brav et al., 2005; Dittmar, 2008; Skinner, 2008). The introduction of Rule 10b-18 in the US allows firms to buy back some portions of their shares on the open market, which functions as the safe harbor of repurchasing their prices under some legal conditions.

Repurchases have attractive properties for both shareholders and firms against dividends. First, repurchases taxed as capital gains incur a lower tax cost than do dividends taxed as ordinary income (Bagwell and Shoven, 1989). Second, repurchases compensate for employee stock ownership and are preferable for mergers (Fama and French, 2001). Third, repurchases can raise undervalued prices (Haruvy et al., 2014; Dittmar and Field, 2015).

2. Present value framework

We first define real gross return $1+R_{t+1}$, held from the beginning of time t to the end of time $t+1$ as

$$1+R_{t+1}=\frac{P_{t+1}+DPS_{t+1}}{P_t}, \tag{1}$$

where P_t denotes a real (inflation-adjusted) price of a share in a firm at the end of time t, and DPS_t denotes a real dividend per share during period t. Following Miller and Modigliani (1961), we restate the return on one share of equity in (1) by multiplying by the number of shares outstanding N_t measured at the end of time t:

$$1 + R_{t+1} = \frac{V_{t+1} + D_{t+1} - (I_{t+1} - S_{t+1})}{V_t}, \tag{2}$$

where $V_t = P_t \cdot N_t$ is a market price of the firm at the end of time t; $V_{t+1} = P_{t+1} \cdot N_{t+1}$ is an *ex-dividend* (or *ex-payout*) price at the end of time $t+1$; $D_t = DPS_t \cdot N_t$ is a total dividend paid to shareholders during period t; and I_t and S_t are respectively the total amounts of share issuances and repurchases during period t such that $I_t - S_t = P_t \cdot (N_t - N_{t-1})$. Without any other payout method, (2) is algebraically equivalent to (1). Henceforth, lowercase letters denote logs of variables, for example, $v_t = \log(V_t)$, $d_t = \log(D_t)$, $s_t = \log(S_t)$, and $i_t = \log(I_t)$.

We then take natural logarithms on both sides of (2):

$$\begin{aligned} r_{t+1} &= \log\left(\frac{V_{t+1} + D_{t+1} + S_{t+1} - I_{t+1}}{V_t}\right) = \log\left(1 + \frac{D_{t+1}}{V_{t+1}} + \frac{S_{t+1}}{V_{t+1}} - \frac{I_{t+1}}{V_{t+1}}\right) + v_{t+1} - v_t \\ &= \log(1 + \exp(d_{t+1} - v_{t+1}) + \exp(s_{t+1} - v_{t+1}) - \exp(i_{t+1} - v_{t+1})) + v_{t+1} - v_t, \end{aligned} \tag{3}$$

where $r_t = \log(1 + R_t)$ denotes a continuously compounded real return on equity. For economic tractability, we attempt to linearize the three ratios on the right-hand side (RHS): a log dividend-price ratio ($d_{t+1} - v_{t+1} \equiv dv_{t+1}$), a log repurchase-price ratio ($s_{t+1} - v_{t+1} \equiv sv_{t+1}$), and a log issuance-price ratio ($i_{t+1} - v_{t+1} \equiv iv_{t+1}$).

Next, applying a first-order Taylor expansion to the non-linear log term in (3) yields

$$r_{t+1} \approx k_3 + \rho_3\left[v_{t+1} + \left(\frac{1}{\rho_1} - 1\right)d_{t+1} + \left(\frac{1}{\rho_2} - \frac{1}{\rho_1}\right)s_{t+1} - \left(\frac{1}{\rho_2} - \frac{1}{\rho_3}\right)i_{t+1}\right] - v_t, \tag{4}$$

where ρ_1, ρ_2, and ρ_3 are log-linear coefficients, and k_3 is a log-linear intercept. The detailed derivation is given in Appendix A. At a mechanical level, log-linear approximation (4) represents the log of total payouts given by $\log(V_{t+1} + D_{t+1} + S_{t+1} - I_{t+1})$ as a weighted average of log price and log payout variables such that their weights add up to one:

$$\rho_3 \left[1 + \left(\frac{1}{\rho_1} - 1\right) - \left(\frac{1}{\rho_2} - \frac{1}{\rho_1}\right) - \left(\frac{1}{\rho_2} - \frac{1}{\rho_3}\right)\right] = 1.$$

We assume that log price and log payouts on the RHS of (4) are of integrated order one I(1) (i.e., non-stationary) and thus rearrange approximate identity (4) as

$$r_{t+1} \approx k_3 + \rho_3 \delta_{t+1} + \beta_1 \Delta d_{t+1} + \beta_2 \Delta s_{t+1} - \beta_3 \Delta i_{t+1} - \delta_t, \tag{5}$$

where $\beta_1 = \frac{\rho_3(1-\rho_1)}{\rho_1(1-\rho_3)} = \frac{\bar{D}}{\bar{D}+\bar{S}-\bar{I}}$, $\beta_2 = \frac{\rho_3(\rho_1-\rho_2)}{\rho_2\rho_1(1-\rho_3)} = \frac{\bar{S}}{\bar{D}+\bar{S}-\bar{I}}$, $\beta_3 = \frac{\rho_3-\rho_2}{\rho_2(1-\rho_3)} = \frac{\bar{I}}{\bar{D}+\bar{S}-\bar{I}}$; Δ denotes the first difference operator; Δd_t is dividend growth during period t; Δs_t and Δi_t are respectively repurchase growth and issuance growth during period t; and $\delta_t \equiv v_t - \beta_1 d_t - \beta_2 s_t + \beta_3 i_t$ is the log *total payout ratio* measured at the end of time t. Here, the three beta coefficients—a historical proportion of each payout to total cash flows—obey the following relationship:

Proposition 1: Present value restriction

The payout proportion weights β_1, β_2, and β_3 should add up to one:

$$\beta_1 + \beta_2 - \beta_3 = 1.$$

Proof: It is straightforward to substitute $\beta_1 = \frac{\bar{D}}{\bar{D}+\bar{S}-\bar{I}}$, $\beta_2 = \frac{\bar{S}}{\bar{D}+\bar{S}-\bar{I}}$, and $\beta_3 = \frac{\bar{I}}{\bar{D}+\bar{S}-\bar{I}}$ into the restriction.

Suppose that Proposition 1 holds with data. Then, we can obtain the ex-ante present value identity without the constant term $k_3/(1-\rho_3)$:

$$\delta_t \approx E\left[\sum_{j=0}^{\infty}(\rho_3)^j \beta_1 \Delta d_{t+1+j} + \beta_2 \Delta s_{t+1+j} - \beta_3 \Delta i_{t+1+j} - r_{t+1+j} | \mathcal{F}_t\right], \qquad (6)$$

where $E[\cdot\,|\mathcal{F}_t]$ is the expectation conditioned on an information set $\mathcal{F}_t = [\delta_t, \Delta d_t, \Delta s_t, \Delta i_t, r_t]$ available at the end of time t. In present value identity (6), log prices are *cointegrated* with log dividends, repurchases, and issuances, given that the four ex-ante RHS variables are stationary; total payout ratio δ_t is the optimal linear forecast of changes in dividends, repurchases, and issuances, and expected returns over the indefinite future. For example, high prices relative to total payouts (i.e., $v_t > \beta_1 d_t + \beta_2 s_t - \beta_3 i_t$) indicate high expected dividend growth, high expected repurchase growth, low expected issuance growth, low expected return, or their combination. Proposition 1 serves to link our framework to the existing literature (Miller and Modigliani, 1961; Campbell and Shiller, 1988; Grullon and Michaely, 2002; Larrain and Yogo, 2008). We want to refer interested readers to Appendix B.

3. Data

Following Boudoukh et al. (2007), we use U.S. nonfinancial firms listed on the New York Stock Exchange (NYSE), the American Stock Exchange (AMEX), and the Nasdaq Stock Market (NASDAQ) at the intersection of COMPUSTAT fundamental annual and the Center for Research in Security Prices (CRSP) monthly profile. The linkage utilizes the CRSP/COMPUSTAT Merged Database (CCM) on available linking information. We select a sample period from calendar years

1971 to 2014 since share repurchase data are only available after 1971 onward. Now, we specify how to construct the data.

First, each firm with superscript i should have available payout information at time t from COMPUSTAT for its fiscal year ending in calendar year $t-1$ for the following variables:

- Dividend D_t^i: Common/Ordinary Equity-Total (DVC).
- Repurchase S_t^i: Purchase of Common and Preferred Stock (PRSTKC).
- Issuance I_t^i: Sale of Common and Preferred Stock (SSTK).[1]

We then aggregate these payout variables into millions of dollars as $D_t = \sum_i D_t^i$, $S_t = \sum_i S_t^i$, and $I_t = \sum_i I_t^i$ to constitute supposedly the market portfolio. Among them, it is clear that repurchases have gradually gained popularity and are most sensitive to the business cycle (Figure 1).

[INSERT FIGURE 1 HERE]

Second, each firm should have financial data from CRSP:

- Market price V_t^i: market cap. (PRC × SHROUT) at the end of June of calendar year t.[2]
- Stock return r_{t+1}^i: ordinary stocks' continuously compounded return ($\log(1+\mathrm{RET})$) from July of calendar year t to June of calendar year $t+1$.[3]

[1] Some studies include any decreases and increases in the value of the net number of preferred stocks outstanding (PSTKRV) as the proxies of repurchases and issuances, respectively. In this paper, we do not include PSTKRV because it decreases the total number of firms and sample period on available information.
[2] Share prices (PRC) and the number of shares outstanding (SHROUT) in CRSP are unadjusted variables for splits. For this reason, we use Cumulative Factor to Adjust Prices (CFACPR) and Cumulative Factor to Adjust Shares/Vol (CFACSHR) to obtain the adjusted market capitalization.
[3] We exclude all certificate, ADRs, SBIs, REITs, and *etc*, which are not within share codes of 10 and 11.

This construction ensures that fiscal-year-end payout information (i.e., D_t, S_t, and I_t) should be acknowledged prior to the realized return. Subsequently, we aggregate V_t^i as $V_t = \sum_i V_t^i$, and calculate the CRSP-portfolio value-weighted real return as r_t from July of calendar year $t-1$ to June of calendar year t and excess return $er_t = r_t - r_{f,t}$, where $r_{f,t}$ is the real return on the three-month Treasury bills at the end of June of calendar year t. Upon data construction, we deflate all monthly data by using the (not seasonally adjusted) Consumer Price Index level (CPIind; December of 1972 = 100) and the Rate of Change for all urban consumers (CPIret) for each firm at a fiscal-year-end month from the CRSP Treasury and Inflation Index Table.

Here, it is important to note that we collect *actual* dividend payments retrieved from the intersection of COMPUSTAT fundamental annual and CRSP monthly profile. In contrast, Cochrane (2008) obtains CRSP dividend series from the CRSP returns, which captures *all* payments including the actual dividends, cash mergers, liquidation, and even the repurchase amount. Notwithstanding the different dividend measures, it will become evident that our results implied by price-dividend ratio $\eta_t = v_t - d_t$ are qualitatively analogous to those of Cochrane (2008).

Next, we use the data to calculate total payout ratio $\delta_t = v_t - \beta_1 d_t - \beta_2 s_t + \beta_3 i_t$ measured at the end of June of calendar year t through the beta formulas (Proposition 1):

$$\delta_t = v_t - 1.06 \cdot d_t - 0.54 \cdot s_t + 0.60 \cdot i_t, \tag{7}$$

where $\beta_1 = \frac{\bar{D}}{\bar{D}+\bar{S}-\bar{I}} = 1.06$, $\beta_2 = \frac{\bar{S}}{\bar{D}+\bar{S}-\bar{I}} = 0.54$, and $\beta_3 = \frac{\bar{I}}{\bar{D}+\bar{S}-\bar{I}} = 0.60$ as a result of $\rho_1 \approx 0.975$, $\rho_2 \approx 0.962$, and $\rho_3 \approx 0.976$. Table 1 presents summary statistics for the data.

[INSERT TABLE 1 HERE]

The simple insight from the beta estimates in (7) is that *cash distributions do matter for price variation* against Miller and Modigliani (1961): they contend that how to distribute cash flows is *not* relevant to price valuation (see Appendix B). The rationale behind this is quite simple because total payout ratio δ_t and price–dividend ratio $\eta_t = v_t - d_t$ move in a different manner (Figure 2); they also have different statistics (Table 1). If cash distributions were to exactly offset each other so that they did not affect prices at all, δ_t should be obviously the same as η_t.

[INSERT FIGURE 2 HERE]

The resulting payout relevance to price variation can elucidate that the two ratios should embody distinct information about long-horizon expected returns inherent in prices today. For example, the first-order autocorrelation of price–dividend ratio η_t is about 0.93. A half life τ of this fluctuation is about 9.5 years such that $0.93^\tau = 0.5$. This number indicates that the price–dividend ratio is highly persistent, so small but *persistent* expected return variation can lead to huge price variation. Therefore, Fama (2014) claims that the price–dividend ratio is a noisy proxy for expected returns because dividend cash flows do not seem to function as a priced factor any longer (Cochrane, 2008). In contrast, total payout ratio δ_t just exhibits about 1.8-year half life τ. Hence, small but *less persistent* return variation does not necessarily correspond to large price variation in the end, which essentially hinges on cash flow predictability. In the next section, we focus on the joint forecastability issue in detail.

4. Vector autoregression

Going forward, we work with "centered" variables by subtracting sample means from the variables because constant terms have nothing to do with any VAR restrictions. For the sake of clarity of exposition, we proceed to deliver our main ideas using total cash flows throughout this paper. Our primary purpose is to compare the economic significance between total cash flows and returns through the extended information set beyond dividends and returns, and *not* to scrutinize which individual cash flow is the most dominant source of price variation. Alternatively, we try to examine the potential effect of non-dividend cash flows by analyzing how much dividend cash flows account for price variation and then compare the answer with our results.

