Cash Flow News and Stock Price Dynamics*

Davide Pettenuzzo[†] Brandeis University

Riccardo Sabbatucci[‡] Stockholm School of Economics

Allan Timmermann[§] UC San Diego

November 11, 2019

Abstract

We develop a new approach to modeling dynamics in cash flow data extracted from daily firm-level dividend announcements. We decompose daily cash flow news into a persistent component, jumps, and temporary shocks. Empirically, we find that the persistent cash flow component is a highly significant predictor of future growth in dividends and consumption. Using a log-linearized present value model, we show that news about the persistent dividend growth component helps predict stock returns consistent with asset-pricing constraints implied by this model. News about the daily dividend growth process also helps explain concurrent return volatility and the probability of jumps in stock returns.

Keywords: High-frequency cash flow news; predictability of dividend growth; present value model; dynamics and predictability of stock returns; Bayesian modeling

*We are grateful to the editor, Stefan Nagel, an Associate Editor, and two anonymous referees for many valuable suggestions on our paper. We also received helpful comments from Jules van Binsbergen, Chris Polk, Federico Bandi, and participants at the Utah Winter Finance Conference 2019, CEPR Gerzensee 2018, BI-SHoF 2018 and conferences in Toronto (Western University, April 2018), Toulouse (Toulouse School of Economics, May 2018), and University of Chicago (May 2019). We are grateful to INQUIRE Europe for financial support.

[†]International Business School, Brandeis University. **Email:** dpettenu@brandeis.edu

[‡]Department of Finance, Stockholm School of Economics. **Email:** riccardo.sabbatucci@hhs.se

[§]Rady School of Management, UC San Diego. **Email:** atimmermann@ucsd.edu

1 Introduction

On most days, a multitude of firms announce cash flow news, but the number of firms, as well as the industries they belong to, can vary greatly over time. Such variation gives rise to a highly irregular cash flow news process and complicates investors' attempt to infer the underlying growth rate of cash flows for individual firms, industries, and for the economy as a whole. This is important because the resulting cash flow growth estimates are a key driver of investors' forecasts of future cash flows, their assessment of cash flow risks, and, ultimately, of how stocks are priced.[1]

While information extracted from firms' cash flow announcements is critical to understanding investors' cash flow expectations and, in turn, movements in stock prices, relatively few studies analyze predictability of cash flows and, in most cases, focus on quarterly or annual changes in aggregate dividends or earnings.[2] However, data aggregated in this manner may conceal the rich dynamic patterns in cash flows recorded at a higher frequency which reduces our ability to study important questions such as how fast cash flow growth responds to changes in the underlying state of the economy.

Several challenges complicate attempts to measure daily cash flow dynamics. First, most firms' cash flows have a pronounced seasonal component related to weather patterns and holiday sales. Second, the number of firms announcing cash flow news on any given day can fluctuate between as little as zero to more than one hundred firms. Third, the particular date on which a firm announces dividends can vary widely from year to year, requiring that close attention be paid to constructing daily proxies that account for firm specific effects. Fourth, there is considerable heterogeneity across individual firms' cash flow processes. The combined effect of these factors is that daily cash flow news tends to be very lumpy.

[1] Patton and Verardo (2012) develop a rational learning model to explain the patterns in firms' betas observed around earnings announcements. Their model contains unobserved firm-specific and common earnings innovation terms and investors' extraction of these components is modeled as a Kalman filtering problem. Savor and Wilson (2016) develop a learning model in which investors decompose cash flow news into firm-specific and market-wide components. Positive covariances between the cash flow process of individual firms and of the broader market imply that bad (good) news on individual firms' cash flows result in reduced (increased) forecasts of aggregate cash flows. In turn, this cash flow learning channel implies that stock returns of announcing firms and of the aggregate market are positively correlated, justifying an announcement risk premium for exposure to individual firms' cash flows. These models do not allow for jumps in cash flows, although in practice this is an important feature of earnings and dividend data.

[2] Cochrane (2008) finds little evidence of dividend growth predictability, while van Binsbergen and Koijen (2010), Kelly and Pruitt (2013) and Jagannathan and Liu (2019) find that growth in dividends is predictable.

To address these challenges, we develop a new approach to measure and model dynamics in high frequency (daily) cash flows. We deal with firm-level heterogeneity and seasonality effects by taking a bottom-up approach that starts from changes in individual firms' dividends on a given day relative to their payments over the same quarter during the previous year. In contrast with conventional smoothed estimates, only data on those firms that announce dividend news on a given day are used to update the dividend growth estimate, ensuring that our measure is timely in picking up changes in the cash flow process.[3] Moreover, by computing a dollar-weighted growth estimate, we account for variation in the size of the firms paying dividends on any given day.[4] Our dividend growth measure uses dividend announcements by publicly traded firms as opposed to dividend payments which form the basis for the CRSP measure of dividend growth conventionally used in the finance literature. This is an important distinction because dividend announcements precede dividend payments by several weeks, giving our dividend measure a significant timing advantage and more closely aligns our measure with movements in market prices following dividend news.[5]

We account for lumpiness in daily values of year-on-year changes in firm-matched cash flows, by decomposing cash flow news into a slowly evolving component that picks up time variation (predictability) in the mean of the cash flow process, a transitory component whose volatility is allowed to change over time, and large jumps whose probability of occurring can depend on the number of firms that announce cash flow news on a given day. All three components turn out to be important for capturing predictability in the dividend growth process and evolution in the uncertainty that surrounds growth in cash flows.

An important test of our approach is whether it can be used to generate more accurate forecasts of cash flows than existing methods. Empirically, we find that our estimate of the persistent dividend growth component is a strong predictor of future dividend growth. Moreover, the predictive power of our approach compares favorably to alternative predictors

[3]To illustrate the loss in information from the common practice of aggregating cash flow news over the most recent 12-month period and updating this on, say, a monthly basis, suppose that firms' announcement dates are uniformly distributed across calendar dates. Every month when the cash flow estimate gets updated, the same weight is assigned to firms announcing cash flows close to the cutoff date and firms whose announcement date happened almost one year previously. This weighting automatically makes the resulting growth estimate stale and also introduces serial correlation in the estimate.

[4]Cash flow news is often announced after the regular trading sessions in the stock market have closed and so aggregating across firms that announced cash flows within a 24-hour interval – as opposed to modeling, say, hourly cash flow news – seems appropriate.

[5]Announced dividends precede actual dividend payments by approximately 42 days, on average.

of dividend growth proposed by van Binsbergen and Koijen (2010) and Kelly and Pruitt (2013). We also find that our measure of the persistent dividend growth component is a positive and significant predictor of future growth in GDP and aggregate consumption. In sharp contrast, "raw" dividend growth, as well as the individual jump or transitory shock components, are very noisy and turn out not to have any predictive power over future dividend growth. Our results suggest that firms closely monitor the state of the economy and adjust their dividend policies in anticipation of changes in slow-moving economic indicators such as GDP and consumption growth.

A key limitation of empirical tests of asset pricing models is that while high-frequency data are available on movements in individual and aggregate stock prices, cash flows of individual firms are observed at much lower frequencies (typically quarterly). The absence of high-frequency cash flow data reduces researchers' ability to estimate and test asset pricing models which rely on the joint dynamics of stock prices, expected returns and cash flow growth expectations.

The second part of our paper addresses these challenges by using our new daily cash flow estimates to study predictability of stock returns in the context of the log-linearized present value approach developed by Campbell and Shiller (1988a). The present value model offers several advantages. As shown by Cochrane (2008) and van Binsbergen and Koijen (2010), the present value model implies a set of cross-equation restrictions that tie the coefficients of the variables in the prediction model for dividend growth to the coefficients on the corresponding variables in the return prediction model. In particular, any variable that helps predict future dividend growth should also, in general, have the ability to forecast stock returns, after controlling for predictor variables such as the log dividend-price ratio. We find strong empirical evidence that variation in the persistent dividend growth component is a strong and significant predictor of daily stock returns and that the cross-equation restrictions implied by the present value model are supported by the data. Moreover, we show that our new dividend growth measure can be used to predict stock returns several days into the future. At the monthly horizon, our persistent dividend growth measure provides more accurate out-of-sample forecasts of stock market returns than all but one of the fourteen predictor variables considered by Goyal and Welch (2008).

We also study whether the different components extracted from the dividend process

help explain concurrent (same-day) dynamics in stock returns. We find strong evidence that positive shocks to the persistent dividend growth component are associated with higher same-day mean stock returns, lower stock market volatility, and a reduced likelihood of outliers (jumps) in stock returns, again in a manner that is consistent with the cross-equation restrictions implied by the present value relation. Higher uncertainty about dividend growth translates into higher stock market volatility as well as a higher chance of a jump (outlier) in stock returns, suggesting that daily dividend news is an important driver of dynamics in aggregate stock market returns and is not just limited to affecting the mean of stock returns.

Our paper is related to a literature that estimates the effect of new information on stock prices on days with news releases using event-study methodology, see e.g., Cutler et al. (1989) and McQueen and Roley (1993). A limitation of this approach is that news stories are qualitative (e.g., "world events"), heterogeneous across news categories (news on housing starts versus monetary policy shocks), and, sometimes, rare (monetary policy shocks). In addition, the effect of these news may be state-dependent. For example, good news about the economy can be bad news to stock prices in a high-growth state (McQueen and Roley (1993)) and may depend on whether the economy is operating above or below trend. Macroeconomic news can also be difficult to decipher because they comprise a bundle of information about cash flows and expected returns (Boyd et al. (2005)).

The methodology developed in our paper is also related to papers in the asset pricing literature such as Chib et al. (2002) and Eraker et al. (2003a) which estimate models of stock return dynamics with stochastic volatility and jumps. There are three key differences between our approach and these papers. First, to the best of our knowledge, no existing study has attempted to model the high-frequency dynamics in dividends using such methods, let alone estimate and test a model as general as ours. Second, our paper tests asset pricing implications implied by the present value model, using a much higher data frequency (daily) than previously attempted. Third, our paper models the transmission from daily cash flow news–in the form of mean, volatility, or jump components–to contemporaneous dynamics in stock market returns, including variation in the volatility and jump probability of returns.

The outline for the paper is as follows. Section 2 introduces our data and explains how we construct a daily cash flow index from dividend announcement data. Section 3 explains our econometric modeling approach and reports empirical model estimates. Section 4 analyzes

dividend growth predictability. Section 5 develops the present value framework to explore the implications of predictability of dividend growth for return predictability, while Section 6 analyzes the effect of news about the dividend growth process on concurrent stock returns and Section 7 concludes. Additional technical details, extensions and robustness tests are reported in a set of web appendices.

2 Data

We start our analysis by explaining how we construct our daily dividend growth series and describing the data sources that we use. Our analysis of daily cash flows focuses on growth in dividends which, as pointed out by Kelly and Pruitt (2013), has been the focus of a large literature on asset pricing.[6] Because earnings can be negative, defining growth in earnings poses challenges that are quite different from those arising when studying dividends.

The biggest effect of dividend news on asset prices is likely to come through their information content, so we focus on dividends as initially *announced* as opposed to the actual dividend payments. However, in the web appendix we also undertake an analysis of daily dividends viewed from the perspective of the payment date which allows us to compare the information effect to the direct cash flow effect from dividend payments.

2.1 Sample construction

Our sample includes all ordinary cash dividends declared by firms with common stocks (share codes 10 and 11) listed on NYSE, NASDAQ, or AMEX from 1926 to 2016.[7] We do not exclude special dividends from our measure. For example, Costco (permno 87055), which usually pays around 30 cents per quarter in dividends, declared two "special dividends" during our sample: $7 per share on 28 November 2012 and $5 per share on 30 January 2015. Those two dividends, with CRSP codes 1232 and 1272, respectively, are included in our dividend measure. What is *not* included in our measure is non-ordinary dividends such as M&A cash flows, buybacks and new issues. We also require firms to have valid stock prices and a valid figure for the number of shares outstanding when dividends are announced.

[6] See, e.g., Campbell and Shiller (1988a), Cochrane (1992), Lettau and Ludvigson (2005), Koijen and Nieuwerburgh (2011), and Maio and Santa-Clara (2015).

[7] Ordinary cash dividends have CRSP distribution codes below 2000.

Furthermore, we make sure there are no duplicate observations in the dataset and that each firm pays only one dividend at any point in time.[8] Overall, our sample consists of 503,591 declared dividends.[9]

Corporate dividends have a strong firm-specific component and often display pronounced seasonal variation. Our analysis therefore computes dividend growth by comparing same-firm, same-(fiscal) quarter, year-on-year changes in cash flows. To this end, let D_t^i be the total dividends declared by firm i on day t, calculated as the dividend per share times the number of outstanding shares. Moreover, let I_t^i be an indicator variable that equals one if company i announces quarterly dividends on day t, and otherwise takes a value of zero, while $\tilde{t}$ is the associated same-quarter, prior-year dividend announcement date for firm i.[10] For example, a company may have declared dividends on May 17, 2014 while it declared the corresponding quarter's prior-year dividends on May 9, 2013, in which case t is May 17, 2014 and $\tilde{t}$ is May 9, 2013.

Aggregating across firms, the total dollar value of dividends paid out on day t is $\sum_{i=1}^{N_t} I_t^i D_t^i$, where N_t is the number of (publicly traded) firms in existence on day t. Similarly, the total value of dividends paid out by the *same* set of firms for the *same* fiscal quarter during the prior year is given by $\sum_{i=1}^{N_t} I_{\tilde{t}}^i D_{\tilde{t}}^i$. Taking the ratio of these two numbers, we obtain a measure of the aggregate, year-on-year (gross) growth in dividends:[11]

$$G_t = \frac{\sum_{i=1}^{N_t} I_t^i D_t^i}{\sum_{i=1}^{N_t} I_{\tilde{t}}^i D_{\tilde{t}}^i}. \tag{1}$$

Note that the number of firms used in this calculation – as well as their identity – changes on a daily basis and from year to year as firms shift their dividend announcement dates. Only firms for which $I_t^i = I_{\tilde{t}}^i = 1$ are included in this calculation, ensuring that the *same* firms

[8] There are instances in CRSP in which a company declares or pays multiple dividends on the same day, using different distribution codes but still classified as ordinary dividends. We aggregate such dividends to convert them into a single dividend. As an example, on November 23, 1983, PPL Corporation (permno 22517) declared two ordinary dividends of 39 and 21 cents.

[9] Following a recent update, CRSP no longer provides the dividend declaration date prior to 1962 and data until 1964 appear to be incomplete. Nonetheless, we also have an older version of the database in which the declared dividend dates start in 1926. As a consequence, we have 101,476 pre-1964 observations and 402,115 post-1964 observations.

[10] $\tilde{t}$ depends on the firm, i, and same-quarter, next-year date, t, so a more precise notation is $\tilde{t}(i,t)$.

[11] Only seven percent of individual firms' year-on-year dividend growth observations in our sample are constant, suggesting that firms often change their dividends, even marginally, every year.

are used in both the numerator and denominator of the ratio. Equation (1) uses the dollar amount paid in dividends by individual firms, implicitly applying value weights since large firms tend to have larger dividend payouts.[12]

As a first illustration of our data, Figure 1 provides a plot of the number of firms, as well as the dollar dividend and the (net) growth rate from equation (1) during a single quarter (Q2 2014). The top panel shows substantial intra-quarter variation in the number of firms announcing dividends. During this particular quarter, the maximum number of firms announcing dividends on any one day was 68 (on April 24), while the minimum number was one (on June 22). On several days more than 50 firms announced dividends.

The middle panel in Figure 1 shows the total value of dividends declared on individual days. This depends on the number and size of firms announcing dividends as large firms tend to announce bigger dividends.[13] Lastly, the bottom panel in Figure 1 shows the daily net dividend growth during the quarter. Peaks in this measure need not coincide with days where most firms announce dividends (top panel) or days on which the overall amount of dividends announced (middle panel) peaked. This is because the dividend growth rate depends on dividends announced by the same group of firms during the prior year as reflected in the denominator of equation (1). For example, the gross dividend growth rate on June 22 (1.15) is generated by a single firm announcing dividends on that day: the firm announced \$155m in dividends in Q2, 2014 and \$135m for Q2, 2013. Thus, variation in daily dividend growth

[12] An alternative approach that more explicitly accounts for heterogeneity in firm size is to first define individual firms' cash flow growth as

$$\tilde{G}_t^i = \begin{cases} \frac{D_t^i}{D_{\bar{t}}^i} & \text{if} \quad I_t^i = 1 \\ 0 & \text{otherwise} \end{cases} .$$

In a second step we can use individual firms' market capitalization to aggregate the cash flow growth rates across firms that pay dividends on day t :

$$\begin{aligned} \tilde{G}_t &= \sum_{i=1}^{N_t} I_t^i \omega_t^i \tilde{G}_t^i, \quad \text{where} \\ \omega_t^i &= \frac{MktCap_t^i \times I_t^i}{\sum_{i=1}^{N_t} MktCap_t^i \times I_t^i} \end{aligned}$$

is the weight on company i in the daily year-on-year value-weighted dividend growth calculation. Results based on this alternative measure, available in the internet appendix, are very similar to those based on the measure in equation (1).