4.1. Cross-equation restriction on VAR

4.1.1. Null hypothesis

Let us start with the *constant-expected-real-return model* of Campbell and Shiller (1988) as

$$\delta_t \approx E\left[\sum_{j=0}^{\infty}(\rho_3)^j \left[\Delta\phi_{t+1+j} - r_{t+1+j}\right]\middle|\mathcal{F}_t\right] = E\left[\sum_{j=0}^{\infty}(\rho_3)^j \,\Delta\phi_{t+1+j}\middle|\mathcal{F}_t\right] \equiv \delta_t', \qquad (8)$$

where a single variable, $\phi_t \equiv \beta_1 d_t + \beta_2 s_t - \beta_3 i_t$, denotes total payouts (cash flows), and δ_t' is an unrestricted VAR forecast implied by economic model (8).

To formulate model (8), we first define an endogenous vector as $z_t = [\delta_t, \Delta\phi_t]'$ measured at the end of time t and then represent a first-order VAR:

$$z_{t+1} = A \cdot z_t + \varepsilon_{t+1},$$

where A is the companion matrix, and ε_t is the residual vector. Here, we assume that the endogenous vector represents the state of the economy so that shareholders use them to forecast future market condition. To simplify the exposition, we proceed with VAR(1) in the rest of the analysis, but high-order VAR systems can be stacked into a first-order VAR system as well.

Next, a multi-period forecasting formula $E\left[z_{t+1+j}\middle|\mathcal{F}_t\right] = A^j z_t$ leads to non-linear restrictions on the VAR:

$$\delta_t = e_1' z_t = \sum_{j=0}^{\infty} (\rho_3)^j \, e_2' A^{j+1} z_t = e_2' A[I - \rho_3 A]^{-1} z_t \equiv \delta_t', \tag{9}$$

where $e_1' = [1{,}0]'$ and $e_2' = [0{,}1]'$ are used to pick up δ_t and $\Delta\phi_t$ in z_t, and I is the two-by-two identity matrix.

Hypothesis 1: $\mathbf{H_0}$: $\boldsymbol{\delta_t' = \delta_t}$

A long-run expected-return forecast $E\left[\sum_{j=0}^{\infty} (\rho_3)^j \, r_{t+1+j}\middle|\mathcal{F}_t\right]$ *is constant over time.*

If Hypothesis 1 is correct, δ_t' should put a unit weight on δ_t and a zero weight on $\Delta\phi_t$, implying that δ_t is the optimal forecast of expected future cash flows.

To test Hypothesis 1, we present another two diagnostic statistics as follows. First, a non-linear Wald test statistic of Campbell and Shiller (1988) is

$$\lambda'(\partial\lambda/\partial\gamma\,'\Sigma\,\partial\lambda/\partial\gamma)^{-1}\lambda,$$

where $\lambda = e_1' - e_2' A[I - \rho_3 A]^{-1}$ is defined as the vector of the deviations; γ is the vector of VAR estimates; $\partial\lambda/\partial\gamma$ is numerically estimated as partial derivatives of λ with respect to each component in γ; and Σ is the estimated variance-covariance matrix. The Wald statistic follows a chi-square (χ^2) distribution with n degrees of freedom, where n is the number of endogenous variables in z_t: $n = 2$ in our case. Second, the ratio of standard deviations is $\sigma(\delta_t')/\sigma(\delta_t)$. The standard errors for the two statistics are numerically computed though the standard delta method.

4.1.2. Empirical results

Panel A of Table 2 reports the VAR estimates for the constant-expected-real-return model. Specifically, we estimate companion matrix A through GMM corrected standard errors for heteroscedasticity.

[INSERT TABLE 2]

The VAR estimates show that total payout ratio δ_t provides weak forecastable evidence for total payout growth $\Delta\phi_{t+1}$ with a point estimate of 0.205 and a R^2 of 9.13% (row 2): the significance level is about 6.49%. Even the unclear statistical significance seems surprising because the conventional price–dividend ratios *little* forecast dividend growth with R^2 of 0% (Cochrane, 2008). Now, we turn to investigate how such short-term evidence builds up with horizons toward long-term predictability.

The key implication is that *variation in the total payout ratios comes mostly from predictable changes in total payout growth*. This stands in sharp contrast with the dividend yield implication (Cochrane, 2008); all variation in the price–dividend ratios corresponds to changes in expected returns. In brief, we provide three pieces of evidence in Panel B of Table 2.

In the first column of Panel B, VAR forecast $\delta_t' = e_2' A[I - \rho_3 A]^{-1} z_t$ calculates that the weight of δ_t' on δ_t (0.627) seems to be far from one, although that of δ_t' on $\Delta\phi_t$ (-0.024) is close to zero. The estimated weight vector $[0.627, -0.024]'$ indicates that the rational expectation of future total cash flows largely accounts for roughly two-thirds of price variation, as also evidenced by $\sigma(\delta_t')/\sigma(\delta_t) = 63.3\%$ (column 2). Accordingly, the remaining one-third variation characterized by the vector $[0.373, 0.024]'$ explains how much expected returns move over long horizons. The third column of $\mathrm{H}_0\colon \delta_t' = \delta_t$ also shows the failure to reject Hypothesis 1 at the conventional levels: the *p*-value is about 0.796. In sum, all diagnostic results support the idea that expected returns are time varying (the *p*-value is *not* 100%) but do not move too much.

To take the understanding to the next level, we need to put the relative economic significance into perspective within our information set $\mathcal{F}_t$ as to which of the return or total payout growth is more forecastable, as highlighted in Cochrane (2008). With such a perspective in mind, the right interpretation for Hypothesis 1 is that the presence of cash flow forecastability provides much more significant evidence than does that of return forecastability.

[INSERT TABLE 3 HERE]

As for price–dividend ratio η_t, the lack of dividend growth forecastability is evident in favor of return predictability (Table 3); this conclusion accords well with that of Cochrane (2008). First, dividend growth Δd_{t+1} is not forecastable by η_t: the point estimate is 0.036 with a R^2 of 3.62%, and the significance level is above 25% (row 2, Panel A). Second, subsequent changes in expected dividends just account for about one-third of price variation: the weight of η_t' on η_t is 0.386 (column 1, Panel B) alongside $\sigma(\eta_t')/\sigma(\eta_t) = 38.6\%$ (column 2, Panel B). The rest of the proportion thus points to the time variability of long-term expected returns.

As opposed to the previous evidences, the statistical significance is not clear. We fail to reject Hypothesis 1 for $\mathrm{H}_0\colon \eta_t' = \eta_t$ at the conventional levels (column 3, Panel B): the *p*-value is about 0.545. This controversial result, however, supports that of Boudough et al. (2007), who show that the statistical significance of η_t disappears when including the recent periods. This is also closely associated with the poor out-of-sample performance in the 1990s (Subsection 1.1).

[INSERT FIGURE 3 HERE]

In sum, the price fluctuations implied by the two ratios (Figure 3) give an essential economic message that *non-dividend cash flows are a primary source of price variation*.[4] Clearly, such a wedge takes place *largely* because of the subsequent effect of non-dividend cash flows: $\hat{\beta}_1 \neq 1$. Our VAR results also suggest that the unsolved question of why prices overreact to a *little* return change is in fact an "illusion" caused by quite the small conventional information set with dividends. Therefore, the discount part of prices can mask the non-dividend driving force for a while, which is essentially transitory enough to be predictable. In the next section, we go into the detail of why the total payout and price–dividend ratios yield these different implications.

4.2. Impulse response function

4.2.1. Empirical method

In this section, we display impulse response functions to provide the economic underpinning of the cross-equation-restriction results. Doing so can clarify which revision of expectations of future variables coincides with each shock (Cochrane, 2011). We first utilize total payout ratio δ_t and total payout growth $\Delta\phi_t$ to infer *implied* return r_t through the bivariate VAR (Subsection 4.1.1):

$$z_{t+1} = A \cdot z_t + \varepsilon_{t+1},$$

where $z_t = [\delta_t, \Delta\phi_t]'$ and $\varepsilon_t = [\varepsilon_t^{\delta}, \varepsilon_t^{\phi}]'$.

The VAR allows implied return shock ε^r to have the following form:

[4] We notice that the constant-expected-excess return model of Campbell and Shiller (1988) delivers similar results because the risk-free rate is not variable enough to affect price variation.

$$\varepsilon_{t+1}^{r} = r_{t+1} - E_t r_{t+1} = (e_1'\rho_3 + e_2')\varepsilon_{t+1} = \rho_3 \varepsilon_{t+1}^{\delta} + \varepsilon_{t+1}^{\phi}, \tag{10}$$

where ε^{δ} denotes a payout yield shock, ε^{ϕ} denotes a total payout (cash flow) shock, and ε^{r} denotes a return shock. Error identity (10) is simply obtained from the present value identity:

$$r_{t+1} = \rho_3 \delta_{t+1} + \Delta\phi_{t+1} - \delta_t = [e_1'\rho_3 A + e_2' A - e_1' I] \cdot z_t + \varepsilon_{t+1}^{r}, \tag{11}$$
$$E_t r_{t+1} = E_t[\rho_3 \delta_{t+1} + \Delta\phi_{t+1}] - \delta_t = [e_1'\rho_3 A + e_2' A - e_1' I] \cdot z_t,$$

where $\rho_3 E_t \delta_{t+1} = e_1'\rho_3 A z_t$, $E_t \Delta\phi_{t+1} = e_2' A z_t$, and $\delta_t = E_t \delta_t = e_1' z_t$.

We emphasize that the three contemporaneous shocks obey error identity (10), given that the implied return is governed by dynamic accounting identity (11); all of them are implied by the bivariate VAR. Note that this bivariate VAR has the similar form to Cochrane's (2011) multivariate regressions using dividend yield, and Lettau and Ludvigson's (2005) consumption to wealth ratio (*cay*) as regressors.

As an example, let the companion matrix A in Panel A of Table 2 be

$$A = \begin{bmatrix} b_\delta & c_\delta \\ b_\phi & c_\phi \end{bmatrix}.$$

The return estimates $[b_r, c_r] = [e_1'\rho_3 A + e_2' A - e_1' I]$ of implied return (11) become $[\rho_3 b_\delta + b_\phi - 1, \rho_3 c_\delta + c_\phi]$. The first term $b_r = \rho_3 b_\delta + b_\phi - 1$ is identical to Cochrane's (2008) identity using the dividend yield as a sole predictor, and the second one $c_r = \rho_3 c_\delta + c_\phi$ is the same as Cochrane's (2011) *cay* identity. Therefore, the three variables $[\delta_t, \Delta\phi_t, r_t]$ are redundant; the return data r and error ε^{r} can be inferred from the other two variables $[\delta_t, \Delta\phi_t]$. This approach has an advantage of ensuring identity without resorting to the return dynamics. In sum, our first-order VAR has the following form:

$$\begin{bmatrix} \delta_{t+1} \\ \Delta\phi_{t+1} \\ r_{t+1} \end{bmatrix} = \begin{bmatrix} b_\delta & c_\delta & 0 \\ b_\phi & c_\phi & 0 \\ b_r & c_r & 0 \end{bmatrix} \cdot \begin{bmatrix} \delta_t \\ \Delta\phi_t \\ r_t \end{bmatrix} + \begin{bmatrix} \varepsilon_{t+1}^{\delta} \\ \varepsilon_{t+1}^{\phi} \\ \varepsilon_{t+1}^{r} \end{bmatrix}.$$

Similarly, we can infer a change in prices as

$$\Delta v_{t+1} = (v_{t+1} - \phi_{t+1}) - (v_t - \phi_t) + \Delta\phi_{t+1} = [b_v, c_v] \cdot z_t + \varepsilon^v,$$

where $\delta_t = v_t - \phi_t = e_1' z_t$, $\delta_{t+1} = e_1' A z_t + \varepsilon_{t+1}^{\delta}$, $\Delta\phi_{t+1} = e_2' A z_t + \varepsilon_{t+1}^{\phi}$, $b_v = b_\delta + b_\phi - 1$, and $c_v = c_\delta + c_\phi$; $\varepsilon^v = \varepsilon^\delta + \varepsilon^\phi$ denotes a "price shock." The price estimates $[b_v, c_v]$ and a price shock ε^v are respectively similar to the return estimates $[b_r, c_r]$ and return shock $\varepsilon^r = \rho_3 \varepsilon^\delta + \varepsilon^\phi$. The only discrepancy arises from whether to discount b_δ, c_δ, and ε^δ by ρ_3.

To see this intuitively, we rewrite the present value form as

$$r_{t+1} = \rho_3 v_{t+1} + (1 - \rho_3)\phi_{t+1} - v_t = \rho_3 \delta_{t+1} + \Delta\phi_{t+1} - \delta_t.$$

Remind that the discount factor, $\rho_3 \approx 0.976$, is used as a weight of log prices, so the remaining weight $(1 - \rho_3)$ is for log total cash flows. The price process can thus be formulated by assigning $\rho_3 = 1$ into log prices with *no* cash flow proportion: $r_{t+1} = \Delta v_{t+1}$. We then specify a cumulative response of prices as $v_t = \sum_{j=1}^{t} \Delta v_j$ and the response of total cash flows as $\phi_t = \sum_{j=1}^{t} \Delta\phi_j$.

Now, we are in the position to identify our primary impulse response surprise. Impulse response functions aim to "spark" a *transitory* surprise today and then see how the expected future path of the variables $[\delta_t, \Delta\phi_t, r_t, v_t, \phi_t]$ is moving forward through time (i.e., the term structures of the variables). Accordingly, it is important to single out which surprise stands for a transitory component of the total payout ratios.

In this regard, we simply exploit our cross-equation-restriction results (Subsection 4.1.2) that the total payout growth is more predictable by the total payout ratios than the returns. Therefore, our purpose is to dissect transitory cash flow movements rather than price fluctuations. In contrast, the price–dividend ratios present the opposite logic in favor of the price fluctuations. Hence, our impulse response surprise should have a different form from a "dividend yield shock given no change in dividends," as formulated by Cochrane (2011).

The present value structure along with the empirical results allows for a "payout yield shock with no change in returns" as $[\varepsilon^{\delta}, \varepsilon^{\phi}, \varepsilon^{r}] = [1, -\rho_3, 0]$:

$$\underbrace{r_{t+1} - E_t r_{t+1}}_{\varepsilon^{r}=0} = \rho_3 \left[\underbrace{\delta_{t+1} - E_t \delta_{t+1}}_{\varepsilon^{\delta}=1}\right] + \left[\underbrace{\Delta\phi_{t+1} - E_t \Delta\phi_{t+1}}_{\varepsilon^{\phi}=-\rho_3}\right]. \tag{12}$$

It is clear that the error identity of $\varepsilon^{r} = \rho_3 \varepsilon^{\delta} + \varepsilon^{\phi}$ does not permit one of the shocks in (12) to independently occur. Return variation must come from either payout yield shock ε^{δ} or total payout shock ε^{ϕ} at the very least. Such a shock identification also points to a slight increase in current prices by $\varepsilon^{p} = \varepsilon^{\delta} + \varepsilon^{\phi} = 1 - \rho_3 \approx 0$.

We choose this simple story since it can convey new insight not covered by Cochrane (2011) dealing with the dividend yield shock with no move in dividend cash flows (i.e., $[\varepsilon^{\eta}, \varepsilon^{d}, \varepsilon^{r}] = [1, 0, \rho_1]$). In fact, adding lags to the VAR system might deliver another essential information. Even so, we want to keep up the system going forward in the interest of brevity. To simplify it further, we also run simple univariate regressions on the total payout ratio without c_i terms for $i = \delta, \phi, r$ and find out that the results are qualitatively similar (untabulated results). The similarity shows that b_i terms drive our results onward.

4.2.2. *Error standard deviations and correlations*

We first present error standard deviations on the diagonal and correlations on the off-diagonal for total payout ratio δ_t in Panel A of Table 4. This preliminary analysis is important to select the most plausible economic story for impulse response functions in light of the data. The first three columns report the statistics of ε^{δ}, ε^{ϕ}, and implied ε^{r}, inferred from error identity (10) by using the VAR estimates in Panel A of Table 2.

[INSERT TABLE 4 HERE]

The summary statistics show that the total payout ratio has the largest standard deviation: $\sigma(\varepsilon^{\delta}) \approx 30.5\%$. Sequentially, the total payout growth has about 29% deviation, which is almost twice as volatile as the 15% return deviation: $\sigma(\varepsilon^{\phi}) \approx 29.4\%$ and $\sigma(\varepsilon^{r}) \approx 15.0\%$. As for price–dividend ratio η_t (last three columns), conversely, the return deviation is roughly twice the dividend deviation: $\sigma(\varepsilon^{\eta}) \approx 18.0\%$, $\sigma(\varepsilon^{r}) \approx 15.6\%$, and $\sigma(\varepsilon^{d}) \approx 8.8\%$.

The most important feature of the data is that *total payout shock* ε^{ϕ} *is strongly and negatively correlated with payout yield shock* ε^{δ}: $\mathrm{Corr}(\varepsilon^{\delta}, \varepsilon^{\phi}) \approx -87.3\%$. Meanwhile, payout yield shock ε^{δ} has a relatively small but positive correlation with implied return shock ε^{r}: $\mathrm{Corr}(\varepsilon^{\delta}, \varepsilon^{r}) \approx 14.0\%$. What this matters is that $\mathrm{Corr}(\varepsilon^{\delta}, \varepsilon^{r})$ is of *less* economic significance than $\mathrm{Corr}(\varepsilon^{\delta}, \varepsilon^{\phi})$.

To achieve a better understanding of this point, we restate error identity (10) by multiplying both sides by ε^{δ} and then taking expectations:

$$\mathrm{Cov}[\varepsilon^{\phi}, \varepsilon^{\delta}] = \mathrm{Cov}[\varepsilon^{r}, \varepsilon^{\delta}] - \rho_3 \cdot \sigma^2[\varepsilon^{\delta}].$$

Suppose that total payout and return shocks are uncorrelated: $\text{Cov}[\varepsilon^r, \varepsilon^\delta] = 0$. Then, divide both sides by $\sigma[\varepsilon^\delta]$ and $\sigma[\varepsilon^\phi]$, yielding

$$\text{Corr}[\varepsilon^\phi, \varepsilon^\delta] = -\rho_3 \cdot \sigma[\varepsilon^\delta]/\sigma[\varepsilon^\phi] \approx -101.2\%.$$

In fact, this arbitrary correlation less than -1 is an undefined number, but there is no question that the number is closer to $\text{Corr}(\varepsilon^\delta, \varepsilon^\phi) \approx -87.3\%$ than to $\text{Corr}(\varepsilon^\delta, \varepsilon^r) \approx 14.0\%$. This strong negative correlation means that both price and cash flow movements greatly offset each other; their balance probably emerges as small return variation by (12). In this context, our impulse response shock, $[\varepsilon^\delta, \varepsilon^\phi, \varepsilon^r] = [1, -\rho_3, 0]$, seems quite reasonable to be reflective of the data.

In contrast, the last three columns of Panel A reveal that price–dividend ratio (henceforth, dividend yield) and returns shocks are most highly correlated: $\text{Corr}(\varepsilon^r, \varepsilon^\eta) \approx 87.9\%$. Note our price–dividend ratio $\eta_t = p_t - d_t$ has the exact negative sign of the log dividend yield defined as $d_t - p_t$. The strongest negative correlation of Cochrane (2008) is consistent with our positive correlation. This strong correlation can illustrate that the persistent return variation under debate (Section 1) is a consequence of the weak relationship between price and dividend fluctuations: $\varepsilon^r = \rho_1 \varepsilon^\eta + \varepsilon^d$.

For robustness tests, we also use the CRSP portfolio returns (Section 3) in exchange for implied r and ε^r, and then report the results in Panel B of Table 4. The direct computation, not resorting to implied identities (10) and (11), does not guarantee the exact identity in principle. Nevertheless, the error correlations and standard deviations remain qualitatively similar to those reported in Panel A. Therefore, we proceed with implied r and ε^r to save space in the rest of the paper.

4.2.3. Empirical results

The key result for impulse response analysis is that *a payout yield shock with no change in returns generates forward about two-thirds of long-term cash flow expectations and the remaining one-third of long-term return expectations*. This finding is analogous to those of the cross-equation-restriction test reported in Table 2, demonstrating that our shock identification $[\varepsilon^{\delta}, \varepsilon^{\phi}, \varepsilon^{r}] = [1, -\rho_3, 0]$ is reflective of the data well. For ease of grasp, Figure 4 plots the impulse responses to such a shock.

[INSERT FIGURE 4 HERE]

In the first panel, a payout yield rise ($\varepsilon^{\delta} = 1$) with no contemporaneous move in returns ($\varepsilon^{r} = 0$) should come with a dramatic decline in current payout growth ($\varepsilon^{\phi} = -\rho_3$). It then induces the expected payout growth to rise (by $b_{\phi} = 0.205$ in principle), which is subsequently decaying (by $b_{\delta} = 0.698$) through time. In the second panel, the path of total payouts thus falls first ($\varepsilon^{\phi} = -\rho_3$) and then should increase slowly along with the expected growth rise. Moreover, the payout yield rise causes total payout ratio δ_t to revert slowly.

The payout yield rise also implies a decline in expected returns by the present value logic:

$$\delta_t \approx E_t\left[\sum_{j=0}^{\infty}(\rho_3)^j \, \Delta\phi_{t+1+j} - r_{t+1+j}\right],$$

With no current return move ($\varepsilon^{r} = 0$), subsequent expected returns first drop by $[b_r, c_r] = [-0.114, -0.009]$ and in turn grow up (by $b_{\delta} = 0.698$) as shown in the first panel. In the second panel, prices starting off from almost zero ($\varepsilon^{p} = 1 - \rho_3 \approx 0$) therefore tend to decrease at a slow pace in the form of the cumulative expected returns: $v_t = \sum_{j=1}^{t} r_j$.

The highlight of the findings is that total cash flow paths almost appear to be the mirror image of payout yield paths as characterized by the data: $\mathrm{Corr}(\varepsilon^{\delta}, \varepsilon^{\phi}) \approx -87.3\,\%$. This finding demonstrates that a large fraction of payout yield shocks is expected cash flow news and thus has to do with cash flow predictability in essence (Subsection 4.1.2).

[INSERT FIGURE 5 HERE]

Our finding is contrary to the dividend yield implication shown in Cochrane (2011). To see this in brief, Figure 5 plots the responses to the dividend yield shock with no move in dividends: $[\varepsilon^{\eta}, \varepsilon^{d}, \varepsilon^{r}] = [1,0,\rho_1]$. As expected, a large fraction of dividend yield shocks is expected return news according to the correlation feature of return and dividend yield shocks: $\mathrm{Corr}(\varepsilon^{r}, \varepsilon^{\eta}) \approx 87.9\%$ (Subsection 4.2.2). The dividend yield movements thus look similar to price paths (cumulative expected returns); hence, returns are forecastable by the price–dividend ratios.

[INSERT FIGURE 6 HERE]

What if a payout yield shock holding total payout growth constant, modeled by $[\varepsilon^{\eta}, \varepsilon^{\phi}, \varepsilon^{r}] = [1,0,\rho_3]$, happens? In short, the paths of expected cash flows and returns in the first panel of Figure 6 look similar to those in Figure 4. Such a similar nature confirms that a payout yield move is overall an expected cash flow move. The disparity here is that this shock identification *little* reflects the data feature; the payout yield movements are far away from the cash flow movements against $\mathrm{Corr}(\varepsilon^{\delta}, \varepsilon^{\phi}) \approx -87.3\%$. It should be noted that both price and cash flow paths are the outcome of changes in expected returns because of no current change in cash flows by our shock restriction.

4.3. Variance return decomposition

4.3.1. *Empirical method*

Following Campbell (1991) and Campbell and Ammer (1993), we turn to decompose the sources of return variation. Briefly, present value identity (6) leads to

$$(E_{t+1} - E_t) r_{t+1} = (E_{t+1} - E_t)\left[\Delta\phi_{t+1} + \sum_{j=1}^{\infty} (\rho_3)^j \left\{\Delta\phi_{t+1+j} - r_{t+1+j}\right\}\right],$$

$$\varepsilon_{t+1}^{r} = \varepsilon_{t+1}^{\phi} + N_{t+1}^{\phi} - N_{t+1}^{DR}, \tag{13}$$

where return shock ε^r is associated with total payout shock ε^{ϕ}, news about *future* total payouts N^{ϕ}, and news about future discount rates N^{DR}.

Campbell and Ammer (1993) mention that "total" cash flow news N^{CF} ($N^{CF} = \varepsilon^{\phi} + N^{\phi}$) refers to a capital gain, which has a permanent effect on return variation subsequently: $\varepsilon^r \propto N^{CF}$. In contrast, discount rate news refers to a capital loss, which has an offsetting effect on return variation: $\varepsilon^r \propto -N^{DR}$. When return innovations are positive ($\varepsilon^r > 0$), it signals (a) current payout growth must be higher ($\varepsilon^{CF} > 0$), (b) expected future payout growth must be higher ($N^{\phi} > 0$), (c) future discount rates must be lower ($N^{DR} < 0$), or (d) all combination cases.

The VAR form allows us to express the other news as

$$N_{t+1}^{\phi} = (E_{t+1} - E_t)\left[\sum_{j=1}^{\infty} (\rho_3)^j \Delta\phi_{t+1+j}\right] = e_2' \sum_{j=1}^{\infty} (\rho_3)^j A^j \varepsilon_{t+1} = e_2' \rho_3 A [I - \rho_3 A]^{-1} \varepsilon_{t+1},$$

$$N_{t+1}^{CF} = \varepsilon_{t+1}^{\phi} + N_{t+1}^{\phi} = e_2' [I - \rho_3 A]^{-1} \varepsilon_{t+1},$$

$$N_{t+1}^{DR} = N_{t+1}^{CF} - \varepsilon_{t+1}^{r},$$

where such news is a linear combination of payout yield and total payout shocks: $\varepsilon = [\varepsilon^{\delta}, \varepsilon^{\phi}]'$. As for price–dividend ratio η_t, the variance decomposition exactly works in the same way by replacing δ_t and $\Delta\phi_t$ with price–dividend ratio η_t and dividend growth Δd_t.

Concretely, we present Chave's (2009) variance decomposition of unexpected returns as

$$\mathrm{Var}[\varepsilon^r] = \mathrm{Cov}\left[\varepsilon^r, \varepsilon^{\phi}\right] + \mathrm{Cov}\left[\varepsilon^r, N^{\phi}\right] - \mathrm{Cov}[\varepsilon^r, N^{DR}]. \tag{14}$$

One can obtain (14) by multiplying both sides of (10) by ε^r and then taking expectations.

4.3.2. Empirical results

Panel A of Table 5 presents the calculation of the variance decomposition of unexpected returns for total payout ratio δ_t. Note that all covariance terms are normalized by return variance $\mathrm{Var}[\varepsilon^r]$, so the numbers reported below are shares that add up to one.