[13] The largest amount of dividends declared during Q2 2014, \$7.12bn, happened on April 24, while only \$3.6m of dividends were announced on June 30.

rates reflects both heterogeneity across firms' dividend behavior and shifts in the number of firms announcing dividends on a given day.

2.2 Features of daily dividend growth

Our data span the period 1927-2016, but the first part of the sample is dominated by the Great Depression. For robustness, we therefore split the sample into halves and study both the full sample and the second half of the sample from 1973 to 2016. Figure 2 (top panel) plots $\Delta d_t = \ln(G_t)$ from 1973 to 2016. On days with no dividend announcements, we set the series to zero.[14] The daily dividend growth series is very spiky and dominated by days with unusually large positive or negative dividend growth. There is also evidence of a sustained decline in dividends during the financial crisis.

The features displayed by our daily series of year-on-year growth in dividends in Figure 2 can be summarized as follows: (i) the daily dividend growth series is very lumpy, reflecting variation, both, in individual firms' cash flow growth and in the composition of firms that, on any given day, announce dividends; (ii) daily dividend news is driven by a persistent component which is particularly pronounced during the financial crisis of 2008/09; (iii) the volatility of daily cash flow news changes over time with unusually calm periods interchanged with more volatile periods.

2.3 Comparison with the standard dividend growth measure

The conventional alternative to our bottom-up approach is to extract dividends top-down using market returns with and without dividends as published by the Center for Research in Security Prices (CRSP). Three limitations render this alternative approach unattractive.

First, the daily CRSP index accounts for dividends *distributed* on a particular day but does not show *when* those dividends were announced. This distinction is crucial as firms typically announce dividends several days prior to the payment date and it is the news effect of announced dividends that we would expect to be important for movements in stock prices.

Second, the set of firms announcing or paying dividends on any given day is generally different from the set of firms announcing dividends on the same day one year earlier. As a consequence, year-on-year estimates of dividend growth from daily values of the CRSP index

[14] This happens on 170 days, or 1.4% of the full sample.

are difficult to interpret as they do not control for firm fixed effects and so may get distorted due to changes in the composition of the set of firms paying dividends on a given calendar day, confounding dividend growth information with composition shifts.[15]

Third, the CRSP index contains assets such as ETFs and mutual funds (Sabbatucci (2017)) whose dividends follow a pattern different from that of individual firms.

These points turn out to make a crucial difference to daily dividend growth. To see this, the bottom panel of Figure 2 plots the daily dividend growth rate constructed using the conventional top-down CRSP approach over the sample 1973-2016.[16] While the daily dividend growth series based on our bottom-up methodology (top panel) is affected by occasional jumps, it clearly contains a persistent component linked to the state of the economy. In contrast, the daily dividend growth series constructed from the CRSP return indexes is very noisy throughout the sample.

3 Econometric model

To match the features of the daily dividend growth data noted above, we next develop an econometric model that incorporates multiple components to capture different parts of the dividend dynamics. We accomplish this as follows. First, we account for lumpiness by allowing for a jump component in daily cash flow growth. Since the lumpiness introduced by firm-level heterogeneity in dividend payments is more likely to be diversified away if a large number of firms announce dividends on the same day, we allow the jump intensity to depend on the number of firms announcing dividends. Second, we incorporate a persistent component in the mean growth equation. Third, we account for time-varying volatility by modeling ordinary shocks to daily dividend growth as a stochastic volatility process.

3.1 A components model for daily dividend growth

Our econometric model decomposes the daily dividend growth process, Δd_{t+1}, into three parts, namely (i) a persistent term, μ_{dt+1}, which captures a smoothly evolving mean

[15] Changes in the composition of the dividend-paying firms has a far smaller effect on dividend growth measured at longer time intervals such as a quarter or a year.

[16] To replicate our methodology as closely as possible, the daily dividend measure from CRSP shown here computes the growth rate as the change in aggregate dividends paid on a given day relative to aggregate dividends paid on the same day (or whichever day is closest) during the previous year.

component; (ii) a jump component, $\xi_{dt+1}J_{dt+1}$, where $J_{dt+1} \in \{0,1\}$ is a jump indicator that equals unity in case of a jump in dividends and otherwise is zero, while ξ_{dt+1} measures the magnitude of the jump; (iii) a temporary cash flow shock, ε_{dt+1}, whose volatility is allowed to be time-varying. Adding up these terms, we have

$$\Delta d_{t+1} = \mu_{dt+1} + \xi_{dt+1}J_{dt+1} + \varepsilon_{dt+1}. \tag{2}$$

We next motivate this decomposition and introduce our assumptions on the components.

As we reduce the time interval over which dividends are measured, the data become a lot scarcer, with sizeable day-to-day variation in both the number and types of firms that announce dividends. To minimize the effect of such composition shifts, our main empirical analysis focuses on an aggregate, market-wide measure of dividend growth which offers the broadest coverage of dividend events and minimizes the number of days with very few firms announcing dividends.

To understand how daily changes in the composition of firms announcing dividends can affect our dividend growth measure, suppose the data generating process for aggregate dividend growth, Δd_{At+1}, takes the form

$$\Delta d_{At+1} = \mu_{dt+1} + \sigma_A \varepsilon_{At+1}, \quad \varepsilon_{At+1} \sim \mathcal{N}(0,1), \tag{3}$$

where μ_{dt+1} is the conditional mean of the aggregate dividend growth process at time $t+1$ and σ_A is the volatility of the daily shocks, ε_{At+1}.

We capture any persistence that may be present in the mean of the dividend growth process by assuming that μ_{dt+1} follows a mean-reverting first-order autoregressive process

$$\mu_{dt+1} = \mu_d + \phi_\mu (\mu_{dt} - \mu_d) + \sigma_\mu \varepsilon_{\mu t+1}, \quad \varepsilon_{\mu t+1} \sim \mathcal{N}(0,1), \tag{4}$$

where $|\phi_\mu| < 1$ and $\varepsilon_{\mu t+1}$ is assumed to be uncorrelated with ε_{At+1}. In the special case where $\phi_\mu = 0$, d_{At+1} follows a random walk process whose changes, Δd_{At+1}, are unpredictable.

In practice, we do not observe the aggregate dividend process at the daily frequency. Rather, we observe the dividends announced by a subset of firms on a given day. Shifts in the composition of firms announcing dividends on any given day can introduce measurement issues and must be carefully dealt with when modeling the dynamics in the observed dividend

process. To see how changes in the composition of dividend-announcing firms affect our dividend growth measure, let Δd_{it+1} be the observed value of firm i's year-on-year dividend growth rate on day $t+1$. We decompose this into a systematic component that is driven by firm i's cash flow beta (β_i) times the growth in aggregate cash flows, $\beta_i \Delta d_{At+1}$, and an idiosyncratic component, ε_{it+1}, assumed to be uncorrelated across firms:

$$\Delta d_{it+1} = \beta_i \Delta d_{At+1} + \sigma_i \varepsilon_{it+1}, \quad \varepsilon_{it+1} \sim \mathcal{N}(0,1). \tag{5}$$

Let ω_{it} be the weight on the dividends announced by firm i on day t, measured relative to the total amount of dividends paid out by all firms on this day, so $\omega_{it} = 0$ if firm i does not announce dividends on day t and $\sum_{i=1}^{N_{dt}} \omega_{it} = 1$ on all days, where N_{dt} is the number of firms announcing dividends on day t. Our measure of (announced) dividend growth on day $t+1$, Δd_{t+1}, aggregated across all dividend-announcing firms, can then be computed as

$$\begin{aligned}
\Delta d_{t+1} &= \sum_{i=1}^{N_{dt+1}} \omega_{it+1} \Delta d_{it+1} = \sum_{i=1}^{N_{dt+1}} \omega_{it+1} \left[\beta_i \Delta d_{At+1} + \sigma_i \varepsilon_{it+1}\right] \\
&= \left[\sum_{i=1}^{N_{dt+1}} \omega_{it+1} \beta_i\right] \Delta d_{At+1} + \left[\sum_{i=1}^{N_{dt+1}} \omega_{it+1} \sigma_i \varepsilon_{it+1}\right] \\
&\equiv \beta_{t+1} \Delta d_{At+1} + \sigma_{dt+1} \varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim \mathcal{N}(0,1).
\end{aligned} \tag{6}$$

where $\beta_{t+1} \equiv \sum_{i=1}^{N_{dt+1}} \omega_{it+1} \beta_i$ is the weighted average of cash-flow betas of firms announcing dividends on day $t+1$, and $\sigma_{dt+1} = \left[\sum_{i=1}^{N_{dt+1}} \sigma_i^2 \omega_{it+1}^2\right]^{1/2}$. If all stocks have identical cash flow betas, β_{t+1} in (6) will be constant over time and equal to the average beta, $\bar{\beta}$. Even if there is heterogeneity in firms' cash flow betas, provided that the number of firms announcing dividends each day is large, the weights on individual firms (ω_{it+1}) are small, and the subset of firms announcing dividends on a given day is reasonably random over time, we would expect little time variation in β_{t+1}. This makes it easy to extract a daily estimate of the aggregate dividend growth rate from equation (6) by means of filtering methods.

The number of firms announcing dividends can be low on some days and so the announced dividends can be dominated by a few companies or by companies within the same industry. If cash flow betas vary systematically across industries, this can lead to time variation in the cash flow beta of the subset of dividend-announcing firms.

While these composition effects can lead to time variation in the cash flow beta of the

observed dividend growth process, Δd_{t+1}, such effects are temporary and so $\beta_{t+1} - \bar{\beta}$ will be mean reverting towards zero. Using (6), we can therefore write

$$\begin{aligned} \Delta d_{t+1} &= \bar{\beta}\Delta d_{At+1} + (\beta_{t+1} - \bar{\beta})\Delta d_{At+1} + \sigma_{dt+1}\varepsilon_{t+1} \\ &= \bar{\beta}\left[\mu_{dt+1} + \sigma_A \varepsilon_{At+1}\right] + (\beta_{t+1} - \bar{\beta})\left[\mu_{dt+1} + \sigma_A \varepsilon_{At+1}\right] + \sigma_{dt+1}\varepsilon_{t+1} \\ &\propto \mu_{dt+1} + \frac{(\beta_{t+1} - \bar{\beta})}{\bar{\beta}}\mu_{dt+1} + \left[1 + \frac{\beta_{t+1} - \bar{\beta}}{\bar{\beta}}\right]\sigma_A \varepsilon_{At+1} + \frac{\sigma_{dt+1}}{\bar{\beta}}\varepsilon_{t+1}. \end{aligned} \tag{7}$$

Compared to the aggregate dividend process in (3), there are some additional time-varying terms in the observed dividend growth rate: (i) $\frac{(\beta_{t+1}-\bar{\beta})}{\bar{\beta}}\mu_{dt+1}$; this term can be quite volatile because of random variation in the first part, $(\beta_{t+1}-\bar{\beta})/\bar{\beta}$, driven by the process determining the number and types of firms announcing dividends on any given day. This term tends to be particularly volatile on days where few firms announce dividends, increasing the probability that $|\beta_{t+1}-\bar{\beta}|/\bar{\beta}$ will be large; (ii) $\frac{\beta_{t+1}-\bar{\beta}}{\bar{\beta}}\sigma_A\varepsilon_{At+1}$; this is the product of a random component $(\beta_{t+1} - \bar{\beta})/\bar{\beta}$ and an uncorrelated temporary shock, ε_{At+1}; (iii) $(\sigma_{dt+1}/\bar{\beta})\varepsilon_{t+1}$. The last two components can introduce time-varying and persistent volatility in Δd_{t+1} if firms in certain industries tend to cluster their dividend announcement dates around certain dates and if the volatility of the idiosyncratic dividend growth component differs across industries, as we might expect.

Our model in (2) captures the effect of these components in Δd_{t+1} in two ways. First, the jump component, $\xi_{dt+1}J_{dt+1}$, can handle temporary composition shifts that lead to large variation in the daily cash flow beta of our composite dividend growth measure. It can also be used to absorb outliers in company-specific dividend shocks, ε_{it+1}, on days where individual firms dominate dividend announcements. Because the magnitude of this component is likely to be larger on days with few firms announcing dividends, we allow the jump probability to depend on the number of announcers, N_{dt+1}:

$$\Pr\left(J_{dt+1} = 1\right) = \Phi\left(\lambda_1 + \lambda_2 N_{dt+1}\right), \tag{8}$$

where Φ is the CDF of a standard normal distribution. The magnitude of the jumps is modeled as $\xi_{dt+1} \sim \mathcal{N}\left(0, \sigma_\xi^2\right)$.

Second, our model captures time-varying heteroskedasticity in Δd_{t+1} by modeling the

variance of the residuals in the dividend growth process as a stochastic volatility process:

$$\varepsilon_{dt+1} \sim \mathcal{N}(0, e^{h_{dt+1}}), \tag{9}$$

where h_{dt+1} is the log-variance of ε_{dt+1} which is assumed to follow a mean-reverting process

$$h_{dt+1} = \mu_h + \phi_h \left(h_{dt} - \mu_h\right) + \sigma_h \varepsilon_{ht+1}, \quad \varepsilon_{ht+1} \sim \mathcal{N}(0, 1), \tag{10}$$

where ε_{ht+1} is uncorrelated with both ε_{dt+1} and $\varepsilon_{\mu t+1}$. Stochastic volatility in the dividend growth process (7) may arise due to the clustering of days with firms from similar industries (e.g., financial institutions) announcing dividends.

To summarize, our dividend growth model in (2) accounts for a persistent mean-reverting component, time-varying volatility, and jumps. We evaluate the importance of these features by comparing results from the general model in (2) to a simpler (no-jump) model that ignores jump dynamics and stochastic volatility and so takes the form

$$\Delta d_{t+1} = \mu^{NJ}_{dt+1} + \varepsilon_{dt+1}, \quad \varepsilon_{dt+1} \sim \mathcal{N}(0, \sigma_d^2), \tag{11}$$

where μ^{NJ}_{dt+1} follows the process in (4). This comparison allows us to gauge the importance of incorporating jump dynamics and stochastic volatility for our estimate of μ_{dt+1}.

3.2 Estimation

We adopt a Bayesian approach that uses Gibbs sampling to estimate the model parameters. Details of our estimation procedure are provided in Internet Appendix A while Internet Appendix B documents the convergence properties of our estimation algorithm.

It is worth briefly describing the priors that underlie our model. We choose standard normal-gamma conjugate priors which simplify the process of drawing from the conditional distributions of the model parameters in the Gibbs samplers. Moreover, we specify independent priors for the parameters of both the mean, variance, and jump processes. For almost all of the prior hyperparameters, we use loose and mildly uninformative priors. The main exceptions are the persistence parameters, ϕ_μ and ϕ_h, whose priors we center on 0.98. Further details are provided in Internet Appendix A.

3.3 Empirical estimates

We next present estimates of the parameters of our econometric model and evaluate empirically the behavior of the three components in the dividend growth process.

Table 1 presents parameter estimates for our general dividend growth model in equations (2)-(10) for both the short sample (1973-2016) and the full sample (1927-2016). We focus our discussion on the parameter estimates for the short sample but note that the estimates for the longer sample are very similar.

First consider the parameters determining the mean of the dividend growth process in (4). The long-run mean estimate, $\mu_d = 0.080$, corresponds to an 8% annualized nominal dividend growth rate which is close to the mean of the standard dividend growth measure (extracted from CRSP data) of 7.8% over the same sample period – a figure well within the 90% credible set of [0.062, 0.097]. The persistence parameter for the mean reverting component of the dividend growth process, ϕ_μ, is centered on 0.998, corresponding to a half-life of 350 trading days. While highly persistent, shocks to the mean-reverting component (4) are very small as shown by the estimate $\sigma_\mu = 0.002$. Our model thus identifies a small, highly persistent component in the dividend growth process.

The top and bottom panels in Figure 3 plot the persistent dividend growth component, μ_{dt}, extracted using either the no-jump model (equation 11, top panel) or the general jump model (equation 2, bottom panel). The μ_{dt}^{NJ} series erroneously assigns large daily spikes in the observed series to the persistent component, μ_{dt}. In contrast, the jump model succeeds in separating temporary spikes (noise) in the daily dividend series from the persistent component μ_{dt} which, consequently, is far smoother. Indeed, values of the persistent dividend growth component extracted from the general model fall on a far narrower scale than the unfiltered dividend growth series, ranging from just below zero to 0.15. The financial crisis in 2008-09 is associated with a notable drop in the persistent dividend growth component which, for the only time in our sample, turns negative, followed by a notable bounce-back in the second half of 2009 and 2010.