[INSERT TABLE 5 HERE]

The variance decomposition shows that *total cash flow news* N^{CF} *is the more dominant component of return variation* ε^r *than discount rate news* N^{DR}. Clearly, the first two columns show that about 79% of the variance of unexpected returns comes from the covariance with news about total cash flows: $\mathrm{Cov}\left[\varepsilon^{\phi}, \varepsilon^r\right] + \mathrm{Cov}\left[N^{\phi}, \varepsilon^r\right] \approx 78.5\%$. Accordingly, the remaining 21% of the return variance is attributed to news about discount rates (column 3): $\mathrm{Cov}[N^{DR}, \varepsilon^r] \approx 21.4\%$. The correlations also exhibit the similar tendency: $\mathrm{Corr}[N^{CF}, \varepsilon^r] \approx 76.1\%$ (column 4) and $\mathrm{Corr}[N^{DR}, \varepsilon^r] \approx -30.3\%$ (column 5).

[INSERT FIGURE 7 HERE]

Specifically, total cash flow news N^{CF} tends to catch up to the overall trend of return variation ε^r (Graph A, Figure 7). In contrast, discount rate news N^{DR} may have a second-order effect on ε^r (Graph B, Figure 7). These are essentially a matter of the fact that the total payout ratios are less persistent than the price-dividend ratios alongside cash flow forecastability. This lesser persistence permits the ratios quickly to revert (Figure 4), and such a mean reverting feature in favor of cash flow predictability assigns the bulk of information to N^{CF} and the remainder to N^{DR}.

In particular, return shock ε^r displays one peak and trough each: (a) the peak around 1983 when the Rule 10b-18 was enacted to allow for a share repurchase program legally, and (b) the trough around the 2008 global financial crisis. Here, total cash flow news N^{CF} reaches one additional peak around 1984; Boudoukh et al. (2007) cast doubt on the stationarity of dividend yields by a structural break. We contend that the structural break has to do with repurchase regulation change (a); share repurchases have gained popularity since then.

Price–dividend ratio η_t tells us quite different economic significance (Panel B, Table 5). The first two columns present that the covariance with total dividend news accounts for about 39% of return variance: $\mathrm{Cov}[\varepsilon^r, \varepsilon^d] + \mathrm{Cov}[\varepsilon^r, N^d] \approx 38.5\%$. Hence, the remaining 61% comes from discount rate news (column 3): $\mathrm{Cov}[\varepsilon^r, N^{DR}] \approx 61.5\%$. Such calculations are analogous to those of Campbell (1991) and Campbell and Ammer (1993)—a large portion of the variance of unexpected returns corresponds to changing expectations of future returns N^{DR}. Their correlations also demonstrate that discount rate news outweighs total dividend news a bit (columns 4 and 5): $\mathrm{Corr}[\varepsilon^r, N^{DR}] \approx -86.0\%$ and $\mathrm{Corr}[\varepsilon^r, N^{CF}] \approx 72.7\%$. By the similar logic, all the results boil down to the fact that the price–dividend ratios are highly persistent in favor of return predictability

(Cochrane, 2008). As opposed to Figure 6, discount rate news thus has a first-order effect on ε^r, whereas total dividend news has the second-order effect.

Another important finding is that total payout shock ε^ϕ explains roughly less than half the variance of unexpected returns (column 1, Panel A, Table 5): $\text{Cov}\left[\varepsilon^r, \varepsilon^\phi\right] \approx 45.4\,\%$. The implication is that return shock ε^r reacts *largely* to contemporaneous cash flow news ε^ϕ (Graph A, Figure 8). In contrast, dividend shock ε^d accounts for almost zero share ($\text{Cov}(\varepsilon^d, \varepsilon^r) \approx -0.04\%$) of the return variance (column 1, Panel B), implying that dividend news ε^d have *little* effect on a change in returns (Graph B, Figure 8).

[INSERT FIGURE 8 HERE]

To comprehend these precisely, we revisit error identity (10):

$$\text{Cov}\left[\varepsilon^r, \varepsilon^\phi\right] \approx 45.4\% = \text{Var}\left[\varepsilon^\phi\right] + \rho_3 \cdot \text{Cov}[\varepsilon^\delta, \varepsilon^\phi],$$

$$\text{Cov}[\varepsilon^r, \varepsilon^d] \approx -0.04\% = \text{Var}[\varepsilon^d] + \rho_1 \cdot \text{Cov}[\varepsilon^\eta, \varepsilon^d].$$

In the first equation, the high covariance on the left-hand side (LHS) occurs because the variance of total payout shocks exceeds the negative covariance between shocks to total payouts and payout yields: $\left|\text{Var}\left[\varepsilon^\phi\right]\right| > \rho_3 \cdot \left|\text{Cov}[\varepsilon^\delta, \varepsilon^\phi]\right|$ (Panel A, Table 4). However, the nearly zero covariance on the LHS of the second equation elucidates that the variance of dividend shocks is smaller than the negative covariance: $|\text{Var}[\varepsilon^d]| < \rho_1 \cdot |\text{Cov}[\varepsilon^\eta, \varepsilon^d]|$. As such, one can scrutinize $\text{Cov}[N^{CF}, \varepsilon^r]$ and $\text{Cov}[N^{DR}, \varepsilon^r]$ through error identity (10), but we do not present them here to save space. For robustness tests, we also report the calculations of Campbell's variance decomposition in Appendix C, and find similar results.

5. Concluding remarks

This study presents additional insight that cash flows can explain a large fraction of both price and return variations, which is against the conventional wisdom. The unsolved question of why prices overreact to even small return variation is still the core of various puzzles (e.g., the equity premium puzzle). We fill the gap through robust empirical investigation via our present value framework.

Obviously, non-dividend cash flows are business cycle indicators. Therefore, their transitory fluctuations—essentially for cash flow forecastability—affect mean-reversion in prices substantially. This conclusion can be identifiable from the strong correlation between a payout yield shock and a total cash flow shock. For future work, it would be interesting to seek which individual payout method mainly causes such a secure connection to price movements.

This study hinges on a small number of samples due to the availability of share repurchases: It covers slightly above 40 annual data points from 1971. Besides, one may think that our analysis is subject to sample bias. We contend that such sample bias is not severe though. Figure 1 can support our contention because regardless of any subsamples, the general data feature seems invariant that non-dividend cash flows fluctuate more over time than do dividend cash flows.

Our work can open up a new stage that other payment means (e.g., mergers & acquisitions) could be incorporated to further enhance our understanding of the risk-return relationship. Of course, such a candidate should be transitory enough to attenuate the discount part, sometimes called the excess volatility (Shiller's volatility literature) or bubbles (behavioral finance).

References

Allen, Franklin, and Roni Michaely, 2003, Chapter 7 payout policy, in M. Harris G.M. Constantinides, and R. M. Stulz, eds.: *Handbook of the economics of finance* (Elsevier).

Bagwell, Laurie Simon, and John B. Shoven, 1989, Cash distributions to shareholders, *The Journal of Economic Perspectives* 3, 129-140.

Boudoukh, Jacob, Roni Michaely, Matthew Richardson, and Michael R. Roberts, 2007, On the importance of measuring payout yield: Implications for empirical asset pricing, *The Journal of Finance* 62, 877-915.

Boudoukh, Jacob, Matthew Richardson, and Robert F. Whitelaw, 2008, The myth of long-horizon predictability, *The Review of Financial Studies* 21, 1577-1605.

Brav, Alon, John R. Graham, Campbell R. Harvey, and Roni Michaely, 2005, Payout policy in the 21st century, *Journal of Financial Economics* 77, 483-527.

Campbell, John Y., 1991, A variance decomposition for stock returns, *The Economic Journal* 101, 157-179.

Campbell, John Y., and John Ammer, 1993, What moves the stock and bond markets? A variance decomposition for long-term asset returns, *The Journal of Finance* 48, 3-37.

Campbell, John Y., and Robert J. Shiller, 1987, Cointegration and tests of present value models, *Journal of Political Economy* 95, 1062-1088.

Campbell, John Y., and Robert J. Shiller, 1988, The dividend-price ratio and expectations of future dividends and discount factors, *The Review of Financial Studies* 1, 195-228.

Chaves, Denis B., 2009, What Explains the Variance of Prices and Returns: Time-series vs. Cross-setion, Available at SSRN: https://ssrn.com/abstract=1508433 or http://dx.doi.org/10.2139/ssrn. 1508433.

Cochrane, John H., 2008, The dog that did not bark: A defense of return predictability, *The Review of Financial Studies* 21, 1533-1575.

Cochrane, John H., 2011, Presidential address: Discount rates, *The Journal of Finance* 66, 1047-1108.

Dittmar, Amy K., 2008, Corporate cash policy and how to manage it with stock repurchases, *Journal of Applied Corporate Finance* 20, 22-34.

Dittmar, Amy K., and Laura Casares Field, 2015, Can managers time the market? Evidence using repurchase price data, *Journal of Financial Economics* 115, 261-282.

Fama, Eugene F., 2014, Two pillars of asset pricing, *The American Economic Review* 104, 1467-1485.

Fama, Eugene F., and Kenneth R. French, 2001, Disappearing dividends: Changing firm characteristics or lower propensity to pay?, *Journal of Financial Economics* 60, 3-43.

Gordon, M. J., 1959, Dividends, earnings, and stock prices, *The Review of Economics and Statistics* 41, 99-105.

Goyal, Amit, and Ivo Welch, 2003, Predicting the equity premium with dividend ratios, *Management Science* 49, 639-654.

Grullon, Gustavo, and Roni Michaely, 2002, Dividends, share repurchases, and the substitution hypothesis, *The Journal of Finance* 57, 1649-1684.

Haruvy, Ernan, Charles N. Noussair, and Owen Powell, 2014, The impact of asset repurchases and issues in an experimental market, *Review of Finance* 18, 681-713.

Larrain, Borja, and Motohiro Yogo, 2008, Does firm value move too much to be justified by subsequent changes in cash flow?, *Journal of Financial Economics* 87, 200-226.

Lettau, Martin, and Sydney C. Ludvigson, 2005, Expected returns and expected dividend growth, *Journal of Fianncial Economics* 76, 583-626.

Miller, Merton, H., and Franco Modigliani, 1961, Dividend policy, growth, and the valuation of shares, *The Journal of Business* 34, 411-433.

Robertson, Donald, and Stephen Wright, 2006, Dividends, total cash flow to shareholders, and predictive return regressions, *Review of Economics and Statistics* 88, 91-99.

Shiller, Robert J., 2014, Speculative asset prices, *American Economic Review* 104, 1486-1517.

Skinner, Douglas J., 2008, The evolving relation between earnings, dividends, and stock repurchases, *Journal of Financial Economics* 87, 582-609.

Stambaugh, Robert F., 1999, Predictive regressions, *Journal of Financial Economics* 54, 375-421.

Welch, Ivo, and Amit Goyal, 2008, A comprehensive look at the empirical performance of equity premium prediction, *The Review of Financial Studies* 21, 1455-1508.

Wright, Stephen, 2004, Measures of stock market value and returns for the U.S. Nonfinancial corporate sector, 1900–2002, *Review of Income and Wealth* 50, 561-584.

Table 1. Summary statistics

Δd_t denotes log dividend growth, Δs_t denotes log repurchase growth, Δi_t denotes log issuance growth, r_t denotes log real CRSP portfolio return, and er_t denotes log excess return. We compute the log total payout ratio as $\delta_t = v_t - \beta_1 d_t - \beta_2 s_t + \beta_3 i_t$ (see (7)) and the log price–dividend ratio as $\eta_t = v_t - d_t$. The first and second rows in Panel A report means and standard deviations for the variables from calendar years 1972 to 2014. The third and fourth rows show t-statistics of the ADF test and adjusted t-statistics of the PP test with endogenous lags automatically selected by the Schwarz Criterion. Panel B shows contemporaneous correlations between the variables. Significant statistics used in Panels A and B at the 5% level are highlighted in bold face.

Panel A. Means, standard deviations, and results of unit root tests

	Δd_t	Δs_t	Δi_t	r_t	er_t	δ_t	η_t
Mean	0.037	0.103	0.034	0.099	0.098	3.653	3.649
Standard deviation	0.089	0.396	0.291	0.155	0.154	0.422	0.470
ADF	**−6.524**	**−6.384**	**−9.660**	**−6.525**	**−6.626**	**−2.962**	−1.174
PP	**−6.529**	**−7.049**	**−18.906**	**−6.534**	**−6.639**	**−2.934**	−1.138

Panel B. Contemporaneous correlations

Corr	Δd_t	Δs_t	Δi_t	r_t	er_t	η_t
Δs_t	0.293					
Δi_t	0.160	−0.038				
r_t	−0.128	0.191	0.015			
er_t	−0.129	0.184	0.013	**1.000**		
δ_t	−0.155	−0.150	**0.380**	−0.015	−0.015	
η_t	−0.013	0.024	0.062	0.239	0.237	0.258

Table 2. Constant-expected-real-return model for total payout ratios

The endogenous vector, $z_t = [\delta_t, \Delta\phi_t]'$, consists of total payout ratio δ_t and total payout growth $\Delta\phi_t$ (Subsection 4.1.1). All endogenous variables, measured at the end of June of calendar year t, are redefined as "centered" variables to be of mean zero by subtracting sample means from 1972 to 2014. Panel A reports VAR estimates, GMM corrected errors for heteroscedasticity under the estimates, and R^2s (%). Significant coefficients at the 5% level in Panel A are highlighted in bold face. Panel B reports VAR forecast δ_t' (see (9)), ratio of standard deviations $\sigma(\delta_t')/\sigma(\delta_t)$, and significance level for the Wald test of $H_0: \delta_t' = \delta_t$ (Hypothesis 1). We also report asymptotic standard errors numerically estimated under the estimates of δ_t' and $\sigma(\delta_t')/\sigma(\delta_t)$ in parentheses. We set the baseline discount factor to $\rho_3 \approx 0.976$.

Panel A. VAR estimates

Dependent variable	Endogenous variable δ_t (S.E.)	 $\Delta\phi_t$ (S.E)	R^2 (%)
δ_{t+1}	**0.698** (0.117)	0.044 (0.142)	47.23
$\Delta\phi_{t+1}$	0.205 (0.108)	-0.052 (0.149)	9.13

Panel B. Diagnostic statistics

$\delta_t' = e_2' A[I - \rho_3 A]^{-1} z_t$ (S.E.)	$\sigma(\delta_t')/\sigma(\delta_t)$ (S.E.)	χ^2 for $H_0: \delta_t' = \delta_t$ (*p*-value)
$\delta_t' = 0.627 \cdot \delta_t - 0.024 \cdot \Delta\phi_t$ (0.635) (0.237)	0.633 (0.588)	0.457 (0.796)

Table 3. Constant-expected-real-return model for price–dividend ratios

The endogenous vector, $z_t = [\eta_t, \Delta d_t]'$, consists of price–dividend ratio η_t and dividend growth Δd_t (see Appendix B). All endogenous variables, measured at the end of June of calendar year t, are redefined as "centered" variables to be of mean zero by subtracting sample means from 1972 to 2014. Panel A reports VAR estimates, GMM corrected standard errors for heteroscedasticity under the estimates, and R^2s (%). Significant coefficients at the 5% level in Panel A are highlighted in bold face. Panel B reports VAR forecast η_t', ratio of standard deviations $\sigma(\eta_t')/\sigma(\eta_t)$, and significance level for the Wald test of $H_0: \eta_t' = \eta_t$ (Hypothesis 1). We also report asymptotic standard errors numerically estimated under the estimates of η_t' and $\sigma(\eta_t')/\sigma(\eta_t)$ in parentheses. We set the baseline discount factor to $\rho_1 \approx 0.975$.

Panel A. VAR estimates

Dependent variable	Endogenous variable: η_t (S.E.)	Endogenous variable: Δd_t (S.E)	R^2 (%)
η_{t+1}	**0.930** (0.063)	0.111 (0.333)	85.65
Δd_{t+1}	0.036 (0.031)	-0.028 (0.238)	3.62

Panel B. Diagnostic statistics

$\eta_t' = e_2' A[I - \rho_1 A]^{-1} z_t$ (S.E.)	$\sigma(\eta_t')/\sigma(\eta_t)$ (S.E.)	χ^2 for $H_0: \eta_t' = \eta_t$ (*p*-value)
$\eta_t' = 0.386 \cdot \eta_t + 0.013 \cdot \Delta d_t$ (0.703) (0.399)	0.386 (0.694)	1.213 (0.545)

Table 4. Error standard deviations and correlations between contemporaneous shocks

The first three columns of Panel A report the results for payout yield shock ε^{δ}, total payout shock ε^{ϕ}, and implied return shock ε^{r}. Specifically, the first two errors ε^{δ} and ε^{ϕ} are computed as residuals from first-order VAR $z_{t+1} = Az_t + \varepsilon_{t+1}$ in Panel A of Table 2, where $\varepsilon_{t+1} = [\varepsilon^{\delta}_{t+1}, \varepsilon^{\phi}_{t+1}]'$. Implied ε^{r} is then estimated from error identity $\varepsilon^{r} = \varepsilon^{\phi} + \rho_3 \varepsilon^{\delta}$, where $\rho_3 \approx 0.976$. The last three columns report those of ε^{η}, ε^{d}, and implied ε^{r} for price–dividend ratio η_t. Panel B directly computes return shock $\varepsilon^{r}_{t+1} = r_{t+1} - E_t r_{t+1}$. In particular, we use log CRSP portfolio return r_{t+1} (Section 3). W then compute its ex-ante return as $E_t[r_{t+1}] = \rho_3 E_t[\delta_{t+1}] + E_t[\Delta\phi_{t+1}] - \delta_t$ or $E_t[r_{t+1}] = \rho_1 E_t[\eta_{t+1}] + E_t[\Delta d_{t+1}] - \eta_t$ through the first-order VAR.

Panel A. Error standard deviations and correlations using an implied return shock

	δ	$\Delta\phi$	Implied r		η	Δd	Implied r
δ	30.5	−87.3	36.1	η	18.0	−50.3	87.9
$\Delta\phi$	−87.3	29.4	14.0	Δd	−50.3	8.8	−3.0
Implied r	36.1	14.0	15.0	Implied r	87.9	−3.0	15.6

Panel B. Error standard deviations and correlations using the direct computation

	δ	$\Delta\phi$	CRSP r		η	Δd	CRSP r
δ	30.5	−87.3	28.6	η	18.0	−50.3	89.9
$\Delta\phi$	−87.3	29.4	20.6	Δd	−50.3	8.8	−11.3
CRSP r	28.6	20.6	14.8	CRSP r	89.9	−11.3	15.3

Table 5. Chave's variance decomposition of unexpected returns

Panel A presents the three covariance terms in (14) and the two correlation terms of implied return shock ε^r with news about total total payouts N^{CF} and news about future discount rates N^{DR} (Subsection 4.3.1). Here, we use total payout ratio δ_t and the VAR estimates in Panel A of Table 2. Panel B reports those using price–dividend ratio η_t and the VAR estimates in Panel A of Table 3. Note all covariance terms are normalized by return variance $\mathrm{Var}[\varepsilon^r]$, so the numbers reported below are shares that add up to one. The standard errors in parentheses are calculated based on the standard delta method (Subsection 4.1.1).

Panel A. Variance decomposition for total payout ratios

$\mathrm{Cov}(\varepsilon^\phi, \varepsilon^r)$ (S.E.)	$\mathrm{Cov}(N^\phi, \varepsilon^r)$ (S.E.)	$-\mathrm{Cov}(N^{DR}, \varepsilon^r)$ (S.E.)	$\mathrm{Corr}(N^{CF}, \varepsilon^r)$ (S.E.)	$\mathrm{Corr}(N^{DR}, \varepsilon^r)$ (S.E.)
0.454	0.331	0.214	0.761	−0.303
(0.000)	(0.420)	(0.422)	(0.235)	(0.288)

Panel B. Variance decomposition for price–dividend ratios

$\mathrm{Cov}(\varepsilon^d, \varepsilon^r)$ (S.E.)	$\mathrm{Cov}(N^d, \varepsilon^r)$ (S.E.)	$-\mathrm{Cov}(N^{DR}, \varepsilon^r)$ (S.E.)	$\mathrm{Corr}(N^{CF}, \varepsilon^r)$ (S.E.)	$\mathrm{Corr}(N^{DR}, \varepsilon^r)$ (S.E.)
-0.000	0.386	0.615	0.727	−0.860
(0.000)	(0.703)	(0.703)	(0.266)	(0.142)

Figure 1. Logs of the fiscal-year-end payout

The figure plots log dividend d_t, log repurchase s_t, and log issuance i_t from 1971 to 2014; for data construction, see Section 3. We also mark three important historical policies that have an influence on firms' distribution patterns: (a) Rule 10b-18 of the Securities Exchange Act (SEA) in November 1982, (b) Jobs and Growth Tax Relief Reconciliation Act (JGTRRA) in May 2003, and (c) the amendment of Rule 10b-18 in November 2003.

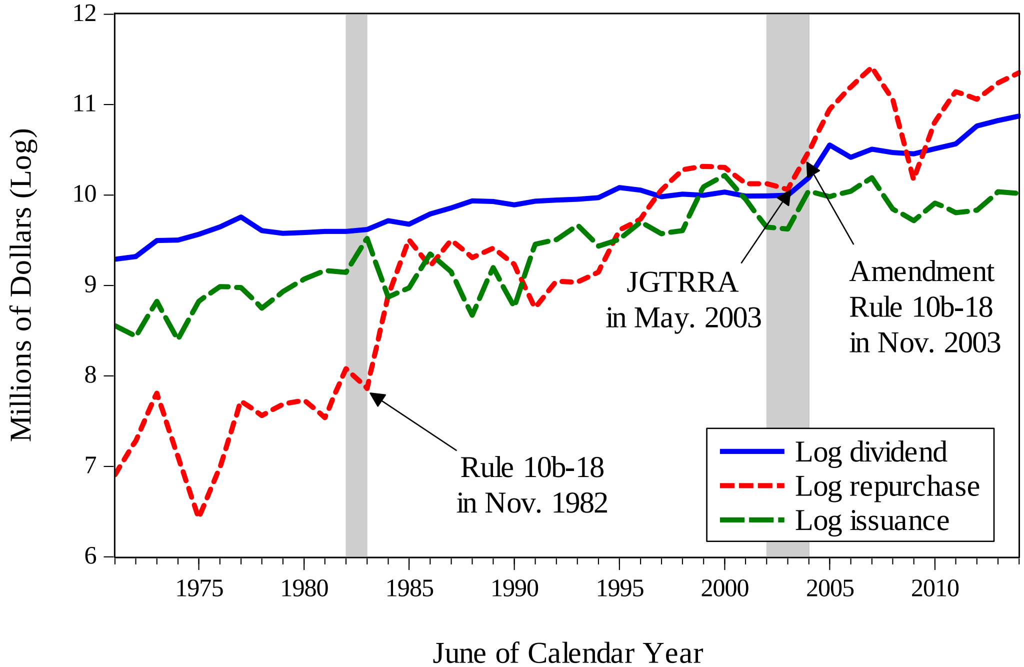


Figure 2. Present value relation of prices to corporate payouts

The figure plots log total payout ratio $\delta_t = v_t - \beta_1 d_t - \beta_2 s_t + \beta_3 i_t$ (see (7)) and log price–dividend ratio $\eta_t = v_t - d_t$ from 1971 to 2014. We use log market price v_t at the end of June of calendar year t and log dividend d_t, repurchase s_t, and issuance i_t in fiscal year ending in calendar year $t-1$ (Section 3). All variables are redefined as "centered variables" to be of mean zero by subtracting the sample means.

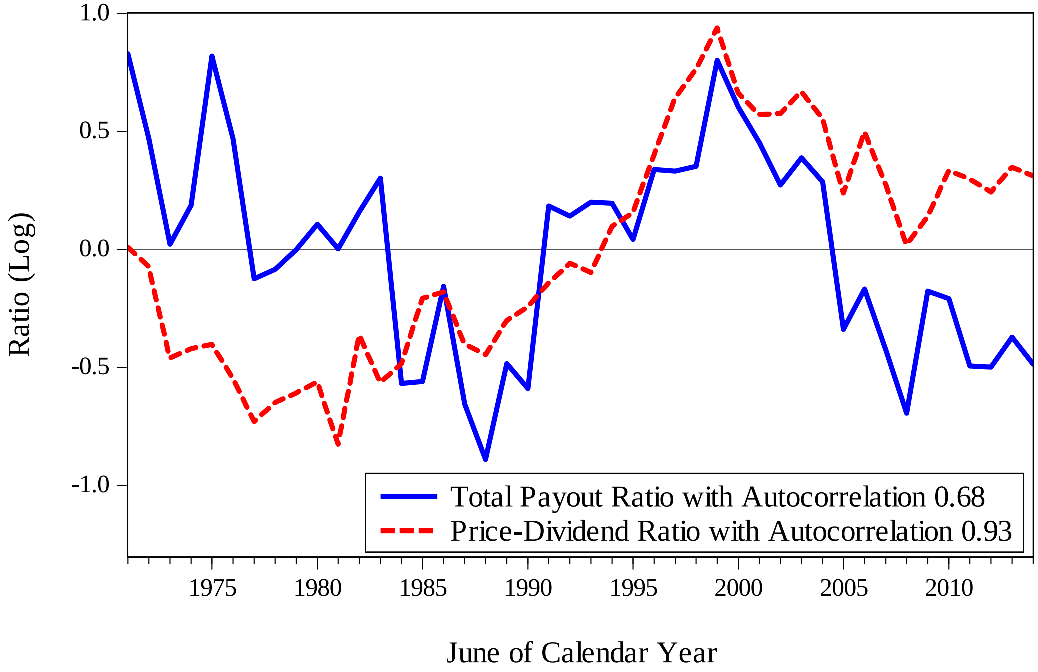


Figure 3. Present value of subsequent changes in expected future cash flows

The figure plots three prices from calendar years 1972 to 2014. The first is the real CRSP portfolio $\exp(v_t)$ at the end of June of calendar year t (Section 3). The second $\exp(\beta_1 d_t + \beta_2 s_t - \beta_3 i_t + \delta_t')$ is prices justified by subsequent changes in expected future total payouts with $\rho_3 \approx 0.976$. The third $\exp(d_t + \eta_t')$ is prices justified by subsequent changes in expected future dividends with $\rho_1 \approx 0.975$. All variables are redefined as "centered variables" to of be mean zero by subtracting the sample means.