Our empirical analysis uses the dividend growth measure in (1), but one could alternatively have chosen to use a market (cap-) weighted growth measure. To briefly address whether our choice of dividend growth measure makes a difference to the estimates of our components model, we construct the daily dividend series as described in

footnote 12 and plot the resulting estimate of μ_{dt} along with our original estimate in Figure 4. The two series are virtually indistinguishable with a correlation of 98.11%, so the dividend weighting scheme makes little or no difference to our results.

To empirically quantify the importance to our estimates of μ_{dt} of including days with very few firms announcing dividends – presumably days with a higher likelihood of being dominated by "noise" due to the potential for large composition shifts – we re-estimate our model using only observations on dividend growth on days with at least five announcing firms, a requirement that leads us to drop around 10% of the original sample. The correlation between our original μ_{dt} series (constructed without requiring at least five announcing firms) and this new estimate of μ_{dt} (constructed with at least five announcing firms) on days where they overlap is 99.64%, so our model seems to succeed in isolating the effect of outliers in the dividend growth data caused by a scarcity of firms announcing dividends. Allowing for jumps in the dividend growth process is key to ensuring this robustness property. To see this, we extracted μ_{dt} in a similar way, using only days with at least five announcing firms, but now for a model without jumps and stochastic volatility. Comparing this series to an estimate from the same model constructed without the requirement of at least five announcing firms, the correlation is now reduced to only 72%, suggesting that the no-jump model is far more strongly affected by days with few firms announcing dividends.

Our estimates of equation (10) show that the stochastic volatility process is moderately persistent with an autoregressive parameter, ϕ_h, whose estimated mean is 0.833.

Finally, consider the jump component extracted from our model. The negative estimates of the jump intensity parameters (λ_1 and λ_2) imply that a jump occurs every sixty days on average with lower jump probabilities on days where a large number of firms announce dividends. On many days in the sample, the jump probability indicator, J_{dt}, shown in the top panel of Figure 5, is close to one. On such days, spikes in daily dividend growth are attributed mostly to jumps rather than to clusters with high volatility from the transitory component, ε_{dt}, in equation (2). Jumps can be very large in magnitude, as shown in the bottom panel of Figure 5, which displays the estimated jump size, ξ_{dt}. Indeed, the estimated standard deviation of the jump size (σ_ξ) has a mean of 2.761 which is almost four times larger than the estimated mean of σ_h (0.75). Shocks to daily dividend growth originating from the jump component thus tend to be far bigger than the regularly occurring ε_{dt} shocks.

To get a sense of how sensitive the dividend growth jump probability is to the number of firms announcing dividends on a given day, N_{dt}, Figure 6 plots the jump intensities for three values of N_{dt} chosen to match the 25th, median and 75th percentiles of the distribution of the daily number of announcing firms. On days with a large number of announcing firms (75th percentile, or 36 firms on average), the jump intensity distribution is centered on a number a little over 0.005, corresponding to a jump on average every 200 days. On days with a typical number of announcing firms (median, or 22 firms), the jump intensity is centered around its average value near 0.016, implying a jump every 60 days. Finally, on days with a small number of announcing firms (25th percentile, or 12 firms), the probability of a jump is centered just below 0.03, corresponding to a jump in dividend growth every 35 days.

3.4 Firm heterogeneity and composition effects

Earlier studies have noted evidence of heterogeneity in the cash flow process across firms with different characteristics. For example, Vuolteenaho (2002) and Cohen et al. (2009) link firms' book-to-market ratios to cross-sectional variation in their cash flow betas, finding that value stocks have higher cash flow betas than growth stocks.

To address concerns about heterogeneity in cash flow betas caused by daily shifts in the composition of the subset of dividend-announcing firms, we construct a number of daily dividend series that only include data on firms within specific sectors or firms with certain book-to-market ratio or size characteristics.[17] In turn, we apply our methodology separately to each of these disaggregate series and extract a persistent dividend growth component. Provided that cash-flow betas are more homogeneous within each of these groups, composition effects due to daily variation in the set of dividend-announcing firms matter less.

Figure 7 shows the μ_{dt} estimates obtained from subsets of firms sorted on high versus low book-to-market ratios (top panel) or firms sorted into small, medium, and large size categories (bottom panel). These plots confirm that there are cross-sectional differences in the behavior of the persistent dividend components (μ_{dt}) extracted from our model.

[17]For example, each month we sort all firms in our sample by their book-to-market (BM) ratios and form two portfolios consisting of firms whose BM ratios are above ("high") or below ("low") the median BM ratio. To be included in the sort in a given month, a stock needs to have reported book and market values and to be traded on the NYSE, AMEX or NASDAQ.

Specifically, the persistent dividend growth component displays much stronger time-series variation for value firms than for growth firms (top panel) and for small firms compared with large or medium-sized firms (bottom panel). Comparing the individual μ_{dt} series with our aggregate, market-wide estimate of μ_{dt}, we find correlations of 97.80% for large firms, 88.38% for medium firms, 71.67% for small firms, 90.40% for firms with high book-to-market ratios, and 92.89% for firms with low book-to-market ratios.

In Internet Appendix C, we also estimate the persistent dividend growth component (μ_{dt}) for the five Fama-French industries and compute parameter estimates for our decomposition applied to these series. The parameter estimates and time-series plots show that the basic features of the dividend process remain the same across very different industries.

These results demonstrate that while the predictable dividend growth component varies across firms with different characteristics and in different industries, some features are common across firms: estimates of ϕ_μ close to unity (indicating a highly persistent component in dividend growth), similar jump sizes, and negative dependence between the jump probability and the number of firms announcing dividends on a given day.

4 Predictability of dividend growth

Predictability of dividend growth features prominently in discussions of asset pricing models. Cochrane (2008) finds little evidence of predictability of US dividends, while studies such as van Binsbergen and Koijen (2010), Kelly and Pruitt (2013), Jagannathan and Liu (2019) argue that dividend growth is, to some extent, predictable.[18] The parameter estimates from our dividend model show that the daily dividend growth process contains a small, but highly persistent component and this section explores the implications of our results for predictability in dividend growth.

4.1 Predictive regressions

Existing studies on dividend growth predictability use time-aggregated dividends measured over longer horizons than our daily interval. To explore whether our estimate of the

[18] A recent literature uses dividend futures to estimate the term structure of dividends. van Binsbergen et al. (2012) and van Binsbergen and Koijen (2017) recover prices of dividend strips and show that their expected returns are higher than those on the underlying index.

persistent dividend growth component can predict dividends–and to make our results directly comparable to existing ones–we use the conventional top-down approach to construct monthly and annual measures of dividend growth from CRSP data, denoted Δd_t^{CRSP}.[19]

Next, we estimate a predictive regression of future dividend growth, Δd_{t+1}^{CRSP}, on the persistent dividend component measured at the end of the previous period, μ_{dt}, the log dividend-price ratio extracted from CRSP, dp_t, and current and lagged dividend growth:

$$\Delta d_{t+1}^{CRSP} = \alpha + \beta \mu_{dt} + \gamma dp_t + \sum_{j=1}^{3} \rho_j \Delta d_{t+1-j}^{CRSP} + \varepsilon_{t+1}. \tag{12}$$

We include the log dividend-price ratio in the regression because this has been suggested as a predictor of cash flow growth in a variety of studies (e.g., Cochrane (1992), Cochrane (2008), and Lettau and Nieuwerburgh (2008)).

Panel A of Table 2 shows that the persistent component of dividend growth, μ_{dt}, has strong predictive power over future dividend growth recorded at the quarterly horizon. In the shorter post-1973 sample, the lagged persistent dividend growth component obtains a t-statistic of 4.6 after accounting for the effect of lagged dividend growth and the lagged dividend-price ratio. Moreover, at 0.26 the R^2 is quite high. These findings are robust to the sample period. Starting the sample in 1927, the coefficient on μ_{dt} obtains a t-statistic of 2.5 and the R^2 value of the predictive regression rises to 0.37. The coefficient on the lagged dividend-price ratio is not significant in any of these regressions, while the first two lags of dividend growth are significant in the quarterly models.

The predictive power of μ_{dt} over future dividend growth is very similar at the annual and quarterly horizons in the post-1973 sample, though somewhat weaker at the annual frequency in the longer sample that begins in 1927. Still, μ_{dt} remains highly statistically significant at the annual horizon for both samples and this result is again robust to the presence of lagged dividend growth and the dividend-price ratio in the predictive regression.

[19] Most researchers extract aggregate dividends, D_t, from CRSP as the difference between the cum-dividend return (VWRETD), R_t^{cum}, and the ex-dividend return (VWRETX), R_t^{ex}, multiplied by the previous ex-dividend index level, P_{t-1}^{ex}, i.e., $D_t = (R_t^{cum} - R_t^{ex}) \times P_{t-1}^{ex}$. Using the resulting aggregate dividend series, the log dividend growth rate can be computed as $\Delta d_t^{CRSP} = ln\left(\frac{D_t}{D_{t-1}}\right)$.

4.2 Comparison with alternative predictors of dividend growth

Previous studies have reported evidence of predictability in dividend growth. Campbell and Shiller (1988b) develop a vector autoregressive (VAR) model for the joint dynamics in dividend growth, the price-dividend ratio and the cyclically adjusted price-earnings ratio (CAPE). Lags of all three variables obtain significant coefficients in the dividend growth equation, suggesting significant in-sample predictability of dividend growth at the annual frequency over the period 1871-1987. Similarly, Jagannathan and Liu (2019) propose a three-variable VAR that includes dividend growth, the dividend-earnings (payout) ratio, and a latent component that can capture the long-run risk surrounding future dividend growth. Although their analysis focuses on large stocks in the S&P500 index, their findings on predictability of annual dividend growth are commensurable with ours. Finally, Donaldson and Kamstra (1996) propose a neural net model to capture non-linear predictability in discounted future dividend growth. Computing stock prices as the present value of expected discounted future dividends, they can match the boom and crash in stock prices experienced during the Great depression.

The two studies most closely related to ours are van Binsbergen and Koijen (2010) and Kelly and Pruitt (2013). For example, van Binsbergen and Koijen (2010) use a latent variable Kalman filtering approach to estimate a log-linearized present value model consisting of expected returns and expected dividend growth rates for the aggregate stock market and identify a predictable component in dividend growth.

Kelly and Pruitt (2013) assume that individual firms' stock returns and log cash flow growth rates are a linear function of a set of unobserved common factors which can be estimated using a three-pass regression (partial least squares) methodology. In turn, cash flow growth can be projected on the common factors to generate a dividend growth forecast. Empirically, Kelly and Pruitt (2013) find strong in-sample evidence of predictability in annual cash flow growth while their out-of-sample results are somewhat mixed; in the Depression-era (1930-1940), dividend growth appears to be highly unstable and hard to predict while out-of-sample predictability is stronger over the sample 1940-2010.

We next compare our dividend growth estimates to results based on these two approaches.[20] To this end, the top panel in Figure 8 plots realized values of annual

[20]We are grateful to Seth Pruitt for sharing data and computer code which allowed us to replicate the

dividend growth against the persistent growth component estimated from our model, μ_{dt}, (sampled annually) and the van Binsbergen and Koijen (2010) measure, g_t^{VBK}. The bottom panel plots monthly dividend growth against our persistent dividend growth series, μ_{dt}, (sampled monthly) and the Kelly and Pruitt (2013) estimate, g_t^{KP}. While these dividend growth estimates are clearly correlated, there are also some notable differences. Notably, our persistent dividend growth measure shows a sharper decline during the global financial crisis compared with the two alternative estimates.

Panel B in Table 2 reports more formal results from regressions of the observed future dividend growth on the growth estimate implied by the three approaches. Because van Binsbergen and Koijen (2010) study cash-reinvested, annual dividend growth while Kelly and Pruitt (2013) use monthly dividend growth extracted from CRSP, their growth estimates are not directly comparable. We therefore report separate results for the annual and monthly series used in the two studies. To make comparisons between the three approaches easier to interpret and to relate the results to other parts of our paper, we use the 1973-2016 sample adopted throughout much of our analysis.

All three growth estimates clearly have predictive power over future dividends in the univariate regressions. For example, the growth estimate of van Binsbergen and Koijen (2010) obtains a t-statistic of 2.16 with an R^2 value of 13% in the annual sample. In comparison, the t-statistic on our μ_{dt} estimate is 4.04 and the associated R^2 value is 25%. Including both the μ_{dt} and g_t^{VBK} measures in the regression, we continue to observe a large t-statistic on μ_{dt} (3.97), while the t-statistic on the estimate of van Binsbergen and Koijen (2010) drops to 0.91 with an R^2 value of 27%. This demonstrates the predictive power possessed by our estimate of the persistent growth component.

In the monthly dividend growth regressions, the growth estimate of Kelly and Pruitt (2013) generates a t-statistic of 6.41 and an R^2 value of 20%. In comparison, the t-statistic obtained when instead we use our μ_{dt} measure is 7.23 and the R^2 value is 31%. Including both μ_{dt} and g_t^{KP} as predictors in the regression, μ_{dt} obtains a t-statistic of 5.12 while the t-statistic of the growth estimate of Kelly and Pruitt (2013) declines to 1.58.

These results show that the persistent component in dividend growth extracted from daily dividend announcements possesses strong predictive power over actual dividend growth

results in Kelly and Pruitt (2013).

at both the monthly and annual frequencies. Moreover, our μ_{dt} estimate adds substantial predictive power to existing dividend growth estimates.

To formally test and compare the predictive power of the three dividend growth estimates, we run a pair of forecast encompassing regressions:

$$\begin{aligned} \Delta d_{t+1}^{CRSP} &= \alpha + \beta_1 \mu_{dt} + (1-\beta_1) g_t^{VBK} + \varepsilon_{t+1}, \\ \Delta d_{t+1}^{CRSP} &= \alpha + \beta_1 \mu_{dt} + (1-\beta_1) g_t^{KP} + \varepsilon_{t+1}. \end{aligned} \tag{13}$$

The larger is β_1, the greater the weight on our dividend growth estimate and the smaller the weight on the competing model estimate. Specifically, a value of $\beta_1 = 1$ suggests that μ_{dt} dominates (encompasses) either g_t^{VBK} (top regression) or g_t^{KP} (bottom regression).[21]

The top row of Panel C in Table 2 shows that the estimate of β_1 equals 0.87 in the encompassing regression that includes μ_{dt} and g_t^{VBK}. The persistent dividend growth estimate from our model thus obtains a weight of 87% while the weight on the van Binsbergen and Koijen (2010) estimate equals 13%. Moreover, the estimated weight on μ_{dt} is statistically significant at the 1% level. In the second regression, the weight on μ_{dt} is 108% which is significantly different from zero, while the weight on the Kelly and Pruitt (2013) dividend growth estimate is -8% which is not significantly different from zero.

We conclude from these regressions that our dividend growth estimate adds significant predictive value to state-of-the-art measures proposed in the existing finance literature.

4.2.1 Out-of-sample forecasts

The analysis performed so far uses data up to 2016 and so does not address whether our approach could have been used in real time to generate accurate forecasts of dividend growth. To address this point, we conduct an out-of-sample forecasting experiment that only uses historically available data to estimate the parameters of our model and generate forecasts. Specifically, using an expanding estimation window and an initial warm-up period from 1973-1986, we re-estimate our model every week and construct real-time, out-of-sample daily forecasts over the 1986-2016 sample. Each month, we then take the last daily value of our forecast and compare this to the monthly out-of-sample forecasts generated using the approach of Kelly and Pruitt (2013). Finally, we evaluate the accuracy of these forecasts

[21] Note that $g_t = E_t \Delta d_{t+1}$ is the forecast of dividend growth next period.

using either univariate regressions or the forecast encompassing regressions in (13).

A univariate regression of dividend growth on our out-of-sample, real-time estimate, μ_{dt}, generates an R^2 of 27.76% (t-stat of 6.70), while the same regression using the out-of-sample forecast of Kelly and Pruitt (2013) produces an R^2 of 2.88% (t-stat of 2.38).[22] Similarly, in the forecast encompassing regression, the weight on the μ_{dt} series increases from 1.08 to 1.25 with a t-statistic of 12, while the weight on the Kelly and Pruitt (2013) forecast is -0.25. Our dividend growth forecasts thus perform very well even out-of-sample.[23]

To better assess the periods in which our $\Delta\mu_{dt}$ measure produces more accurate forecasts than existing alternatives, we follow Goyal and Welch (2008) and inspect the out-of-sample cumulative sum of squared forecast error differentials (CSSED) from a forecasting model based on $\Delta\mu_{dt}$ versus the Kelly-Pruitt dividend growth measure. Using $\Delta\mu_{dt}$ leads to more accurate forecasts than the Kelly-Pruitt model between the start of the sample (1973) and 1980, while our model underperforms during a brief spell in the late 1990s. During the 16-year period after 2000, our $\Delta\mu_{dt}$ measure consistently produces more accurate forecasts of dividend growth than the forecasting model based on the Kelly-Pruitt measure.

4.3 Data frequency and dividend dynamics

Prior work on predictability of dividend growth has focused on long horizons using monthly, quarterly, or annual data, while our results pertain to the daily frequency. To more formally evaluate the role of the data frequency in our analysis, we estimate our components model using dividend data at the longer frequencies, extract the corresponding estimates of the persistent growth component, μ_{dt}. and replicate the analysis in Panel A of Table 2. Figure 9 shows that the dividend growth estimates extracted from the lower frequency data possess many of the features of the daily estimates. However, they also tend to be very smooth compared to the estimates of μ_{dt} extracted from daily data. This means that the lower frequency estimates miss some of the important troughs and peaks in the dividend growth process. For example, the severity of the fall in dividend growth during the global financial crisis is completely missed by the quarterly and annual estimates, although the monthly

[22] Results ending in 2007, excluding the recent financial crisis, are qualitative similar with an R^2 of 25.47% (t-stat of 9.01) for our measure, and an R^2 of 4.05% (t-stat of 2.15) for Kelly and Pruitt (2013).

[23] Since our out-of-sample forecasts are generated recursively, our findings are not overly sensitive to the Great Recession which produced a considerable amount of cash flow news, see, e.g., Campbell et al. (2013).

estimate fares better in this regard. Table 3 shows that our daily μ_{dt} estimate outperforms the monthly, quarterly and yearly measures in terms of predictive accuracy, generating higher R^2 values even when the basis for the comparison is the monthly, quarterly, and annual forecast horizons used to estimate the lower-frequency measures.

Next, we perform forecast encompassing regressions, regressing monthly, quarterly, or annual dividend growth on the daily growth estimate plus one of either the monthly, quarterly, or annual growth estimates subject to the constraint that the coefficients ("weights") sum to unity. At all three horizons, we find that our daily dividend growth estimate gains a substantially larger weight than the corresponding low-frequency dividend growth estimates.

These results show that the data frequency is important to our findings and that applying our decomposition on daily data yields stronger prediction results than applying it on data recorded at lower frequencies. The reason for this finding is that the daily dividend growth estimate better captures shifts in the momentum of dividend growth which tends to be smoothed out in time-series averages based on longer periods.

The twin effects of using daily data on growth in announced dividends *and* decomposing changes in this series into a smoothly evolving persistent component, temporary jumps and (small) shocks with time-varying volatility is what combines to yield better forecasts of dividend growth. So *both* the higher data frequency *and* our decomposition approach help produce more accurate forecasts of dividend growth.

Past studies suggest that dividends are sticky as firms are cautious in raising dividends and very hesitant in reducing them (e.g., Brav et al. (2005)). This behavior might imply that the extent to which dividend growth is predictable differs on the upside and downside. To explore this point, we extend the regression in the first column of Table 3 to include μ_{dt} interacted with top and bottom quintile dummies defined according to whether the prior-month dividend growth rate was in the highest or lowest 20% of historical dividend growth rates, respectively. We find that, indeed, the slope coefficient on μ_{dt} increases significantly from 0.17 on average to 0.28 during downside months, indicating downside stickiness.

Our analysis of predictive dynamics in dividend growth is based on announced dividends rather than the commonly used paid-out dividends which lag dividend announcements. To shed light on the importance of this distinction, we sample both

announced and paid dividends at the monthly, quarterly and annual horizons and estimate our components model separately on these series. Next, we produce forecasts of dividend growth for the ensuing period. We find that predictability of dividend growth based on the forecasts that use announced dividends is stronger than if we use forecasts based on paid dividends. Interestingly, some advantage from using announced rather than distributed dividends to generate forecasts remain even at the annual horizon. Further details are provided in a web appendix.

4.4 Cash flow news and economic activity

We next examine the relation between our estimate of the persistent component of dividend growth news and two measures of macroeconomic growth, namely GDP and consumption growth, both of which have been examined by authors such as Liew and Vassalou (2000) and Bansal and Yaron (2004).[24] Specifically, to evaluate the statistical significance of these relations, we estimate predictive regressions

$$\Delta y_{t+1} = \alpha + \beta_1 \mu_{dt} + \beta_2 \Delta y_t + \varepsilon_{t+1}, \tag{14}$$

where Δy_{t+1} is the future change in either log GDP or log consumption. We include one lag of the dependent variable, Δy_t, to control for persistence in consumption or GDP growth. Table 4 reports the results from the regression in (14). First consider the results based on the estimate of μ_{dt} extracted from the daily data (Panel A). In the univariate regressions, our persistent dividend growth measure, μ_{dt}, generates positive coefficients of 0.15 with t-statistics of 5.68 and 6.58 for GDP growth and consumption growth, respectively. Moreover, with R^2 values of 23.5% and 31.1%, μ_{dt} clearly has strong predictive power over future GDP and consumption growth. The coefficient on μ_{dt} remains statistically significant once we add a lag of the dependent variable, although the t-statistic declines to a value around three.

To evaluate the importance of the data frequency, we also show results for cases where μ_{dt} is extracted from either monthly dividend growth data (Panel B) or quarterly dividend growth data (Panel C). While the μ_{dt} estimates extracted from the lower frequency data

[24]The Gross Domestic Product series is downloaded from FRED and is seasonally adjusted. Consumption expenditures are the sum of non-durable consumption plus services from Table 2.3.5 of the National Income and Product Accounts (NIPAs) from the Bureau of Economic Analysis (BEA) website.

remain statistically significant in helping predict both GDP and consumption growth, the R^2 values decline notably–particularly in the quarterly data for which we obtain R^2 values of 5.9% and 9.4%, compared with 23.5% and 31.1% for the daily data. High-frequency data can, thus, significantly boost the predictive power of our dividend growth estimate.

We conclude from this evidence that our persistent cash flow measure μ_{dt} helps predict variation in macroeconomic growth. This is consistent with our earlier finding that μ_{dt} predicts future dividend growth and shows that this result carries over to broader measures of economic growth. Our findings indicate that firms adjust their dividend payments in anticipation of changes in economic conditions measured by slow-moving economic indicators such as GDP and consumption growth.

4.5 Relation to other activity measures

Our daily dividend growth measure reflects general macroeconomic conditions and so can be viewed as an economic indicator similar to existing measures such as the macroeconomic uncertainty measure of Jurado et al. (2015), the economic policy uncertainty measure of Baker et al. (2016), the ADS business conditions index of Aruoba et al. (2009), the credit spread indicator of Gilchrist and Zakrajsek (2012), and "noise" in the Treasury market (Hu et al. (2013)).[25] Previous research has addressed whether these measures can be used to predict the state of the economy, especially during recessions and financial crises, so we next explore the relation between μ_{dt} and these alternative measures.

Panel A of Table 5 shows estimates of the correlations between the persistent dividend component μ_{dt} and these daily measures of financial and macroeconomic conditions. μ_{dt} has a highly significant negative correlation of -0.54 with the VIX, suggesting that dividend growth is lower in times with high uncertainty, which tends to coincide with economic recessions. μ_{dt} also has a significantly negative correlation of -0.20 with the policy uncertainty index of Baker et al. (2016) and a negative correlation of -0.51 with the liquidity noise index of Hu et al. (2013), indicating that firm payouts are lower in times with greater uncertainty.

[25] Aruoba et al. (2009) measure economic activity at the daily frequency using a variety of stock and flow data observed at mixed frequencies. Their approach extracts the state of the business cycle from a latent factor that affects all observed variables. Jurado et al. (2015) provide econometric estimates of time-varying macroeconomic uncertainty and show that important uncertainty episodes appear far more infrequently than indicated by popular uncertainty proxies. However, when such episodes do occur, they tend to be larger, more persistent, and more correlated with real economic activity.

Finally, μ_{dt} has a significantly positive correlation of 0.20 with the ADS index of Aruoba et al. (2009) and with the daily inflation index of Cavallo and Rigobon (2016) (correlation of 0.76). These findings show that our persistent dividend growth measure is significantly negatively correlated with risk proxies, e.g., stock market volatility and policy uncertainty, but positively correlated with economic growth and inflation.

Panel A in Table 5 uses levels and so the correlation estimates described above are driven by common, persistent factors reflecting the state of the economy. Panel B highlights short-run effects by reporting the correlations between daily changes in the underlying indexes. Changes in our daily dividend growth index are only significantly correlated with changes in daily inflation, suggesting that daily variation in our new dividend growth measure is not captured by daily variation in other financial and macroeconomic variables but constitutes a new source of (high frequency) information.

5 Implications for return predictability

A long-standing debate in asset pricing is concerned with the importance of cash flows expectations as a source of variation in stock returns. For example, Cochrane (2008) argues that dividend growth is largely unpredictable by means of variables such as the dividend-price ratio. If true, this has important asset pricing implications because it implies that movements in the dividend-price ratio must instead reflect time-varying risk premia.

Finding that the dividend-price ratio fails to predict future dividend growth does not, of course, rule out that other variables–such as our persistent dividend growth measure–can predict dividend growth and, in turn, stock returns. Whether such return predictability from our persistent dividend growth measure can be established–and whether it is consistent with basic principles of asset pricing–is what we set out to explore in this section.

5.1 A predictive present value model

We use a simple reduced-form present value model to explore the implications of dividend growth predictability for predictability of stock returns. This model provides a structured framework that allows us to quantify and test the implications of dividend growth

predictability for predictability of stock returns.[26]

Our analysis of return dynamics uses the following notations. We refer to p_{t+1} as the logarithm of the stock price on day $t+1$ while d^{cf}_{t+1} is the logarithm of the dividends *paid out* (distributed) on day $t+1$ and, as before, d_{t+1} is the logarithm of dividends *announced* on day $t+1$. The present value model refers to paid-out cash flows (cf) and the two dividend measures are very different at the daily horizon–recall that, on average, dividend announcements precede distributions by 42 days. At longer horizons such as a quarter or a year, differences between announced and distributed dividends tend to be much smaller.[27] Finally, $\Delta d^{cf}_{t+1} = d^{cf}_{t+1} - d^{cf}_{t}$ denotes the growth rate in distributed dividends, while r_{t+1} is the log return on day $t+1$.

Following Campbell and Shiller (1988a), consider the approximate log-linearized present value model

$$r_{t+1} \approx k + \rho(p_{t+1} - d^{cf}_{t+1}) + \Delta d^{cf}_{t+1} + (d^{cf}_{t} - p_t), \tag{15}$$

where $\rho = 1/(1+\exp(\overline{d-p})$ and $k = -\ln(\rho) - (1-\rho)\ln(1/\rho - 1)$ are log-linearization constants. At the daily frequency, the average dividend yield is very small and so $\rho = 0.99998$ is extremely close to unity. Using this value for ρ, the log-linearization in (15) is highly accurate with a correlation between exact and approximate returns of 0.9998.

Again, we emphasize the distinction between announced and distributed dividends for the daily present value model: at the daily horizon, the correlation between growth in announced and distributed dividends (Δd_{t+1} and Δd^{cf}_{t+1}) is essentially zero (-0.0072). Moreover, if we (incorrectly) use announced instead of distributed dividends in (15), the accuracy of the present value approximation deteriorates sharply.

Studying return predictability in the context of the present value model in (15) offers a number of important advantages. First, previous studies of return predictability such as Campbell and Shiller (1988a), Cochrane (2008) and van Binsbergen and Koijen (2010), have used this model.[28] Adopting a similar framework makes it easier to compare our findings to

[26] Other, less parametric approaches to asset pricing are also available, e.g., the pricing kernel bounds proposed by Hansen and Jagannathan (1997). An important advantage of the present value model is that it lends itself to empirical tests through a set of linear restrictions on the model parameters.

[27] The correlations between log-dividends calculated using announced and paid-out dividends are 33.23%, 72.45% and 83.61% at the monthly, quarterly and annual frequencies, respectively.

[28] Jagannathan and Liu (2019) embed a dividend growth VAR specification in an asset pricing model with recursive Epstein-Zin preferences and find significant predictability of annual stock returns from a stock yield variable, particularly in the presence of investor learning about dividend growth dynamics.

evidence from the existing literature.

Second, the present value model (15) has strong economic implications that directly link predictability of dividend growth to predictability of stock returns not only from popular predictors such as the log dividend-price ratio but also from other potential predictors such as news about our persistent dividend growth component. In particular, (15) implies that if a variable helps predict dividend growth, Δd_{t+1}^{cf}, this same variable should also predict returns–unless this variable has an equivalent and opposite effect on the price-dividend ratio, $p_{t+1} - d_{t+1}^{cf}$. This latter scenario arises if a variable forecasts equal movements in expected cash flow growth and discount rates so that the net effect on return predictability is zero.

Third, the present value model has implications for return predictability at different horizons. Because (15) holds regardless of the horizon, if a variable helps forecast dividend growth at multiple horizons, we would, in general, expect the same variable to possess predictive power over stock returns across these horizons. This is relevant here, given our empirical evidence that the persistent dividend growth component helps predict dividend growth at short as well as long horizons.

At the long forecast horizons conventionally used in the finance literature, it is common to predict the dividend yield, $d_{t+1}^{cf} - p_{t+1}$, because it is easy to extract a highly persistent component from the dividend-price ratio which is very smooth; see, e.g., Cochrane (2008). Conversely, at the daily horizon, $d_{t+1}^{cf} - p_{t+1}$ tends to be dominated by large variations in daily dividend distributions which, however, are temporary and so do not reflect any predictability of long-term dividend prospects.

To deal with this and obtain a more stable forecasting model, we instead predict daily *changes* in the dividend-price ratio $\Delta\left(d_{t+1}^{cf} - p_{t+1}\right) = (d_{t+1}^{cf} - p_{t+1}) - (d_t^{cf} - p_t)$. From an econometric perspective, the main effect of predicting $\Delta\left(d_{t+1}^{cf} - p_{t+1}\right)$ rather than $d_{t+1}^{cf} - p_{t+1}$ is to anchor daily variation in the dividend-price ratio on a random walk, imposing a coefficient of unity on $d_t^{cf} - p_t$. Imposing a similar constraint on the dividend yield has been found by Ferreira and Santa-Clara (2011) to reduce estimation error and lead to more accurate forecasts of stock returns. The constraint can be viewed as a sensible economic prior since we would expect the dividend-price ratio to contain a highly persistent (near unit root) component at the daily frequency.

Following this discussion, we approximate (15) as follows,

$$r_{t+1} \approx k - \Delta \left(d_{t+1}^{cf} - p_{t+1} \right) + \Delta d_{t+1}^{cf}. \tag{16}$$

The approximation in (16) is very accurate as a consequence of ρ being very close to unity at the daily horizon: the correlation between r_{t+1} and the approximation in (16) is 0.9997 in our sample.

Note that Δd_{t+1}^{cf} can be dropped from (16), so $r_{t+1} \approx k + (p_{t+1} - p_t)$, suggesting a predictive regression that simply projects capital gains, Δp_{t+1}, on any candidate predictors. However, daily capital gains and stock returns are very similar, so regressing capital gains rather than returns on the predictors does not provide additional economic insights. Conversely, stock prices can move because of changes in valuations (reflected in the dividend-price ratio) or because of cash flow news (dividend growth) and our decomposition in (16) allows us to identify predictability in these separate components. This offers the potential for new economic insights on return predictability at the daily horizon and for testing the validity of the cross-equation restrictions implied by the present value model (15).

Next, consider the list of predictors. Our first predictor is the lagged change in the persistent dividend growth component, $\Delta \mu_{dt}$. Using μ_{dt} as a predictor might seem the natural choice. However, μ_{dt} is constructed to capture the predictive component in *announced* dividends whereas the present value model (16) is based on *distributed* dividends which are uncorrelated with announced dividends at the daily frequency. Moreover, given its very high persistence, μ_{dt} is not well suited for predicting daily stock returns which are not very persistent. Conversely, $\Delta \mu_{dt}$ effectively measures innovations to μ_{dt} and so captures news about changes in the state of the economy.