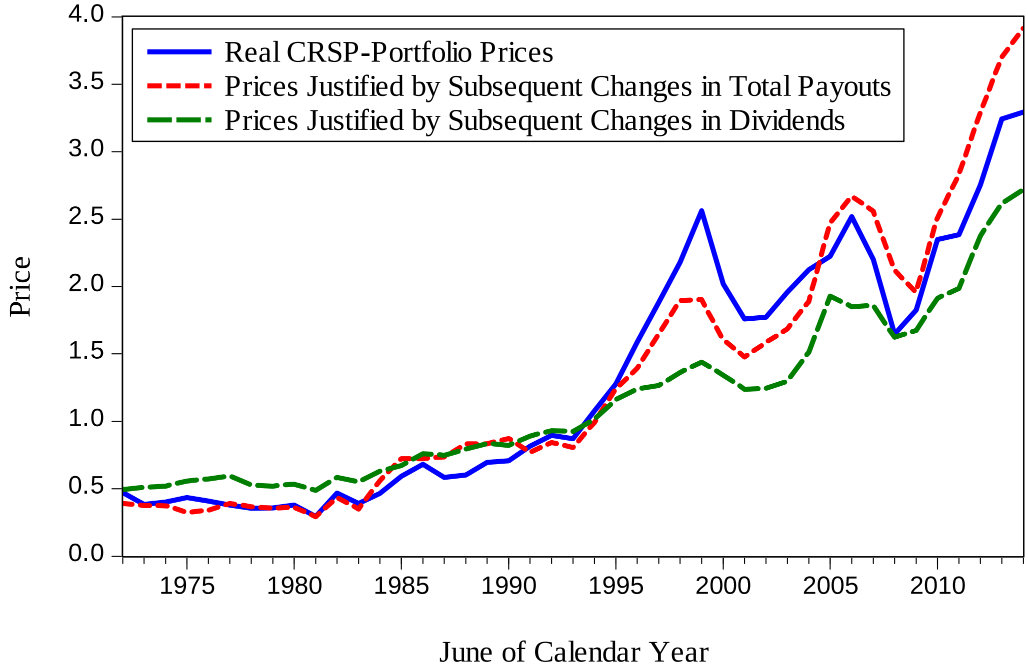


Figure 4. Impulse response functions for a payout yield shock with no change in returns

The impulse response functions are based on the VAR estimates in Panel A of Table 2 and the two error identities: $\varepsilon^r = \varepsilon^\phi + \rho_3\varepsilon^\delta$ and $\varepsilon^v = \varepsilon^\phi + \varepsilon^\delta$, where $\rho_3 \approx 0.976$. The implied return data are generated from the present value identity: $r_{t+1} = \rho_3\delta_{t+1} + \Delta\phi_{t+1} - \delta_t$. We then look at how the pure cash flow shock, $[\varepsilon^\delta, \varepsilon^\phi, \varepsilon^r] = [1, -\rho_3, 0]$, changes (first panel) log equity return r_t and total payout growth $\Delta\phi_t$, and (second panel) total payout ratio δ_t, price v_t, and total payout ϕ_t through time. The vertical dashed line represents the starting time of the shock.

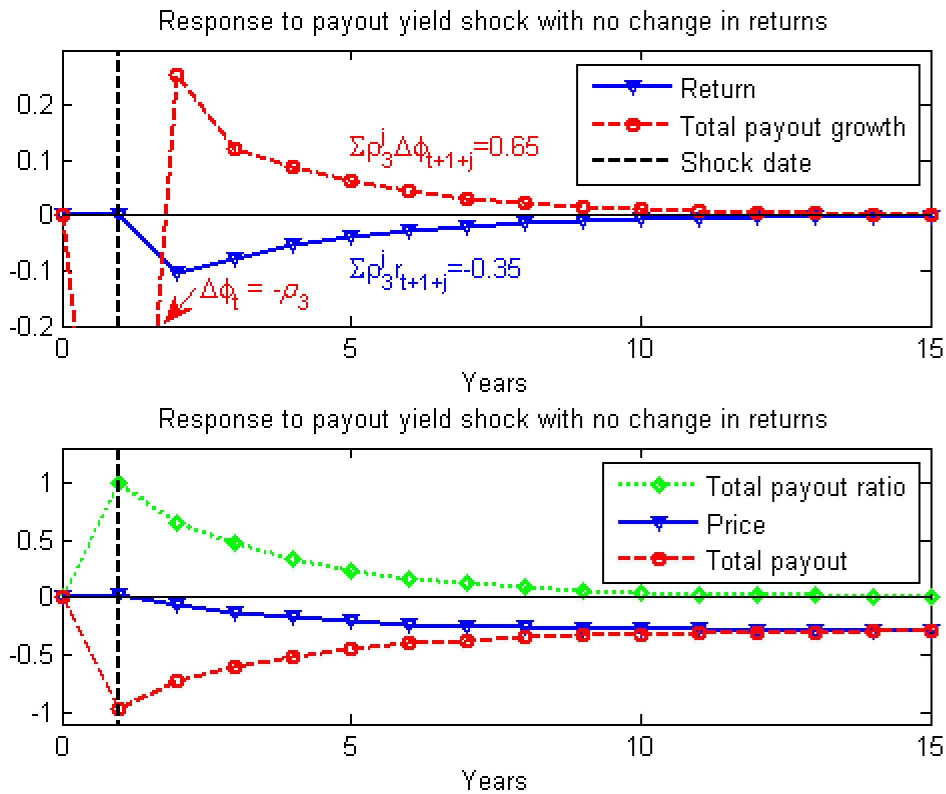


Figure 5. Impulse response functions for a dividend yield shock with no change in dividends

The impulse response functions are based on the VAR estimates in Panel A of Table 3 and error identity $\varepsilon^r = \varepsilon^d + \rho_1\varepsilon^\eta$, where $\rho_1 \approx 0.975$. The implied return data are generated from the present value identity: $r_{t+1} = \rho_1\eta_{t+1} + \Delta d_{t+1} - \eta_t$. We then look at how the dividend yield shock with no move in dividends, $[\varepsilon^\eta, \varepsilon^d, \varepsilon^r] = [1, 0, \rho_1]$, changes (first panel) log equity return r_t and dividend growth Δd_t, and (second panel) price–dividend ratio η_t, price v_t, and total dividend d_t through time. The vertical dashed line represents the starting time of the shock.

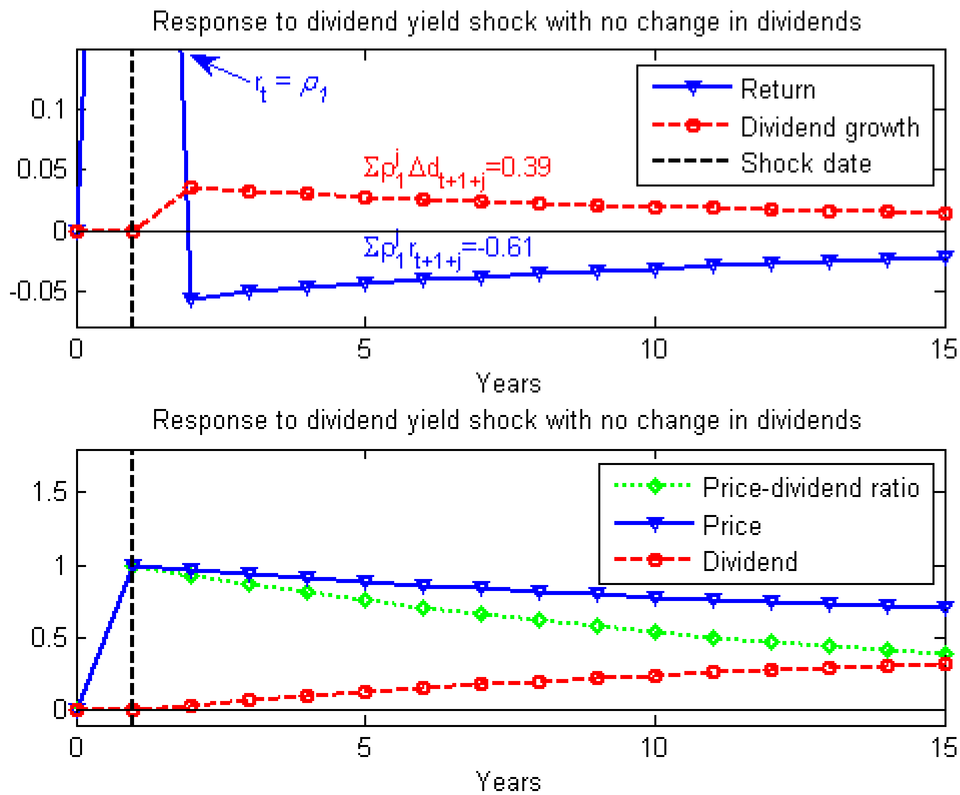


Figure 6. Impulse response functions for a payout yield shock with no change in cash flows

The impulse response functions are based on the VAR estimates in Panel A of Table 2 and the two error identities: $\varepsilon^r = \varepsilon^\phi + \rho_3\varepsilon^\delta$ and $\varepsilon^v = \varepsilon^\phi + \varepsilon^\delta$, where $\rho_3 \approx 0.976$. The implied return data are generated from the present value identity: $r_{t+1} = \rho_3\delta_{t+1} + \Delta\phi_{t+1} - \delta_t$. We then look at how the pure cash flow shock, $[\varepsilon^\delta, \varepsilon^\phi, \varepsilon^r] = [1, 0, \rho_3]$, changes (first panel) log equity return r_t and total payout growth $\Delta\phi_t$, and (second panel) total payout ratio δ_t, price v_t, and total payout ϕ_t through time. The vertical dashed line represents the starting time of the shock.

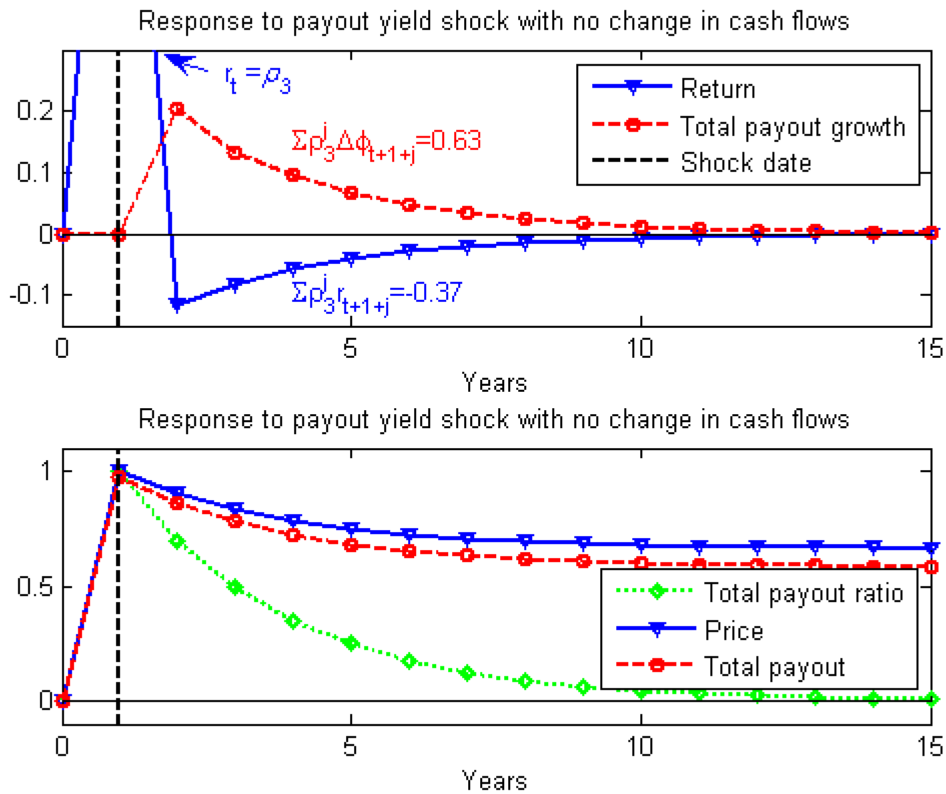
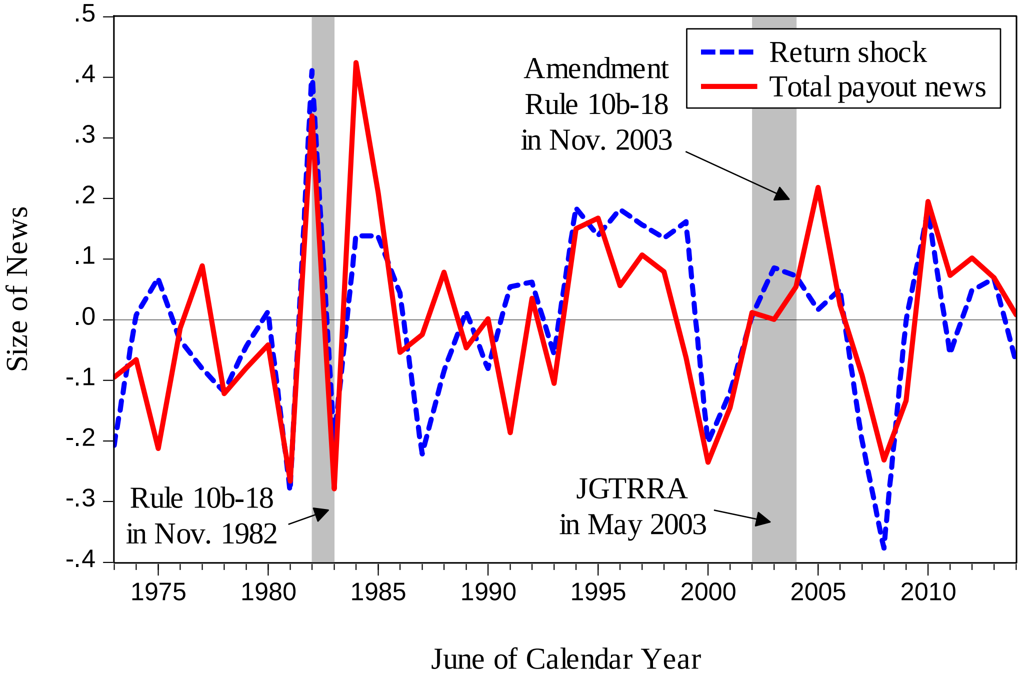


Figure 7. Return shock and news for total payout ratios

Graph A plots return shock ε^r and news about total payouts N^{CF} from 1973 to 2014. Graph B plots ε^r and news about future discount rates $N^{DR} = N^{CF} - \varepsilon^r$. Calculations are based on the VAR estimates in Panel A of Table 2 and error identity $\varepsilon^r = \varepsilon^\phi + \rho_3 \varepsilon^\delta$, where $\rho_3 \approx 0.976$.