Our second predictor is the lagged dividend-price ratio which has been extensively used in the finance literature to forecast variation in the investment opportunity set and so is important to include. To see why, note from (15) that variation in the log dividend-price ratio should reflect changes in expected future returns and/or changes in expected future dividend growth:

$$d_t^{cf} - p_t \propto E_t \left[\sum_{j=0}^{\infty} \rho^j \left(r_{t+j+1} - \Delta d_{t+j+1}^{cf} \right) \right]. \tag{17}$$

At the frequencies typically considered in empirical studies–monthly, quarterly, or annual–

the dividend-price ratio tends to be very smooth and highly persistent. This makes it well-suited as a predictor of slow movements in the investment opportunity set operating at longer horizons. In contrast, because of the lumpiness in daily dividends, the daily dividend-price ratio is very volatile. To obtain a predictor variable with persistence features more similar to the conventional dividend-price ratio obtained at lower frequencies, we measure the dividend-price ratio as a smoothed 30-day moving average of announced dividends divided by the current stock price, denoted $\overline{d_t - p_t}$.[29]

Finally, we also include the lagged value of the daily dividend growth, Δd_t^{cf}, in order to account for short-term reversals in the daily dividend growth process as days with very large dividend distributions tend to be followed by days with much smaller dividend payments.

Following this discussion, our daily prediction model for growth in distributed dividends takes the form

$$\Delta d_{t+1}^{cf} = \theta_0 + \theta_\mu \Delta \mu_{dt} + \theta_{dp} \left(\overline{d_t - p_t}\right) + \theta_d \Delta d_t^{cf} + \varepsilon_{dt+1}. \tag{18}$$

Similarly, we predict changes in the daily dividend-price ratio using the same set of predictors:

$$\Delta \left(d_{t+1}^{cf} - p_{t+1}\right) = \gamma_0 + \gamma_\mu \Delta \mu_{dt} + \gamma_{dp} \left(\overline{d_t - p_t}\right) + \gamma_d \Delta d_t^{cf} + \varepsilon_{t+1}^{cf}. \tag{19}$$

Using equations (18) and (19) in the log-linearized return equation (16), we have

$$\begin{aligned} r_{t+1} &\approx k + \theta_0 - \gamma_0 + (\theta_\mu - \gamma_\mu)\Delta \mu_{dt} + (\theta_{dp} - \gamma_{dp}) \left(\overline{d_t - p_t}\right) \\ &+ (\theta_d - \gamma_d)\Delta d_t^{cf} + \varepsilon_{dt+1} - \varepsilon_{t+1}^{cf}. \end{aligned} \tag{20}$$

This equation can be rewritten as

$$r_{t+1} = \lambda_0 + \lambda_\mu \Delta \mu_{dt} + \lambda_{dp}(\overline{d_t - p_t}) + \lambda_d \Delta d_t^{cf} + \varepsilon_{rt+1}, \tag{21}$$

where $\varepsilon_{rt+1} = \varepsilon_{dt+1} - \varepsilon_{t+1}^{cf}$. Comparing (21) to (20), it follows that the present value model

[29] Our empirical results are robust to using other windows to smooth the dividend process.

imposes the following cross-equation restrictions on the parameters of the return equation:

$$\begin{aligned} \lambda_0 &= k + \theta_0 - \gamma_0, \\ \lambda_\mu &= \theta_\mu - \gamma_\mu, \\ \lambda_{dp} &= \theta_{dp} - \gamma_{dp}, \\ \lambda_d &= \theta_d - \gamma_d. \end{aligned} \tag{22}$$

The restriction on λ_{dp} implies that if, at any horizon, $\overline{d_t - p_t}$ predicts future dividend growth ($\theta_{dp} \neq 0$ in (18)), it must, in general, also predict future returns ($\lambda_{dp} \neq 0$). Similarly, if, after controlling for the effect of the log dividend-price ratio, some other predictor such as $\Delta\mu_{dt}$ forecasts dividend growth ($\theta_\mu \neq 0$ in (18)), then the present value model requires that this variable also forecasts returns ($\lambda_\mu \neq 0$) unless $\theta_\mu = \gamma_\mu$, in which case the predictive effects from $\Delta\mu_{dt}$ on Δd^{cf}_{t+1} and $\Delta\left(d^{cf}_{t+1} - p_{t+1}\right)$ cancel out. This case arises if the variable predicts both higher (lower) dividend growth and higher (lower) discount rates in equal amounts, see (17).[30]

An important implication of (21) and (22) is that dividend news will, in general, also contain information about discount rates. Hence, care should be exercised when interpreting the effect of dividend news on variables such as stock returns and the dividend-price ratio since such news may also have sizeable discount rate effects. This interpretation is reinforced by our finding that the persistent dividend growth component helps predict consumption growth which, of course, is an important component of the pricing kernel in standard consumption-based asset pricing models.

5.2 Empirical findings

Before proceeding with the empirical analysis of the present value model, it is important to note that the behavior of the components in (15) is quite different at the daily frequency than at the longer horizons conventionally used in empirical studies.

First, as noted previously, the distinction between announced and distributed dividends is

[30]The model in Cochrane (2008) is a special case of our specification which is obtained by omitting $\Delta\mu_{dt}$ and Δd^{cf}_t as predictor variables (i.e., setting $\lambda_\mu = \lambda_d = 0$), assuming that the distributed and announced dividends are the same, and setting $\rho \approx 1$ which, as we have shown, holds to a very close approximation for daily data.

not important at longer horizons but is crucial at the daily horizon for which the correlation is only 52.12% and 0.72% for the two dividend series in log-levels and log-changes, respectively.

Second, the contribution to return volatility of the individual return components in (15) is very different at the daily and longer frequencies. For example, in quarterly data, Δd^{cf}_{t+1} is quite smooth and its volatility is roughly six times smaller than the volatility of the dividend-price ratio, $d^{cf}_{t+1} - p_{t+1}$. Conversely, in daily data, the volatility of Δd^{cf}_{t+1} is almost 50% *higher* than the volatility of $d^{cf}_{t+1} - p_{t+1}$ due to shifts in the composition of firms paying dividends on any given day. Hence, the volatility of dividend growth is nearly an order of magnitude greater at the daily than at the quarterly horizon, measured relative to the importance of variation in the dividend-price ratio.

Third, days with unusually large dividend payments (large, positive dividend growth) are typically followed by days with smaller dividend payments (negative dividend growth). This produces negative first-order autocorrelation in daily dividend growth–the first-order autocorrelation in Δd^{cf}_t is -0.35–which does not carry over beyond one day, however, and so is not helpful for predicting long-term dividend growth.

These observations are important for interpreting our empirical estimates and suggest that we should only expect to find limited predictability of dividend growth and in the log dividend-price ratio at the daily frequency by means of our three predictors.

The top row in Table 6 presents estimates from the predictive dividend growth regression, (18). We show results for the long sample, 1927-2016, but the estimates are very similar in the shorter sample, 1973-2016. Both the lagged dividend-price ratio and the lagged daily dividend growth obtain significantly negative coefficients. The negative coefficients on $\overline{d_t - p_t}$ and Δd^{cf}_t are largely a consequence of the short-term negative serial correlation (reversal) in daily dividend growth which also explains the large t-statistic on Δd^{cf}_t. Conversely, the coefficient on news about persistent growth in announced dividends ($\Delta \mu_{dt}$) fails to be significant. This is not surprising given the lead time of dividend announcements relative to dividend distributions.

Turning to the prediction equation for the change in the dividend-price ratio, (19), the second row in Table 6 shows that the coefficient on $\Delta \mu_{dt}$ is positive but insignificant. The coefficient on the dividend-price ratio is also insignificant, whereas lagged growth in cash flows, Δd^{cf}_t, produces a negative and significant coefficient, again as a result of short-term

reversals in daily dividend distributions.

Finally, in the prediction equation for daily stock returns, (21), only news about the persistent dividend growth component, $\Delta\mu_{dt}$, turns out to be significant. The estimated coefficient on $\Delta\mu_{dt}$ is 1.28 with a t-statistic of 3.31. This indicates that positive news about persistent dividend growth is associated with larger stock returns the following day.[31]

To see if these results are consistent with the asset pricing implications that follow from the present value model, the bottom row in Table 6 reports the outcome of Wald tests of the cross-equation restrictions in (22). We fail to reject any of the coefficient restrictions, suggesting that the relatively precisely estimated coefficient on $\Delta\mu_{dt}$ in the return equation (21) is consistent with the present value model. In fact, the estimated coefficient on $\Delta\mu_{dt}$ (1.28) is very close to the value implied by the present value model (1.21).

Similarly, the present value model implies small coefficients on both $\overline{d_t - p_t}$ and Δd_t^{cf} in the return equation which is again consistent with our estimates of (21).[32] Although these variables are significant predictors of variation in both daily dividend growth and changes in the dividend-price ratio, the two effects cancel out and so these variables do not have predictive power over returns. This case arises when the predictors correlate with future cash flow growth and discount rates in the same amount but in opposite directions.

Internet Appendix C assesses the economic value of return predictability from our μ_{dt} measure. We find that using information on μ_{dt} leads to an increase of 2.32% in the annualized certainty equivalent return for a mean-variance investor with moderate risk aversion.

5.3 Return predictability at longer horizons

A large empirical literature in finance examines return predictability at longer horizons such as a month or a quarter. To make our results comparable to this literature and to explore if our new dividend growth measure possesses predictive power over returns at longer horizons, we undertake two exercises.

First, we regress multi-day stock returns $r_{t+2::t+h} = \sum_{\tau=t+2}^{t+h} r_\tau$ on $\Delta\mu_{dt}$ in addition to a

[31] To be consistent with standard ways of testing the present value model, we use one-period-ahead returns, r_{t+1}, in these regressions. However, the results are robust if, instead, we skip one day and use r_{t+2}, as the results in Table 7 show.

[32] Similar results hold in the shorter sample, 1973-2016, where, again, we fail to reject the cross-equation restrictions in (22).

range of alternative measures of dividend growth or dividend dynamics such as the "raw" daily dividend growth, Δd_t, news about persistent dividend growth extracted from a model without a jump component, $\Delta\mu_{dt}^{NJ}$, jumps in the dividend growth process, $J_{dt}\xi_t$, the stochastic volatility component extracted from the dividend growth process, $h_{dt}/2$, and temporary shocks to dividend growth, ε_{dt+1}. These regressions use a conservative approach and skip one day to compute stock returns which therefore start on day $t+2$. We do so in order to show that our results are not due to price slippage effects.

Table 7 shows results for different models and forecast horizons ranging from one through three, five, and 21 trading days. We find that $\Delta\mu_{dt}$ has significant predictive power over stock returns at one, three, and five day horizons as well as one month (21 trading days) ahead in time. This finding strongly depends on allowing for jumps in the dividend growth process to extract $\Delta\mu_{dt}$; we find essentially no return predictability from the alternative measure that excludes jumps, $\Delta\mu_{dt}^{NJ}$. We also find little-to-no return predictability from raw daily dividend growth, Δd_t or from any of the other predictors.

Second, we consider the more common monthly horizon for which return predictability from variables such as the dividend-price ratio has been established in the literature. In particular, we compare return predictability from $\Delta\mu_{dt}$ to predictability from a set of 14 standard predictors used by Goyal and Welch (2008). Following common practice, we focus on univariate predictive regressions conducted one predictor at a time. We use the 20-year window 1973:01–1992:12 as an initial training sample and compute out-of-sample return forecasts for the remaining period 1993:01–2016:12, employing an expanding estimation window which adds new observations as they become available.

We consider out-of-sample R^2 values and average utility gains. The out-of-sample R^2 measures the percent reduction in mean squared forecast error obtained when conditioning on a predictor relative to using the prevailing mean forecast. Average utility gains can be interpreted as the portfolio management fees that an investor with mean variance preferences and a risk aversion coefficient of five is willing to pay to have access to the forecasts generated using a predictor, again computed relative to the prevailing mean.

The results, reported in Table 8, show that $\Delta\mu_{dt}$ has significant out-of-sample predictive power over stock returns, coming in as ranked second among the 15 predictor variables both in terms of the strength of its out-of-sample predictive power and in terms of utility gains.

Moreover, consistent with findings in the literature (e.g., Dangl and Halling (2012) and Rapach et al. (2010)), the degree of return predictability, along with the size of the utility gains, is larger during recessions than in expansions.

6 Dividend news and return dynamics

The present value model can also be used to quantify and assess the effect of dividend news on concurrent (same day) stock returns. In particular, the model has implications for which of the different dividend growth components is likely to most strongly affect stock returns. Shocks to the dividend growth process should be positively correlated with unexpected returns and shocks to the persistent (predictable) growth component should generally have a larger impact on stock returns than shocks to the temporary growth component. In this section, we set out to test these implications and provide a broader analysis of how the different dividend news components affect not only the mean of stock returns but also the volatility and probability of observing a jump in stock returns.[33]

6.1 Dividend news and contemporaneous stock returns

To analyze how stock returns are affected contemporaneously by shocks to the dividend growth process, we consider two alternative formulations of the prediction equations underlying the present value model. Although both formulations follow from the present value model–and thus are equivalent–considering both allows us to explore the robustness of our results in regards to how we estimate and test the model. For both cases, note that because μ_{dt+1} is highly persistent, $\Delta\mu_{dt+1}$ is essentially identical to the innovation to the μ_{dt+1} process ($\varepsilon_{\mu t+1}$): the two have a correlation of 0.9999. Hence, $\Delta\mu_{dt+1}$ can be interpreted as the news or shock to the predictable (persistent) component of the dividend growth process and so we include this variable in our model for contemporaneous movements in stock returns.[34]

Our first specification decomposes daily stock returns into a weighted average of the

[33] Internet Appendix C reports results from an analysis of the cross-sectional effects of shocks to the persistent component in the dividend growth process.

[34] Empirical results are essentialy identical if we use $\varepsilon_{\mu t+1}$ in our analysis instead of $\Delta\mu_{dt+1}$.

capital gain $(p_{t+1} - p_t)$ and dividend yield $(d^{cf}_{t+1} - p_t)$:

$$\begin{aligned} r_{t+1} &\approx k + \rho p_{t+1} - p_t + (1-\rho) d^{cf}_{t+1} \\ &= k + \rho(p_{t+1} - p_t) + (1-\rho)(d^{cf}_{t+1} - p_t). \end{aligned} \tag{23}$$

Next, we relate $p_{t+1} - p_t$ and $d^{cf}_{t+1} - p_t$ to news about the persistent dividend growth component, $\Delta\mu_{dt+1}$, controlling for movements in the smoothed log dividend-price ratio, $\overline{d_t - p_t}$, which captures investors' forward-looking expectations:

$$\begin{aligned} p_{t+1} - p_t &= \pi_0 + \pi_\mu \Delta\mu_{dt+1} + \pi_{dp}(\overline{d_t - p_t}) + \varepsilon_{pt+1}, \\ d^{cf}_{t+1} - p_t &= \delta_0 + \delta_{dp}(\overline{d_t - p_t}) + \varepsilon^{cf}_{t+1}. \end{aligned} \tag{24}$$

We leave out $\Delta\mu_{dt+1}$ from the equation for $d^{cf}_{t+1} - p_t$ because dividend announcements on day $t+1$ should not affect dividends paid out on the same day due to the lengthy lag between when dividends are announced and distributed.

The premise of our first specification, (23) and (24), is that daily dividend payments, d^{cf}_{t+1}, are largely known ahead of time. This is unique to the daily frequency and does not hold at longer horizons such as a quarter or a year. Because of daily composition shifts in the dividend-announcing firms, this does not imply that we would expect a high R^2 value for the regression of $d^{cf}_{t+1} - p_t$ on $\overline{d_t - p_t}$ in (24). However, it means that investors are able to quite accurately predict growth in next-day dividend distributions. As a result, forecasting r_{t+1} at the daily frequency is largely equivalent to forecasting the capital gain, $p_{t+1} - p_t$.

Combining the present value model in (23) with the regressions in (24), we get

$$r_{t+1} = \lambda_0 + \lambda_\mu \Delta\mu_{dt+1} + \lambda_{dp}(\overline{d_t - p_t}) + \varepsilon_{rt+1}, \tag{25}$$

where $\varepsilon_{rt+1} = \rho\varepsilon_{pt+1} + (1-\rho)\varepsilon^{cf}_{t+1}$. It follows from the present value model (23) that the coefficients in (25) must satisfy the cross-equation restrictions

$$\begin{aligned} \lambda_0 &= k + \rho\pi_0 + (1-\rho)\delta_0, \\ \lambda_\mu &= \rho\pi_\mu, \\ \lambda_{dp} &= \rho\pi_{dp} + (1-\rho)\delta_{dp}. \end{aligned} \tag{26}$$

These restrictions are analogous to the constraints obtained for the predictive present value model in (22) which only uses prior-day (lagged) information.

Our second specification uses the present value model (15) with the key difference that we now use the contemporaneous news shock to the persistent dividend growth component, $\Delta\mu_{dt+1}$, instead of its lagged value, $\Delta\mu_{dt}$. This model therefore takes the form:

$$\begin{aligned} \Delta d^{cf}_{t+1} &= \theta_0 + \theta_\mu \Delta\mu_{dt+1} + \theta_{dp}\left(\overline{d_t - p_t}\right) + \theta_d \Delta d^{cf}_t + \varepsilon_{dt+1}, \\ \Delta\left(d^{cf}_{t+1} - p_{t+1}\right) &= \gamma_0 + \gamma_\mu \Delta\mu_{dt+1} + \gamma_{dp}\left(\overline{d_t - p_t}\right) + \gamma_d \Delta d^{cf}_t + \varepsilon^{cf}_{t+1}, \\ r_{t+1} &= k - \Delta\left(d^{cf}_{t+1} - p_{t+1}\right) + \Delta d^{cf}_{t+1} + \varepsilon_{rt+1}. \end{aligned} \tag{27}$$

Again, the coefficients in (27) are subject to a set of cross-equation restrictions equivalent to those in (22).

6.2 Empirical findings

Panel A of Table 9 reports estimates of (24)-(25) fitted to returns data for the sample 1927-2016. First consider the capital gains equation in (24). At 1.55, the estimated coefficient on $\Delta\mu_{dt+1}$ is highly significant while, conversely, the coefficient on $\overline{d_t - p_t}$ is statistically insignificant. In turn, $\overline{d_t - p_t}$ is highly statistically significant in the prediction equation for $d^{cf}_{t+1} - p_t$ with a coefficient of 0.94, consistent with $\overline{d_t - p_t}$ picking up a persistent component in the daily dividend-price ratio.

Turning to the return regression (25), the coefficient on $\Delta\mu_{dt+1}$ remains highly statistically significant with an estimate of 1.57, suggesting that higher dividend growth news translates into positive stock returns with a strong and accurately estimated effect. The coefficient on $\overline{d_t - p_t}$ is essentially zero with a t-statistic below unity, confirming our earlier finding in Section 5 that movements in this predictor are not important for explaining variation in next-day stock returns, r_{t+1}. The R^2 value of 0.12% shows that news about persistent dividend growth is a non-trivial determinant of daily stock returns.

The bottom row in Panel A of Table 9 reports a set of Wald tests of the cross-equation restrictions in (26). With p-values ranging from 0.13 to 0.98, none of the restrictions gets rejected, suggesting that the effects of news about dividend growth are consistent with the asset pricing implications that follow from the present value model in (23). Panel B

reports estimates from the present value specification in (27). We find very similar coefficient estimates to those obtained from the predictive model analyzed in Section 5. In particular, the estimated coefficient on $\Delta\mu_{dt+1}$ in the return equation equals 1.45 with a highly significant t-statistic of 3.74. This estimate is larger than the estimate obtained from the predictive model (1.28) in Table 6, though a little lower than its value (1.57) obtained from equation (25).

A key motivation for constructing a daily dividend growth measure is that it can help identify the drivers of movements in daily stock returns. From a theoretical perspective, we would expect the three dividend growth components in (2) to have a very different impact on stock prices. For example, temporary shocks to dividend growth should have very little effect on stock prices while shocks to the persistent dividend growth component should have a larger impact. To see if this holds, we consider whether the other components from our dividend growth model (2) help explain movements in mean returns by expanding (25) to

$$r_{t+1} = \lambda_0 + \lambda_\mu \Delta\mu_{dt+1} + \lambda_{dp}(\overline{d_t - p_t}) + \lambda_\varepsilon \varepsilon_{dt+1} + \lambda_h \exp(h_{dt+1}) + \lambda_J J_{dt+1}\xi_{dt+1} + \varepsilon_{rt+1}. \quad (28)$$

In sharp contrast with the significant return effect of $\Delta\mu_{dt+1}$, in unreported results we fail to reject that $\lambda_\varepsilon = \lambda_h = \lambda_J = 0$ and also find that the estimated return effect of these additional terms is very small. Hence, temporary shocks to the dividend growth process as well as movements in dividend growth volatility or jumps in dividend growth fail to have a significant effect on mean returns.[35]

These results suggest that news about the persistent dividend growth rate affects the first moment (mean) of same-day stock returns while news about the temporary components does not. However, dividend news need not be confined to affecting mean returns and could also influence how volatile stock returns are on a given day: we might expect positive news about persistent dividend growth to be associated with calmer markets as it reduces the likelihood of being in a low-growth ("bad fundamentals") state. Movements in the volatility of dividend growth may also spill over to return volatility.

To explore these possibilities, we estimate the following auxiliary regression which uses

[35] Internet Appendix C undertakes an analysis that uses daily dividend distributions rather than dividend announcements. Whereas dividend announcements are significantly positively correlated with same-day stock returns, we find no such effect for dividend distributions.

the squared daily return residual from (25) as a proxy for return variance:

$$\varepsilon_{rt+1}^2 = a_0 + a_1\varepsilon_{rt}^2 + a_2\sigma_{dt+1}^2 + a_3\Delta\mu_{dt+1} + \varepsilon_{\sigma t+1}, \tag{29}$$

where $\sigma_{dt+1}^2 = \exp(h_{dt+1})$ is the stochastic variance series extracted from the daily dividend growth process, Δd_{t+1}. We include the lagged squared return residual, ε_{rt}^2, to account for persistence in the return variance process.

Panel C of Table 9 reports results from this regression. The significantly positive coefficient on prior-day squared return shocks, ε_{rt}^2, picks up persistence in the daily return variance. The variance of the dividend growth process, σ_{dt+1}^2, is strongly positively correlated with return variance, suggesting that higher uncertainty about dividend growth translates into higher same-day return volatility.[36] Consistent with good news about the persistent dividend growth component being associated with calmer markets, the coefficient on $\Delta\mu_{dt+1}$ is negative although, with a p-value of 0.12, it fails to be statistically significant.[37]

Daily dividend news may also affect the probability of observing jumps in stock returns. To explore this possibility, we go through our sample to flag days with unusually large price movements and, following Bandi and Renò (2016), introduce a jump indicator which equals one on days when the absolute value of the (standardized) residuals from the return equation exceed two. We then estimate the following jump test model[38]

$$I_{t+1} = b_0 + b_1\xi_{dt+1}J_{dt+1} + b_2\sigma_{dt+1}^2 + b_3\Delta\mu_{dt+1} + \varepsilon_{t+1}^I$$

Panel D in Table 9 shows that positive news about the persistent cash flow growth component, $\Delta\mu_{dt+1}$, is associated with a reduced probability of jumps in returns with a coefficient that is highly significant. Positive jumps in the dividend growth process tend to

[36] Using textual analysis, Boudoukh et al. (2019) find that, consistent with our results, public information about firm-level news is important for explaining the variance of stock returns.

[37] As we show in Internet Appendix C, μ_{dt} contains significant predictive power also for volatility measures based on high-frequency intra-daily price data and option prices (VIX).

[38] Estimates of a probit model linking this jump indicator to the three components extracted from our daily dividend growth model, i.e., jumps, $\xi_{dt+1}J_{dt+1}$, stochastic variance, σ_{dt+1}^2 and news about persistent dividend growth, $\Delta\mu_{dt+1}$:

$$\Pr\left(I_{t+1}|\,\xi_{dt+1}, J_{dt+1}, \sigma_{dt+1}^2, \Delta\mu_{dt+1}\right) = \Phi\left(b_0 + b_1\xi_{dt+1}J_{dt+1} + b_2\sigma_{dt+1}^2 + b_3\Delta\mu_{dt+1}\right). \tag{30}$$

are very similar.

be associated with a slightly reduced likelihood of observing a jump in stock returns on the same day, although the effect is only marginally significant at the 10% level. Conversely, higher uncertainty about cash flow growth (a larger value of σ^2_{dt+1}) is associated with a highly significant increase in the likelihood of a jump in stock returns on the same day. Greater uncertainty about cash flow growth thus translates into a higher chance of large movements in stock prices and our dividend growth components help to identify cash flow news as a source of jumps in daily stock returns.[39]

7 Conclusion

This paper develops a new methodology for constructing a daily "bottom-up" measure of aggregate cash flow growth based on firms' dividend announcements. In constructing this measure, we address two key challenges. First, individual firms' announced dividends can change by large amounts from one quarter to the next and display strong heterogeneity across firms. Second, the number and type of firms that announce dividends often changes substantially from day to day, leading to large composition shifts. Both effects cause lumpiness in the daily cash flow news process.

We handle this lumpiness by decomposing news on dividend growth into a transitory "normal" shock whose volatility can vary over time, jumps that occur more rarely but whose magnitude tends to be much larger, and a persistent, smoothly evolving component that captures long-run predictive dynamics in the mean of the cash flow growth process. We find that these components are well identified in the daily dividend growth data. Importantly, the persistent mean component captures predictable dynamics in dividend growth which is easily overlooked in the raw dividend growth series which gets dominated by the highly volatile temporary jump component. We show empirically that this persistent dividend growth component can be used to produce more accurate forecasts of future dividend growth than alternative approaches from the finance literature.

While our empirical analysis uses high-frequency (daily) dividend announcements, it also

[39] Our findings are related to an extensive literature in finance that finds evidence of time-varying jump risk in stock returns using data on high-frequency intra-day returns or out-of-the-money options, see, e.g., Bollerslev and Todorov (2011). Similarly, Kelly and Jiang (2014) introduce a new measure of time-varying tail risk extracted from the cross-section of stock returns which they find has strong predictive power over returns on individual stocks and on market returns.

offers new insights into the drivers of stock price dynamics at longer horizons. Indeed, shocks to the persistent dividend growth component identified by our model are long-lasting–with a half life exceeding a year–and lead measures of economic activity traditionally used as proxies for "fundamentals" such as GDP and consumption growth. Moreover, our estimates of persistent dividend growth can be used to produce more accurate forecasts of dividend growth than existing alternatives at both monthly and annual horizons.

Finally, we use a present value model to explore implications for asset pricing and return predictability associated with our results on dividend growth predictability. We find strong evidence that the persistent dividend growth component can be used to forecast stock market returns both at short horizons such as a single day and at longer horizons such as a month. Moreover, the cross-equation restrictions implied by the present value model are supported by the data. Positive news about the persistent dividend growth component are associated with higher mean stock returns, reduced stock market volatility, and a reduced likelihood of a jump in stock returns on the same day. Finally, we find that greater uncertainty about cash flow growth translates into higher volatility for stock prices and we identify news about our dividend growth components as drivers of jumps in daily stock returns.

References

Albert, J. H. and S. Chib (1993): "Bayesian Analysis of Binary and Polychotomous Response Data," *Journal of the American Statistical Association*, 88, 669–679.

Aruoba, S. B., F. X. Diebold, and C. Scotti (2009): "Real-Time Measurement of Business Conditions," *Journal of Business and Economic Statistics*, 27(4), 417–427.

Baker, S. R., N. Bloom, and S. J. Davis (2016): "Measuring Economic Policy Uncertainty," *The Quarterly Journal of Economics*, 131(4), 1593–1636.

Bandi, F. and R. Renò (2016): "Price and volatility co-jumps," *Journal of Financial Economics*, 119, 107 – 146.

Bansal, R. and A. Yaron (2004): "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *The Journal of Finance*, 59(4), 1481–1509.

Barillas, F. and J. Shanken (2018): "Comparing Asset Pricing Models," *The Journal of Finance*, 73(2), 715–754.

Bollerslev, T. and V. Todorov (2011): "Tails, Fears, and Risk Premia," *The Journal of Finance*, 66, 2165–2211.

Bollerslev, T., V. Todorov, and L. Xu (2015): "Tail risk premia and return predictability," *Journal of Financial Economics*, 118, 113 – 134.

Boudoukh, J., R. Feldman, S. Kogan, and M. Richardson (2019): "Information, Trading, and Volatility: Evidence from Firm-Specific News," *The Review of Financial Studies*, 32(3), 992–1033.

Boyd, J. H., J. Hu, and R. Jagannathan (2005): "The stock market's reaction to unemployment news: Why bad news is usually good for stocks," *The Journal of Finance*, 60(2), 649–672.

Brav, A., J. R. Graham, C. R. Harvey, and R. Michaely (2005): "Payout policy in the 21st century," *Journal of Financial Economics*, 77, 483–527.

Campbell, J. and R. Shiller (1988a): "The dividend-price ratio and expectations of future dividends and discount factors," *Review of Financial Studies*, 1, 195–228.

Campbell, J. Y., S. Giglio, and C. Polk (2013): "Hard Times," *The Review of Asset Pricing Studies*, 3(1), 95–132.

Campbell, J. Y. and R. J. Shiller (1988b): "Stock prices, earnings, and expected dividends," *The Journal of Finance*, 43(3), 661–676.

Campbell, J. Y. and L. M. Viceira (2002): *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*.

Cavallo, A. and R. Rigobon (2016): "The Billion Prices Project: Using Online Data for Measurement and Research," *Journal of Economic Perspectives*, 31(1), 151–178.

Chib, S., F. Nardari, and N. Shephard (2002): "Markov chain Monte Carlo methods for stochastic volatility models," *Journal of Econometrics*, 108(2), 281–316.

Cochrane, J. H. (1992): "Explaining the Variance of Price-Dividend Ratios," *The Review of Financial Studies*, 5(2), 243–280.

——— (2008): "The dog that did not bark: a defense of return predictability," *The Review of Financial Studies*, 21(4), 1533–1575.

Cohen, R. B., C. Polk, and T. Vuolteenaho (2009): "The Price Is (Almost) Right," *The Journal of Finance*, 64(6), 2739–2782.

Corsi, F. (2009): "A Simple Approximate Long-Memory Model of Realized Volatility," *Journal of Financial Econometrics*, 7(2), 174–196.

Cutler, D. M., J. M. Poterba, and L. H. Summers (1989): "What moves stock prices?" *The Journal of Portfolio Management*, 15(3), 4–12.

Dangl, T. and M. Halling (2012): "Predictive regressions with time-varying coefficients," *Journal of Financial Economics*, 106(1), 157–181.

Donaldson, R. G. and M. Kamstra (1996): "A New Dividend Forecasting Procedure that Rejects Bubbles in Asset Prices: The Case of 1929's Stock Crash," *The Review of Financial Studies*, 9(2), 333–383.

Eraker, B., M. Johannes, and N. Polson (2003a): "The Impact of Jumps in Volatility and Returns," *The Journal of Finance*, 58(3), 1269–1300.

——— (2003b): "The Impact of Jumps in Volatility and Returns," *The Journal of Finance*, 58(3), 1269–1300.

Ferreira, M. A. and P. Santa-Clara (2011): "Forecasting stock market returns: The sum of the parts is more than the whole," *Journal of Financial Economics*, 100, 514–537.

Gilchrist, S. and E. Zakrajsek (2012): "Credit Spreads and Business Cycle Fluctuations," *American Economic Review*, 102(4), 1692–1720.

Goyal, A. and I. Welch (2008): "A comprehensive look at the empirical performance of equity premium prediction," *The Review of Financial Studies*, 21(4), 1455–1508.

Hansen, L. P. and R. Jagannathan (1997): "Assessing Specification Errors in Stochastic Discount Factor Models," *The Journal of Finance*, 52(2), 557–590.

Hu, G. X., J. Pan, and J. Wang (2013): "Noise as Information for Illiquidity," *The Journal of Finance*, 68(6), 2341–2382.

Jagannathan, R. and B. Liu (2019): "Dividend Dynamics, Learning, and Expected Stock Index Returns," *The Journal of Finance*, 74(1), 401–448.

Johannes, M., R. Kumar, and N. G. Polson (1999): "State Dependent Jump Models: How do US Equity Indices Jump?" *Working Paper*.

Jurado, K., S. C. Ludvigson, and S. Ng (2015): "Measuring Uncertainty," *American Economic Review*, 105(3), 1177–1216.

Kelly, B. and H. Jiang (2014): "Tail Risk and Asset Prices," *The Review of Financial Studies*, 27, 2841–2871.

Kelly, B. and S. Pruitt (2013): "Market Expectations in the Cross-Section of Present Values," *The Journal of Finance*, 68(5), 1721–1756.

Kim, S., N. Shephard, and S. Chib (1998): "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *The Review of Economic Studies*, 65(3), 361–393.

Koijen, R. S. and S. V. Nieuwerburgh (2011): "Predictability of Returns and Cash Flows," *Annual Review of Financial Economics*, 3, 467–491.

Lettau, M. and S. Ludvigson (2005): "Expected returns and expected dividend growth," *Journal of Financial Economics*, 76(3), 583–626.

Lettau, M. and S. V. Nieuwerburgh (2008): "Reconciling the return predictability evidence," *The Review of Financial Studies*, 21(4), 1607–1652.

Lewellen, J., S. Nagel, and J. Shanken (2010): "A skeptical appraisal of asset pricing tests," *Journal of Financial Economics*, 96, 175–194.