Graph A. Return shock and news about total payouts



Graph B. Return shock and news about future discount rates

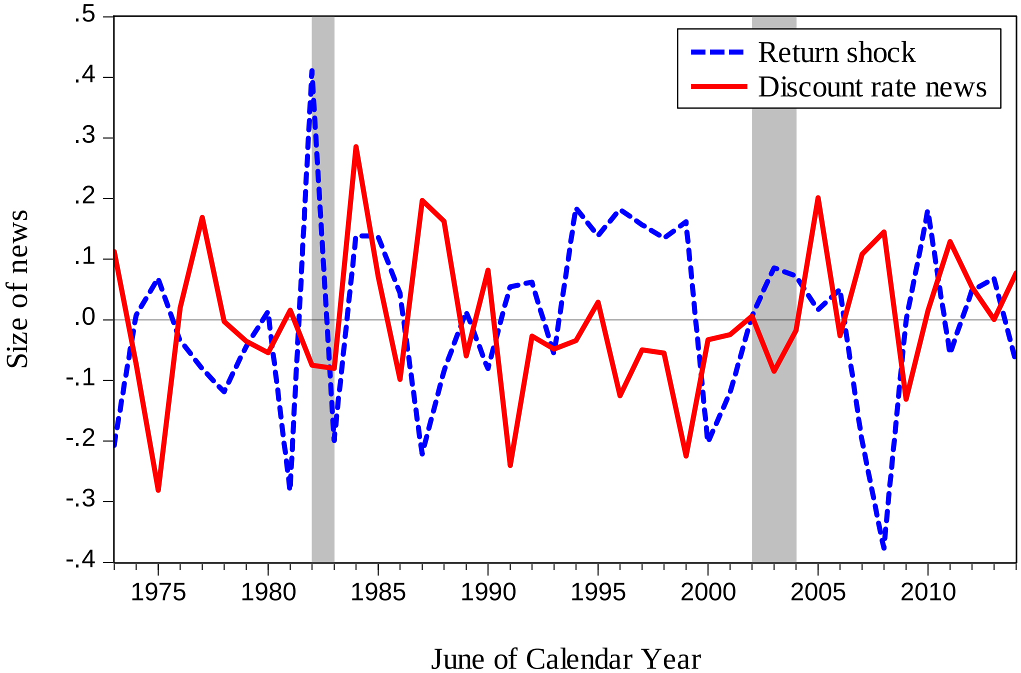
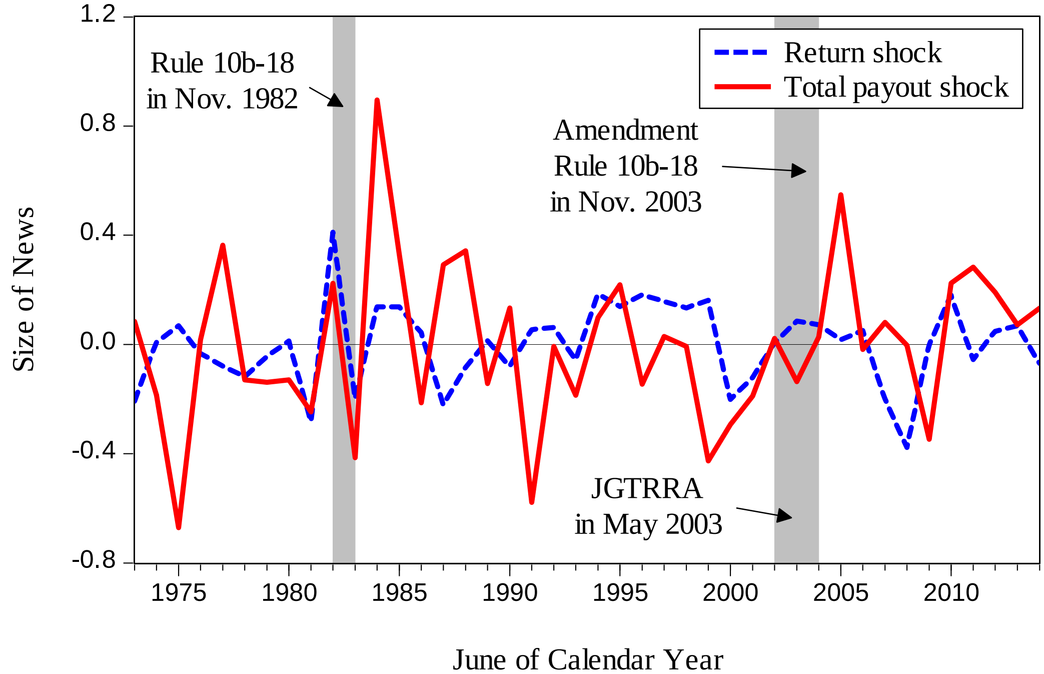


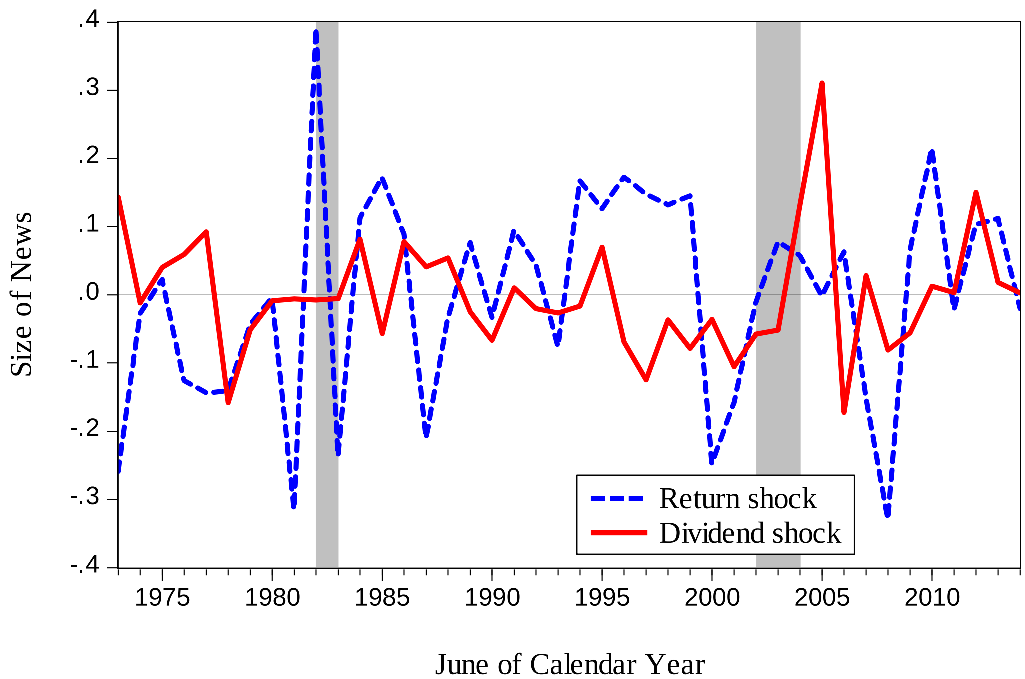
Figure 8. Return shock and cash flow shock

Graph A plots return shock ε^r and total payout shock ε^ϕ from 1973 to 2014. Calculations are based on the VAR estimates in Panel A of Table 2 and error identity $\varepsilon^r = \varepsilon^\phi + \rho_3 \varepsilon^\delta$, where $\rho_3 \approx 0.976$. Graph B plots ε^r and dividend shock ε^d, calculated based on the VAR estimates in Panel A of Table 3 and error identity $\varepsilon^r = \varepsilon^d + \rho_1 \varepsilon^\eta$, where $\rho_1 \approx 0.975$.

Graph A. Return shock and total payout shock



Graph B. Return shock and dividend shock



Appendix

A. Derivation of log-linear approximation

An arbitrary function, $f(x,y,z)$, can be approximated by the first-order Taylor approximation around long-term means $\bar{x}$, $\bar{y}$, and $\bar{z}$ as

$$f(x,y,z) \approx$$
$$f(\bar{x},\bar{y},\bar{z}) + \frac{\partial f}{\partial x}\Big|_{x=\bar{x},y=\bar{y},z=\bar{z}}(x-\bar{x}) + \frac{\partial f}{\partial y}\Big|_{x=\bar{x},y=\bar{y},z=\bar{z}}(y-\bar{y}) + \frac{\partial f}{\partial z}\Big|_{x=\bar{x},y=\bar{y},z=\bar{z}}(z-\bar{z}),$$

where $\partial f/\partial x$, $\partial f/\partial y$, and $\partial f/\partial z$ are respectively partial derivatives with respect to x, y, and z.

First, we apply a first-order Taylor expansion to the non-linear term in (3) around the long-term means of $\overline{dv} \equiv E[dv_{t+1}]$, $\overline{sv} \equiv E[sv_{t+1}]$, and $\overline{iv} \equiv E[ip_{t+1}]$:

$$-\log(\rho_3) + \frac{\exp(\overline{dv})}{1+\exp(\overline{dv})+\exp(\overline{sv})-\exp(\overline{iv})}\left(dv_{t+1} - \overline{dv}\right)$$
$$+\frac{\exp(\overline{dv})}{1+\exp(\overline{dv})+\exp(\overline{sv})-\exp(\overline{iv})}\left(sv_{t+1} - \overline{sv}\right) - \frac{\exp(\overline{iv})}{1+\exp(\overline{dv})+\exp(\overline{sv})-\exp(\overline{iv})}\left(iv_{t+1} - \overline{iv}\right)$$
$$= k_3 + (\rho_3 - 1)v_{t+1} + \rho_3\left[\left(\frac{1}{\rho_1} - 1\right)d_{t+1} + \left(\frac{1}{\rho_2} - \frac{1}{\rho_1}\right)s_{t+1} - \left(\frac{1}{\rho_2} - \frac{1}{\rho_3}\right)i_{t+1}\right]. \qquad \text{(A.1)}$$

where $\rho_1 = \frac{1}{1+\exp(\overline{dv})} = \frac{\bar{V}}{\bar{V}+\bar{D}}$, $\rho_2 = \frac{1}{1+\exp(\overline{dv})+\exp(\overline{sv})} = \frac{\bar{V}}{\bar{V}+\bar{D}+\bar{S}} < \rho_1$, $\rho_3 = \frac{1}{1+\exp(\overline{dv})+\exp(\overline{sv})-\exp(\overline{iv})} = \frac{\bar{V}}{\bar{V}+\bar{D}+\bar{S}-\bar{I}} > \rho_2$, and k_3 is a log-linear constant:

$$k_3 = -\log(\rho_3) - \rho_3\left(\frac{1}{\rho_1} - 1\right)\log\left(\frac{1}{\rho_1} - 1\right) - \rho_3\left(\frac{1}{\rho_2} - \frac{1}{\rho_1}\right)\log\left(\frac{1}{\rho_2} - \frac{1}{\rho_1}\right) + \rho_3\left(\frac{1}{\rho_2} - \frac{1}{\rho_3}\right)\log\left(\frac{1}{\rho_2} - \frac{1}{\rho_3}\right).$$

Plugging (A.1) into log return (3) yields (4):

$$r_{t+1} \approx k_3 + \rho_3 \left[v_{t+1} + \left(\frac{1}{\rho_1} - 1\right) d_{t+1} + \left(\frac{1}{\rho_2} - \frac{1}{\rho_1}\right) s_{t+1} - \left(\frac{1}{\rho_2} - \frac{1}{\rho_3}\right) i_{t+1} \right] - v_t.$$

Second, we solve present value identity (4) forward by repeatedly substituting in future prices until terminal time T:

$$v_t \approx \sum_{j=0}^{T} (\rho_3)^j \left[\rho_3 \left\{ \left(\frac{1}{\rho_1} - 1\right) d_{t+1+j} + \left(\frac{1}{\rho_2} - \frac{1}{\rho_1}\right) s_{t+1+j} - \left(\frac{1}{\rho_2} - \frac{1}{\rho_3}\right) i_{t+1+j} \right\} - r_{t+1+j} \right]$$

$$+ (\rho_3)^{T+1} v_{t+1+T}.$$

The fact that the value of a discount factor, $\rho_3 = \bar{V}/(\bar{V} + \bar{D} + \bar{S} - \bar{I})$, is less than one can justify that rational bubbles do not exist. Hence, the terminal condition holds: $\lim_{T\to\infty} (\rho_3)^{T+1} v_{t+1+T} = 0$. Proposition 1 with the terminal condition delivers the ex-post present value identity

$$\delta_t \approx \sum_{j=0}^{\infty} (\rho_3)^j \left[\{ \beta_1 \Delta d_{t+1+j} + \beta_2 \Delta s_{t+1+j} - \beta_3 \Delta i_{t+1+j} \} - r_{t+1+j} \right]. \qquad \text{(A.2)}$$

The fact that (A.2) holds ex-post ensures that it also holds ex-ante for any information set as

$$\delta_t \approx E\left[\sum_{j=0}^{\infty} (\rho_3)^j \{ \beta_1 \Delta d_{t+1+j} + \beta_2 \Delta s_{t+1+j} - \beta_3 \Delta i_{t+1+j} \} - r_{t+1+j} | \mathcal{F}_t \right]. \qquad \text{(A.3)}$$

Note that (A.3) is the same as present value identity (6).

The substitution of $\rho_1 = \frac{\bar{V}}{\bar{V}+\bar{D}}$, $\rho_2 = \frac{\bar{V}}{\bar{V}+\bar{D}+\bar{S}}$, $\rho_3 = \frac{\bar{V}}{\bar{V}+\bar{D}+\bar{S}-\bar{I}}$ into β_1, β_2, and β_3 leads to a portion of each payout to total payouts:

$$\beta_1 = \frac{\bar{D}}{\bar{D}+\bar{S}-\bar{I}}, \ \beta_2 = \frac{\bar{S}}{\bar{D}+\bar{S}-\bar{I}}, \text{ and } \beta_3 = \frac{\bar{I}}{\bar{D}+\bar{S}-\bar{I}}.$$

Some may argue that each ratio used in the log-linear approximation (i.e., dv_{t+1}, sv_{t+1}, and iv_{t+1}) must be stationary for the derivation. However, Proposition 1 allows multivariate Taylor expansion (A.1) to restate with univariate Taylor expansion (A.4):

$$-\log(\rho_3) - \frac{\exp(\overline{dv})+\exp(\overline{sv})-\exp(\overline{iv})}{1+\exp(\overline{dv})+\exp(\overline{sv})-\exp(\overline{iv})} \cdot (\delta_{t+1} - \bar{\delta}), \tag{A.4}$$

where $\frac{\exp(\overline{dv})+\exp(\overline{sv})-\exp(\overline{iv})}{1+\exp(\overline{dv})+\exp(\overline{sv})-\exp(\overline{iv})} = 1 - \rho_3$, $\bar{\delta} = -\beta_1\overline{dv} - \beta_2\overline{sv} + \beta_3\overline{iv}$, and $k_3 = -\log(\rho_3) - (1 - \rho_3)\bar{\delta}$. Substituting (A.4) into log return (3) delivers the same representation in (6).

B. Comparison with conventional studies

First, *constant variation in expected returns* yields Miller and Modigliani's (1961) payout irrelevance proposition:

$$V_t = E_t\left[\sum_{j=0}^{\infty}\left(\frac{1}{1+R}\right)^j [D_{t+1+j} - (I_{t+1+j} - S_{t+1+j})]\right] = E_t\left[\sum_{j=0}^{\infty}\left(\frac{1}{1+R}\right)^j [X_{t+1+j} - \theta_{t+1+j}]\right],$$

where $\rho_3 = \bar{V}/(\bar{V} + \bar{D} + \bar{S} - \bar{I}) = 1/(1+R)$, X_t is the amount of the firm's earnings during time t, and θ_t denote the amount of the firm's investment; and the *funds identity* is given by

$$D_t - (I_t - S_t) = X_t - \theta_t.$$

Miller and Modigliani (1961) show that prices are independent of dividend payout policy because dividend payments can *exactly* offset net share issuances $I_t - S_t$. What really matters is *not* cash distributions but the firm's earnings power and investment policy. Further, *time variation in expected returns* yields the dynamic version of the conventional proposition:

$$\delta_t = v_t - \alpha_1 x_t + \alpha_2\theta_t \approx E_t\left[\sum_{j=0}^{\infty}(\rho_3)^j[\alpha_1\Delta x_{t+1+j} - \alpha_2\Delta\theta_{t+1+j} - r_{t+1+j}]\right],$$

where $\rho_3 = \frac{1}{1+\exp(\overline{x_t - v_{t+1}}) - \exp(\overline{\theta_t - v_{t+1}})} = \frac{\bar{V}}{\bar{V}+\bar{X}-\bar{\theta}}$ due to $D_t - (I_t - S_t) = X_t - \theta_t$, $\rho_4 = \frac{1}{1+\exp(\overline{x_t - v_{t+1}})} < \rho_3$, $\alpha_1 = \frac{\rho_3(1-\rho_4)}{\rho_4(1-\rho_3)} = \frac{\bar{X}}{\bar{X}-\bar{\theta}}$, and $\alpha_2 = \frac{\rho_3 - \rho_4}{\rho_4(1-\rho_3)} = \frac{\bar{\theta}}{\bar{X}-\bar{\theta}}$ such that $\alpha_1 - \alpha_2 = 1$.

Second, suppose that dividends are the only payout method to shareholders; equivalently, the changes in the number of shares outstanding do not affect prices at all such that $I_t - S_t = P_t \cdot (N_t - N_{t-1}) = 0$:

$$1 + R_{t+1} = \frac{V_{t+1} + D_{t+1} - (I_{+1t} - S_{t+1})}{V_t} = \frac{V_{t+1} + D_{t+1}}{V_t},$$

In this case, Proposition 1 shows that $\beta_1 = 1$ because β_2 and β_3 become zero as a result of $\rho_1 = \rho_2 = \rho_3 = \frac{\bar{V}}{\bar{V}+\bar{D}}$. Subsequently, present value relation (6) can be rewritten as

$$\eta_t \approx E_t\left[\sum_{j=0}^{\infty} (\rho_1)^j [\Delta d_{t+1+j} - r_{t+1+j}]\right],$$

where $\eta_t = v_t - d_t = \log(P_t \times N_t) - \log(DPS_t \times N_t) = p_t - dps_t$ denotes the log price–dividend ratio. If η_t does not differ substantially from δ_t, non-dividend cash flows such as share repurchases and issuances do not influence price variation substantially.

Third, suppose that repurchases can be used as another payout method without external financing such as issuances. As a result, Proposition 1 becomes $\beta_1 + \beta_2 = 1$ alongside $\beta_3 = 0$ and gives rise to

$$v_t - \beta_1 d_t - \beta_2 s_t \approx E_t\left[\sum_{j=0}^{\infty} (\rho_2)^j [\beta_1 \Delta d_{t+1+j} + \beta_2 \Delta s_{t+1+j} - r_{t+1+j}]\right],$$

where $\beta_1 = \frac{\rho_2(1-\rho_1)}{\rho_1(1-\rho_2)} = \frac{\bar{D}}{\bar{D}+\bar{S}}$ and $\beta_2 = \frac{\rho_1 - \rho_2}{\rho_1(1-\rho_2)} = \frac{\bar{S}}{\bar{D}+\bar{S}}$. This present value identity might hold when dividends are perfect substitute for repurchases (Grullon and Michaely, 2002).

Fourth, present value identity (6) also involves Larrain and Yogo's (2008) framework. Following Larrain and Yogo (2008), suppose that ds_t denotes log dividend plus repurchase, $\log(D_t + S_t)$, to avoid potential negative net issuance: $I_t - S_t < 0$. As a result, Proposition 1 delivers $\beta_1 - \beta_3 = 1$:

$$v_t - \beta_1 ds_t + \beta_3 i_t \approx E_t\left[\sum_{j=0}^{\infty}(\rho_3)^j\left[\beta_1 \Delta ds_{t+1+j} - \beta_3 \Delta i_{t+1+j} - r_{t+1+j}\right]\right],$$

where $\beta_1 = \frac{\rho_3(1-\rho_1)}{\rho_1(1-\rho_3)} = \frac{\overline{D+S}}{\overline{D+S}-\bar{I}} > 1$, $\beta_3 = \frac{\rho_3-\rho_1}{\rho_1(1-\rho_3)} = \frac{\bar{I}}{\overline{D+S}-\bar{I}} > 0$, $\rho_3 = \frac{1}{1+\exp(\overline{ds-v})-\exp(\overline{iv})} = \frac{\bar{V}}{\bar{V}+\overline{D+S}-\bar{I}} > \rho_1$ and $\rho_1 = \frac{1}{1+\exp(\overline{ds-v})} = \frac{\bar{V}}{\bar{V}+\overline{D+S}}$. Note that θ of Larrain and Yogo (2008) is matched by β_1 above. Our calculation β_1 is about 1.41, corresponding to $\beta_3 = 0.41$ by Proposition 1. These beta estimates mean that the historical proportion of the issuance amount to dividend plus repurchase amount is about $\frac{\beta_3}{\beta_1} = \frac{\bar{I}}{\overline{D+S}} = 29\%$ on average over our sample periods. In contrast, $\theta(\text{Larrain and Yogo, } 2008)$ is 2.5, implying that the historical proportion of issuances is about 60% on average over their sample periods.

C. Campbell's variance decomposition of unexpected returns

Campbell's variance decomposition of unexpected returns is

$$\mathrm{Var}[\varepsilon^r] = \mathrm{Var}[N^{CF}] + \mathrm{Var}[N^{DR}] - 2 \cdot \mathrm{Cov}[N^{CF}, N^{DR}]. \tag{C-1}$$

The further decomposition of N^{CF} into ε^{ϕ} and N^{ϕ} allows additional variance and covariance terms: $\mathrm{Var}[N^{CF}] = \mathrm{Var}\left[\varepsilon^{\phi}\right] + \mathrm{Var}\left[N^{\phi}\right] + 2 \cdot \mathrm{Cov}[\varepsilon^{\phi}, N^{\phi}]$. However, we do not present them here since we have already studied those effects through Chave's variance decomposition (Subsection 4.3.2). In essence, the two variance decompositions are identical.

Panel A of Table C.1 calculates Campbell's decomposition implications for total payout ratio δ_t. Specifically, slightly more than 100% of return variance $\mathrm{Var}[\varepsilon^r]$ is attributed to the variance of news about total payouts: $\mathrm{Var}[N^{CF}] \approx 106.6\%$, where $N^{CF} = \varepsilon^{\phi} + N^{\phi}$. The remainders correspond to the variance of news about discount rates or the covariance term: $\mathrm{Var}[N^{DR}] \approx 49.9\%$ and $-2 \times \mathrm{Cov}[N^{CF}, N^{DR}] \approx -56.1\%$, but they seem to almost offset each other. We do not report the asymptotic standard errors because they come out to be very large due to the lack of the data samples.

Table C. 1. Campbell's variance decomposition of unexpected returns
Panel A presents the three variance-covariance terms in (C-1) and the correlation between news about future total payouts and discount rates. Here, we use total payout ratio $\delta_t = v_t - \beta_1 d_t - \beta_2 s_t + \beta_3 i_t$ (see (7)) and the VAR estimates in Panel A of Table 2. Panel B reports those using price–dividend ratio $\eta_t = v_t - d_t$ and the VAR estimates in Panel A of Table 3. Note all variance-covariance terms are normalized by return variance $\mathrm{Var}[\varepsilon^r]$, so the numbers reported below add up to one.

Panel A. CF: total payouts

$\mathrm{Var}[N^{CF}]$	$\mathrm{VAR}[N^{DR}]$	$-2 \times \mathrm{Cov}[N^{CF}, N^{DR}]$	$\mathrm{Corr}[N^{CF}, N^{DR}]$
1.066	0.499	−0.561	0.706

Panel B. CF. dividends

$\mathrm{Var}[N^{CF}]$	$\mathrm{VAR}[N^{DR}]$	$-2 \times \mathrm{Cov}[N^{CF}, N^{DR}]$	$\mathrm{Corr}[N^{CF}, N^{DR}]$
0.281	0.511	0.208	−0.275

Our results strongly contrast with the dividend yield implications (Panel B, Table C.1), pointing out that news about discount rates is a main source of return variation. This is consistent with Campbell (1991) and Campbell and Ammer (1993). In particular, slightly over than half the variance of return shocks is attributed to news about discount rates: $\mathrm{Var}[N^{DR}] \approx 51.1\%$. The remaining variance and covariance account for the rest of the return variance with the same positive sign: $\mathrm{Var}[N^{CF}] \approx 28.1\%$ and $-2 \times \mathrm{Cov}[N^{CF}, N^{DR}] \approx 20.8\%$, where $N^{CF} = \varepsilon^d + N^d$.

The key difference between Panels A and B arises from whether cash flow news N^{CF} and discount rate news N^{DR} offset each other or not. More precisely, each covariance is calculated as

$$\text{Cov}[N^{CF}, N^{DR}] = -0.22 \cdot \text{Var}\left[\varepsilon^{\delta}\right] - 0.37 \cdot \text{Cov}\left[\varepsilon^{\delta}, \varepsilon^{\phi}\right] - 0.02 \cdot \text{Var}\left[\varepsilon^{\phi}\right] > 0,$$

$$\text{Cov}[N^{CF}, N^{DR}] = -0.23 \cdot \text{Var}[\varepsilon^{\eta}] - 0.60 \cdot \text{Cov}[\varepsilon^{\eta}, \varepsilon^{d}] + 0.01 \cdot \text{Var}[\varepsilon^{d}] < 0,$$

where $N_{t+1}^{CF} = e_2'[I - \rho_i A]^{-1}\varepsilon_{t+1}$ for $i = 3$ or 1 and $N_{t+1}^{DR}(= N_{t+1}^{CF} - \varepsilon_{t+1}^{r})$ are calculated as residuals (Subsection 4.2.1). In the first equation, the fact that the variance terms $\text{Var}\left[\varepsilon^{\delta}\right]$ and $\text{Var}\left[\varepsilon^{\phi}\right]$ on the RHS must be greater than zero requires strong negative covariance $\text{Cov}\left[\varepsilon^{\delta}, \varepsilon^{\phi}\right]$ on the LHS: $\text{Corr}\left[\varepsilon^{\delta}, \varepsilon^{\phi}\right] \approx -87.3\%$ (Panel A, Table 4).

In the second equation, the negative covariance on the LHS should follow the large variance of dividend yield shocks: $\text{Var}[\varepsilon^{\eta}] > 0$ given $\text{Corr}[\varepsilon^{\eta}, \varepsilon^{d}] \approx -50.3\%$ (Panel A, Table 4). This finding accords well with the conventional view of time-varying discount rates: a dividend shock has almost nothing to do with a dividend yield shock (Cochrane, 2008).