Liew, J. and M. Vassalou (2000): "Can book-to-market, size and momentum be risk factors that predict economic growth?" *Journal of Financial Economics*, 57(2), 221–245.

Maio, P. and P. Santa-Clara (2015): "Dividend Yields, Dividend Growth, and Return Predictability in the Cross-Section of Stocks," *Journal of Financial and Quantitative Analysis*, 50(1-2), 33–60.

McQueen, G. and V. V. Roley (1993): "Stock Prices, News, and Business Conditions," *Review of Financial Studies*, 6(3), 683–707.

Patton, A. J. and M. Verardo (2012): "Does Beta Move with News? Firm-Specific Information Flows and Learning about Profitability," *The Review of Financial Studies*, 25(9), 2789–2839.

Paye, B. S. (2012): ""Déjà vol:" Predictive regressions for aggregate stock market volatility using macroeconomic variables," *Journal of Financial Economics*, 106(3), 527–546.

Rapach, D. E., J. K. Strauss, and G. Zhou (2010): "Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy," *The Review of Financial Studies*, 23(2), 821–862.

Sabbatucci, R. (2017): "Are Dividends and Stock Returns Predictable? New Evidence Using M&A Cash Flows," *Working Paper*.

Savor, P. and M. Wilson (2016): "Earnings Announcements and Systematic Risk," *The Journal of Finance*, 71(1), 83–138.

van Binsbergen, J. H., M. Brandt, and R. S. J. Koijen (2012): "On the Timing and Pricing of Dividends," *American Economic Review*, 102(4), 1596–1618.

van Binsbergen, J. H. and R. S. J. Koijen (2010): "Predictive Regressions: A Present-Value Approach," *The Journal of Finance*, 65(4), 1439–1471.

——— (2017): "The Term Structure of Returns: Facts and Theory," *Journal of Financial Economics*, 124(1), 1–21.

Vuolteenaho, T. (2002): "What Drives Firm-Level Stock Returns?" *The Journal of Finance*, 57(1), 233–264.

Vuong, Q. H. (1989): "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses," *Econometrica*, 57(2), 307–333.

Parameter estimates						
	1927-2016			**1973-2016**		
	Mean	*Std*	*90% Credible Set*	*Mean*	*Std*	*90% Credible Set*
μ_d	0.058	0.016	[0.035,0.080]	0.080	0.012	[0.062,0.097]
ϕ_μ	0.999	0.000	[0.999,0.999]	0.998	0.001	[0.997,0.999]
σ_μ	0.002	0.000	[0.002,0.002]	0.002	0.000	[0.002,0.002]
μ_h	-5.076	0.037	[-5.138,-5.014]	-5.337	0.045	[-5.415,-5.262]
ϕ_h	0.749	0.008	[0.735,0.762]	0.833	0.008	[0.820,0.846]
σ_h	1.320	0.031	[1.269,1.369]	0.752	0.028	[0.706,0.797]
σ_ξ	2.651	0.040	[2.584,2.718]	2.761	0.040	[2.695,2.829]
λ_1	-1.387	0.039	[-1.452,-1.323]	-1.354	0.045	[-1.428,-1.281]
λ_2	-0.054	0.006	[-0.064,-0.045]	-0.024	0.002	[-0.028,-0.021]

Table 1: Parameter estimates for the dividend growth rate model. This table shows parameter estimates for a model fitted to the daily dividend growth series. The equations for the components model, further described in Section 3.1, take the following form:

$$\Delta d_{t+1} = \mu_{dt+1} + \xi_{dt+1} J_{dt+1} + \varepsilon_{dt+1},$$

$$\mu_{dt+1} = \mu_d + \phi_\mu \left(\mu_{dt} - \mu_d\right) + \sigma_\mu \varepsilon_{\mu t+1},$$

$$\varepsilon_{dt+1} \sim \mathcal{N}(0, e^{h_{dt+1}}),$$

$$h_{dt+1} = \mu_h + \phi_h \left(h_{dt} - \mu_h\right) + \sigma_h \varepsilon_{ht+1},$$

$$\Pr\left(J_{dt+1} = 1\right) = \Phi\left(\lambda_1 + \lambda_2 N_{dt+1}\right),$$

$$\xi_{dt+1} \sim \mathcal{N}\left(0, \sigma_\xi^2\right),$$

where μ_{dt+1} captures the mean of the smooth component of the underlying dividend process, $J_{dt+1} \in \{0,1\}$ is a jump indicator that equals unity in case of a jump in dividends and otherwise is zero, ξ_{dt+1} measures the jump size, ε_{dt+1} is a temporary cash flow shock, $\varepsilon_{\mu t+1} \sim \mathcal{N}(0,1)$ is assumed to be uncorrelated at all times with the innovation in the temporary dividend growth component, ε_{dt+1}, and $|\phi_\mu| < 1$. h_{dt+1} denotes the log-variance of ε_{dt+1}, and $\varepsilon_{ht+1} \sim \mathcal{N}(0,1)$ is uncorrelated at all times with both ε_{dt+1} and $\varepsilon_{\mu t+1}$. N_{dt+1} denotes the number of firms announcing dividends on day $t+1$, while Φ stands for the CDF of a standard Normal distribution and $\xi_{dt+1} \sim \mathcal{N}\left(0, \sigma_\xi^2\right)$ captures the magnitude of the jumps. The columns report the posterior mean, standard deviation and 90% credible sets for the parameter estimates.

PANEL A: $\Delta d_{t+1}^{CRSP} = \alpha + \rho_i \sum_{i=1}^{3} \Delta d_{t+1-i}^{CRSP} + \beta \mu_{dt} + \gamma dp_t^{CRSP} + \varepsilon_{t+1}$

	Quarterly (1973-)	**Annual (1973-)**	**Quarterly (1927)**	**Annual (1927)**
μ_{dt}	.36***	2.50***	.14**	.96***
	[4.58]	[3.95]	[2.49]	[3.20]
dp_t	-.01	.03	-.00	.04
	[-1.29]	[0.50]	[-0.73]	[1.11]
Δd_t^{CRSP}	.16**	-.64***	.34***	-.16
	[2.20]	[-4.58]	[5.58]	[-1.07]
Δd_{t-1}^{CRSP}	.07	-.53***	.22***	-.12
	[1.39]	[-3.87]	[4.58]	[-0.84]
Δd_{t-2}^{CRSP}	.03	-.12	.01	.01
	[0.45]	[-0.78]	[0.26]	[0.05]
R^2	26.50%	27.45%	37.42%	6.31%
Observations	169	40	353	86

PANEL B: $\Delta d_{t+1}^{i} = \alpha + \beta \mu_{dt} + \gamma g_t^i + \varepsilon_{t+1}$

	Annual			**Monthly**		
μ_{dt}	1.28***		1.10***	1.39***		1.16***
	[4.04]		[3.97]	[7.23]		[5.12]
g_t^{VBK}		.90**	.34			
		[2.16]	[0.91]			
g_t^{KP}					1.45***	.43
					[6.41]	[1.58]
R^2	25.84%	13.14%	27.27%	31.65%	20.30%	32.67%
Observations	42	42	42	527	527	527

PANEL C: $\Delta d_{t+1}^{i} = \beta_1 \mu_{dt} + (1-\beta_1)\gamma g_t^i + \varepsilon_{t+1}$

β_1	0.87***	1.08***
p-value	(0.00)	(0.00)
$1-\beta_1$	0.13	-0.08
p-value	(0.66)	(0.53)
β_1 (OOS)		1.25***
p-value (OOS)		(0.00)
$1-\beta_1$ (OOS)		-0.25***
p-value (OOS)		(0.00)

Table 2: Dividend growth regressions. Panel A reports results from predictive regressions of dividend growth extracted from CRSP data, Δd_{t+1}^{CRSP}, on the persistent component μ_{dt} estimated from our daily dividend growth model and the log dividend price ratio, dp_t, at quarterly and annual frequencies. Panel B compares the predictive power of our persistent dividend growth component to that of two alternative dividend growth variables. The first measure, g_t^{VBK}, is taken from van Binsbergen and Koijen (2010) and uses cash reinvested dividend growth, measured annually. The second measure, g_t^{KP}, is taken from Kelly and Pruitt (2013) and uses monthly data. Panel C reports results from forecast encompassing regressions which compare the predictive power of our μ_{dt} measure to the two alternative measures, both in-sample and out-of-sample (available only for Kelly and Pruitt (2013)). Square brackets report t-statistics computed using Newey-West standard errors with three lags. Unless otherwise indicated, the sample period is 1973-2016.

$$\Delta d_{t+1}^{CRSP} = \alpha + \beta \mu_{dt}^{i} + \gamma dp_t^{CRSP} + \sum_{i=1}^{3} \rho_i \Delta d_{t+1-i}^{CRSP} + \varepsilon_{t+1}$$

	Monthly			**Quarterly**			**Annual**		
μ_{daily}	.17***		1.14***	.36***		.73***	2.50***		1.72
	[5.77]		[15.01]	[4.58]		[7.64]	[3.95]		[1.44]
$\mu_{monthly}$		.14***	-.14*						
		[6.51]	[-1.90]						
$\mu_{quarterly}$					.22***	.27***			
					[3.19]	[2.87]			
μ_{yearly}								2.64**	-.72
								[2.58]	[-0.60]
dp_t	-.00**	-.00	-.03***	-.01	.00	-.02***	.03	.07	.05
	[-2.49]	[-0.97]	[-9.86]	[-1.29]	[0.20]	[-2.79]	[0.50]	[1.12]	[0.73]
Δd_t^{CRSP}	-.05	-.05	-.58***	.16**	.22**	-.05	-.64***	-.52***	-.52***
	[-0.97]	[-0.94]	[-4.77]	[2.20]	[2.59]	[-0.72]	[-4.58]	[-3.07]	[-2.81]
Δd_{t-1}^{CRSP}	-.07*	-.07**	-.58***	.07	.10	-.10	-.53***	-.36***	-.39*
	[-1.96]	[-2.01]	[-5.69]	[1.39]	[1.65]	[-1.50]	[-3.87]	[-3.01]	[-2.03]
Δd_{t-2}^{CRSP}	.20***	.20***	-.31***	.03	.06	-.13*	-.12	-.08	-.05
	[3.68]	[3.59]	[-3.99]	[0.45]	[1.05]	[-1.84]	[-0.78]	[-0.52]	[-0.26]
R^2	18.29%	17.80%		26.50%	21.76%		27.45%	22.14%	
Observations	525	525	525	169	169	169	40	40	40

Table 3: Dividend growth regressions. This table shows results from predictive regressions of the conventional dividend growth measure extracted from CRSP data, Δd_{t+1}^{CRSP} on the persistent component, μ_{dt}, extracted from daily, quarterly, or annual dividend growth series, and the log dividend-price ratio, dp_t. The first column of each block (monthly, quarterly, annual) regresses dividend growth on a daily estimate of μ_{dt}, measured at the end of the corresponding period. The third column of each block reports forecast encompassing regressions which impose that the coefficients on the daily estimate of μ_{dt} and the monthly, quarterly, or annual estimates of μ_{dt} sum to unity. Square brackets report t-statistics computed using Newey-West standard errors with three lags. The sample period for these regressions is 1973-2016.

Panel A $\Delta y_{t+1} = \alpha + \beta \mu_{daily} + \gamma \Delta y_t + \varepsilon_{t+1}$				
	ΔGDP		$\Delta Consumption$	
μ_{daily}	.15***	.09***	.15***	.04***
	[5.68]	[3.30]	[6.58]	[2.84]
Δy_t		.37***		.67***
		[3.87]		[12.16]
R^2	23.52%	33.56%	31.08%	59.97%
Observations	171	171	171	171

Panel B $\Delta y_{t+1} = \alpha + \beta \mu_{monthly} + \gamma \Delta y_t + \varepsilon_{t+1}$				
	ΔGDP		$\Delta Consumption$	
$\mu_{monthly}$	.13***	.07***	.13***	.04*
	[4.28]	[2.53]	[4.80]	[1.90]
Δy_t		.40***		.68***
		[4.59]		[11.15]
R^2	20.11%	32.37%	27.36%	59.85%
Observations	171	171	171	171

Panel C: $\Delta y_{t+1} = \alpha + \beta \mu_{quarterly} + \gamma \Delta y_t + \varepsilon_{t+1}$				
	ΔGDP		$\Delta Consumption$	
$\mu_{quarterly}$	.08**	.04	.09***	.02
	[2.06]	[1.44]	[2.65]	[1.63]
Δy_t		.49***		.74***
		[5.77]		[17.22]
R^2	5.85%	28.68%	9.36%	58.64%
Observations	171	171	171	171

Table 4: Predictive regressions of GDP and consumption growth on the persistent dividend growth component. This table reports estimates from quarterly predictive regressions of future GDP and consumption growth, Δy_{t+1}, on the persistent dividend growth component, μ_{dt}, estimated from our daily dividend growth model (Panel A) or from dividend data sampled at the monthly (Panel B) or quarterly (Panel C) frequencies. Square brackets show t-statistics computed using Newey-West standard errors with three lags. The sample period used for these regressions is 1973-2015.

PANEL A: Persistent dividend growth and indexes (levels)

	VIX	*PU*	*ADS*	*Liquidity*	*Inflation*	μ_{dt}
VIX	1					
PU	0.34***	1				
	(0.00)					
ADS	-0.49***	-0.26***	1			
	(0.00)	(0.00)				
Liquidity	0.67***	0.19***	-0.54***	1		
	(0.00)	(0.00)	(0.00)			
Inflation	-0.66***	-0.36***	0.50***	-0.63***	1	
	(0.00)	(0.00)	(0.00)	(0.00)		
μ_{dt}	-0.54***	-0.20***	0.20***	-0.51***	0.76***	1
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	

PANEL B: Persistent dividend growth and indexes (changes)

	ΔVIX	ΔPU	$\Delta ADS\ Index$	$\Delta Liquidity$	$\Delta Inflation$	$\Delta\mu_{dt}$
ΔVIX	1					
ΔPU	-0.02	1				
	(0.10)					
$\Delta ADS\ Index$	0.01	0.00	1			
	(0.65)	(0.94)				
$\Delta Liquidity$	-0.01	0.05***	-0.00	1		
	(0.53)	(0.00)	(0.95)			
$\Delta Inflation$	0.03	-0.03	-0.00	-0.00	1	
	(0.24)	(0.27)	(0.89)	(0.94)		
$\Delta\mu_{dt}$	-0.02	-0.00	0.04	-0.01	0.10***	1
	(0.16)	(0.77)	(0.66)	(0.58)	(0.00)	

Table 5: Correlations between the persistent dividend growth component μ_{dt} and macroeconomic and financial activity measures. This table reports correlations between the persistent dividend growth component μ_{dt} extracted from our daily cash flow model and the following daily macroeconomic variables/indicators: the VIX index, the policy uncertainty index of Baker et al. (2016), the ADS index of Aruoba et al. (2009), the liquidity noise index of Hu et al. (2013), and the daily inflation index of Cavallo and Rigobon (2016). Panel A correlates the levels of these variables, while Panel B correlates changes in the variables. Parentheses below the correlation estimates show p-values. Three asterisks indicate statistical significance at the 1% level, two starts indicate significance at the 5% level, while one star indicates significance at the 10% level. Sample periods vary according to the length of the individual series.

	Const.	$\Delta\mu_{dt}$	$\overline{d_t - p_t}$	Δd_t^{cf}	R^2
Δd_{t+1}^{cf}	-.24**	13.93	-.04**	-.36***	12.87%
	(-2.31)	(0.30)	(-2.54)	-58.32	
$\Delta\left(d_{t+1}^{cf} - p_{t+1}\right)$	-.23**	12.72	-.04	-0.36***	12.87%
	(-2.31)	(0.28)	(-2.53)	(-58.32)	
ret_{t+1}	.00	1.28***	.00	-.00	0.08%
	(0.86)	(3.31)	(0.27)	(-0.14)	
$p - value$ of restriction	0.37	0.87	0.13	0.75	

Table 6: Parameter estimates and test statistics for the predictive present value model fitted to daily data. This table reports parameter values and test statistics for a predictive present value model estimated at the daily frequency over the sample period 1927-2016:

$$\Delta d_{t+1}^{cf} = \theta_0 + \theta_\mu \Delta\mu_{dt} + \theta_{dp}(\overline{d_t - p_t}) + \theta_d \Delta d_t^{cf} + \varepsilon_{dt+1}.$$

$$\Delta\left(d_{t+1}^{cf} - p_{t+1}\right) = \gamma_0 + \gamma_\mu \Delta\mu_{dt} + \gamma_{dp}(\overline{d_t - p_t}) + \gamma_d \Delta d_t^{cf} + \varepsilon_{t+1}^{cf}$$

$$r_{t+1} = \lambda_0 + \lambda_\mu \Delta\mu_{dt} + \lambda_{dp}(\overline{d_t - p_t}) + \lambda_d \Delta d_t^{cf} + \varepsilon_{rt+1}$$

with the following restrictions implied by the model

$$\lambda_0 = k + \theta_0 - \gamma_0,\ \lambda_\mu = \theta_\mu - \gamma_\mu,\ \lambda_{dp} = \theta_{dp} - \gamma_{dp},\ \lambda_d = \theta_d - \gamma_d$$

	Return horizon											
	r_{t+2}			$r_{t+2:t+4}$			$r_{t+2:t+6}$			$r_{t+2:t+22}$		
Δd_t	-.00**			-.00			-.00			-.00		
	[-2.19]			[-0.07]			[-0.17]			[-0.32]		
$\Delta\mu_{dt}^{NJ}$		-.00**			-.00			-.00			-.00	
		[-2.20]			[-0.00]			[-0.08]			[-0.26]	
$\Delta\mu_{dt}$			1.14***			3.17***			5.40***			21.60***
			[3.36]			[3.90]			[4.77]			[4.98]
$\xi_{dt}J_{dt}$			-.00			.00			.00			.00
			[-0.04]			[1.49]			[0.53]			[0.37]
$h_{dt}/2$			.00			.00			.00			.01
			[0.13]			[0.85]			[1.52]			[0.72]
ε_{dt+1}			-.00			.00			-.00			-.00
			[-1.16]			[0.37]			[-0.30]			[-0.77]
R^2	0.03%	0.03%	0.09%	0.00%	0.00%	0.22%	0.00%	0.00%	0.37%	0.00%	0.00%	1.27%
Observations	20,965	20,965	20,965	20,963	20,963	20,963	20,961	20,961	20,961	20,945	20,945	20,945

Table 7: Predictive regressions for stock returns. This table shows estimates of predictive regressions using cumulative stock market returns as the dependent variable and the following predictor variables 1) daily growth in announced dividends, (Δd_t); 2) changes in the persistent dividend growth component $\Delta\mu_t^{NJ}$ estimated from a model without jumps and stochastic volatility; 3) changes to the persistent dividend growth component $\Delta\mu_t$ estimated from a model that accounts for jumps and stochastic volatility. The dependent variables are the one-day ahead return r_{t+2} (columns 1-3), the 3-day-ahead cumulative return $r_{t+2:t+4}$ (columns 4-6), the 5-day-ahead cumulative return $r_{t+2:t+6}$ (columns 7-9), and the 1-month (21 trading day) ahead cumulative return $r_{t+2:t+22}$ (columns 10-12). We skip one day in the predictive regressions to account for dividends announced after trading hours. t-statistics, shown in square brackets, are computed using Newey-West standard errors with two lags for 1, 3, and 5 day-ahead cumulative returns and 10 lags for 1-month-ahead cumulative returns. The sample period is 1927-2016.

	Full Sample			Expansions			Recessions		
	$R^2_{OS}(\%)$	$p-val$	$\Delta U(\%)$	$R^2_{OS}(\%)$	$p-val$	$\Delta U(\%)$	$R^2_{OS}(\%)$	$p-val$	$\Delta U(\%)$
$\Delta\mu_{dt}$	0.43	0.14	1.28	-0.02	0.35	0.11	1.68	0.03	11.40
log(DP)	-1.98	0.97	-2.44	-3.06	0.99	-3.63	1.01	0.09	7.66
log(DY)	-2.01	0.96	-2.28	-3.25	0.99	-3.73	1.41	0.04	10.18
log(EP)	-0.85	0.67	0.74	-0.94	0.86	-1.17	-0.61	0.48	17.45
log(DE)	-2.04	0.80	-0.77	-1.29	0.99	-1.24	-4.13	0.65	3.19
SVAR	1.49	0.12	1.30	0.08	0.34	0.10	5.38	0.13	11.80
BM	-0.45	0.80	-0.77	-0.51	0.75	-0.96	-0.28	0.82	1.00
NTIS	-2.57	0.97	-1.37	-1.31	0.72	-0.23	-6.04	0.99	-11.33
TBL	-0.55	0.43	-0.57	0.42	0.12	0.93	-3.22	0.98	-13.20
LTY	-0.33	0.65	-0.14	-0.17	0.50	-0.17	-0.76	0.82	0.37
LTR	-0.20	0.31	-0.31	-0.28	0.33	-0.38	0.01	0.41	-0.04
TMS	-1.20	0.65	-1.33	-1.11	0.51	-0.22	-1.46	0.80	-11.35
DFY	-3.24	0.99	-4.13	-3.02	0.96	-2.66	-3.84	0.96	-17.29
DFR	-2.13	0.41	1.08	-2.17	0.49	-0.25	-2.00	0.41	12.41
INFL	-0.75	0.79	-1.06	-0.33	0.53	-0.30	-1.89	0.89	-7.62

Table 8: Out-of-sample return predictability. This table compares the out-of-sample predictive performance of our $\Delta\mu_{dt}$ measure to that of 14 standard predictors, used by Goyal and Welch (2008) to forecast stock returns, for the period 1993:01–2016:12. The predictor variables are the log dividend price ratio (log(DP)), the log dividend yield (log(DY)), the log earnings price ratio (log(EP)), the log dividend payout ratio (log(DE)), the stock variance (SVAR), the book-to-market ratio (BM), the net equity expansion (NTIS), the Treasury-bill rate (TBL), the long term yield (LTY), the long term rate of return (LTR), the term spread (TMS), the default yield spread (DFY), the default return spread (DFR), and the inflation rate (INFL). For all predictor variable, we use a 20-year initial estimation window from 1973:01 to 1992:12 followed by an expanding estimation window that adds new data points to the sample as they become available. We report both the out-of-sample R^2 value and average utility gains for the full out-of-sample period as well as separately for recession and expansion months, as defined by the NBER. For any given predictor, the out-of-sample R^2 measures the percentage reduction in the mean squared forecast error associated with a model that uses this predictor relative to the mean squared forecast error of forecasts based on the historical mean return. Positive values indicate better performance than this benchmark, while negative values suggest worse performance. The average utility gain can be interpreted as the portfolio management fee that an investor with mean variance preferences and a risk aversion coefficient of five would be willing to pay to have access to the forecasts generated by a particular predictor, again measured relative to the historical average forecasts.

Panel A: Estimates and tests for present value model (I)

	Const.	$\Delta\mu_{d,t+1}$	$\overline{d_t - p_t}$	R^2
$p_{t+1} - p_t$	-.00	1.55***	-.00	0.13%
	(-1.26)	(3.94)	(-1.62)	
$d^{cf}_{t+1} - p_t$	-4.13***		.94***	14.44%
	(-40.49)		(54.80)	
ret_{t+1}	-.00	1.57***	$-.01^{\dagger}$	0.12%
	(-0.09)	(3.97)	(-0.63)	
$p-value$ of restriction	.13	.98	.60	

Panel B: Estimates and tests for present value model (II)

	Const.	$\Delta\mu_{dt+1}$	$\overline{d_t - p_t}$	Δd^{cf}_t	R^2
Δd^{cf}_{t+1}	-.24**	30.28	-.04**	-.36***	12.87%
	(-2.31)	(0.66)	(-2.54)	-58.32	
$\Delta\left(d^{cf}_{t+1} - p_{t+1}\right)$	-.23**	28.86	-.04	-0.36***	12.87%
	(-2.31)	(0.63)	(-2.54)	(-58.32)	
ret_{t+1}	.00	1.45***	.00	-.00	0.11%
	(0.87)	(3.74)	(0.28)	(-0.14)	
$p-value$ of restriction	0.36	0.94	0.13	0.75	

Panel C: Heteroskedasticity test

	Coeff.	*Std. Err.*	*t-stat*	*p-value*
$Intercept$	0.00***	0.00	11.81	0.00
$\varepsilon^2_{r,t}$	0.17***	0.06	3.11	0.00
σ^2_{t+1}	0.00***	0.00	5.56	0.00
$\Delta\mu_{dt+1}$	-0.03	0.02	-1.55	0.12

Panel D: Jump Test

	Coeff.	*Std. Err.*	*t-stat*	*p-value*
$Intercept$	0.04***	0.00	21.49	0.00
$\xi_{dt+1}J_{dt+1}$	-0.02*	0.01	-1.70	0.09
σ^2_{dt+1}	0.31***	0.05	6.82	0.00
$\Delta\mu_{dt+1}$	-23.98**	9.59	-2.50	0.01

Table 9: Estimates and diagnostic tests for the contemporaneous present value model fitted to daily data. Panel A reports parameter estimates and t-statistics for the contemporaneous present value model fitted to daily data:

$$p_{t+1} - p_t = \pi_0 + \pi_\mu \Delta\mu_{dt+1} + \pi_{dp}(\overline{d_t - p_t}) + \varepsilon_{pt+1}$$

$$d^{cf}_{t+1} - p_t = \delta_0 + \delta_{dp}(\overline{d_t - p_t}) + \varepsilon^{cf}_{t+1}$$

$$r_{t+1} = \lambda_0 + \lambda_\mu \Delta\mu_{dt+1} + \lambda_{dp}(\overline{d_t - p_t}) + \varepsilon_{rt+1}$$

with the following restrictions implied by the model

$$\lambda_0 = k + \rho\pi_0 + (1-\rho)\delta_0,\ \lambda_\mu = \rho\pi_\mu,\ \lambda_{dp} = \rho\pi_{dp} + (1-\rho)\delta_{dp}$$

The present value model in Panel B follows the same layout as in Table 6, with the only difference that we have replaced the lagged term $\Delta\mu_{dt}$ with the contemporaneous value $\Delta\mu_{dt+1}$.
Panel C reports estimates from the following estimates on the squared residuals from the return model:

$$\varepsilon^2_{rt+1} = a_0 + a_1\varepsilon^2_{rt} + a_2\sigma^2_{t+1} + a_3\Delta\mu_{dt+1} + \varepsilon^{\sigma}_{t+1}$$

Panel D reports the OLS estimates of the following jump test

$$I_{t+1} = b_0 + b_1\xi_{dt+1}J_{dt+1} + b_2\sigma^2_{dt+1} + b_3\Delta\mu_{dt+1} + \varepsilon^I_{t+1}$$

In all cases, the sample period is 1927-2016. †: coefficient has been multiplied by 100.

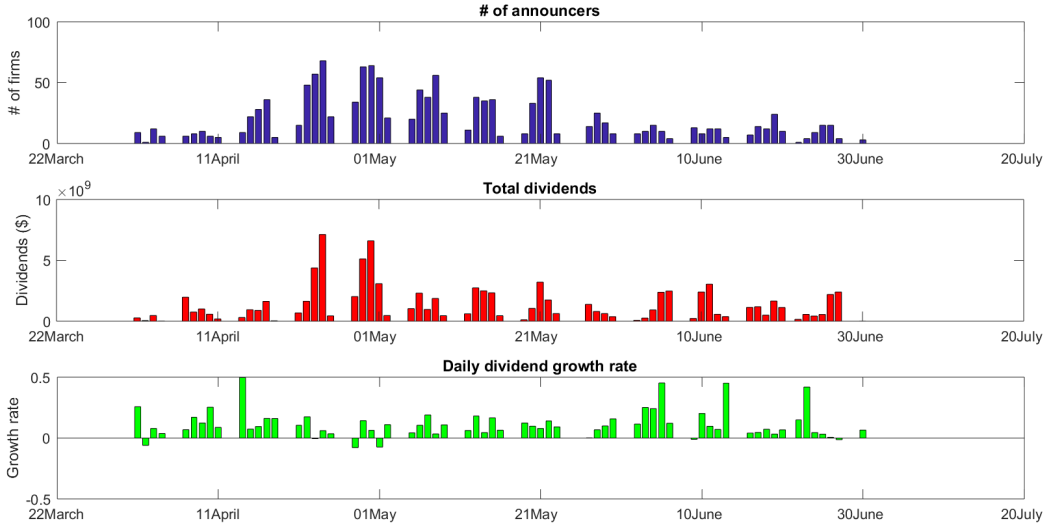

Figure 1: Distribution of dividend announcements within a quarter. This figure plots time-series of dividend announcements for Q2 2014. For every day within this quarter, the top panel shows the number of firms announcing dividends. The middle panel shows the overall nominal amount of dividends announced by those firms (in billion dollars), while the bottom panel shows the daily (net) dividend growth rate.

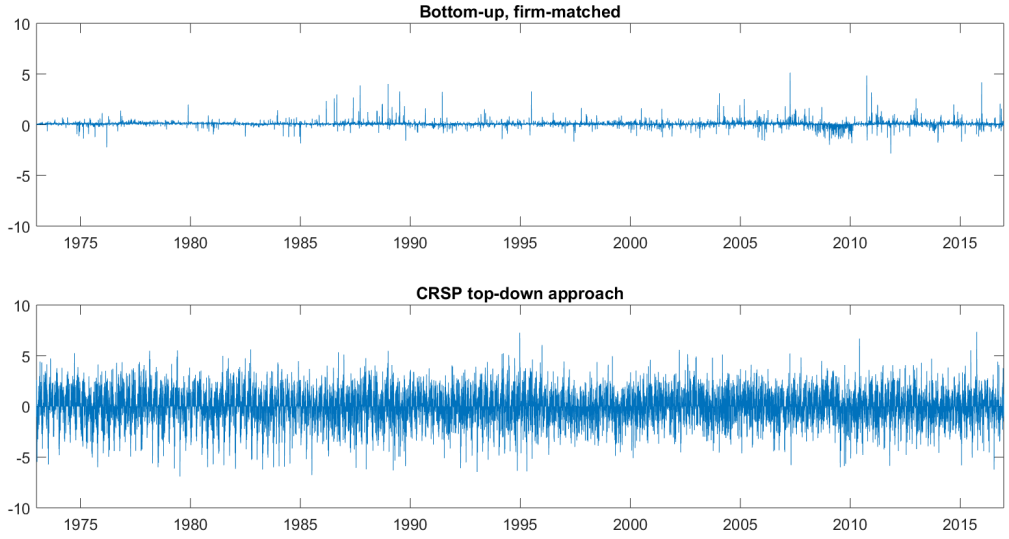

Figure 2: Comparison between our daily "bottom-up" dividend growth series vs. a daily "top-down" dividend growth measure extracted from CRSP. The top panel plots the log of the daily dividend growth series G_t defined in Eq.1. The bottom panel plots the CRSP-extracted daily dividend growth, calculated as dividends (paid out) on an given day divided by dividends distributed on the same day, one year earlier. Both plots use daily data over the sample 1973-2016.

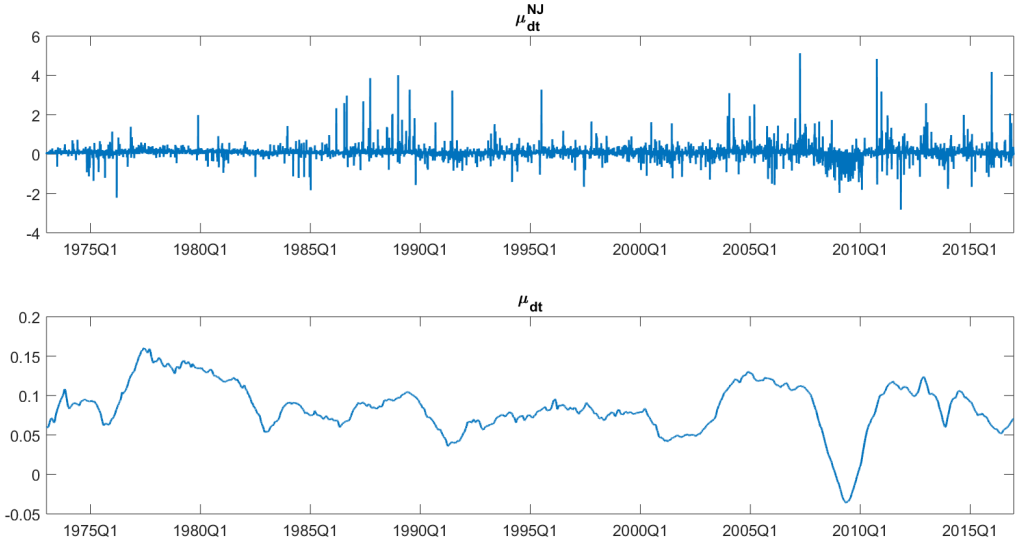

Figure 3: Time series of daily dividend growth and the persistent growth component. The top panel plots the persistent dividend growth component, μ_{dt}^{NJ}, extracted from a model without jumps and stochastic volatility. The bottom panel plots the persistent dividend growth component, μ_{dt}, extracted from the daily dividend series using a model that accounts for jumps and stochastic volatility. All plots use daily data over the sample 1973-2016.

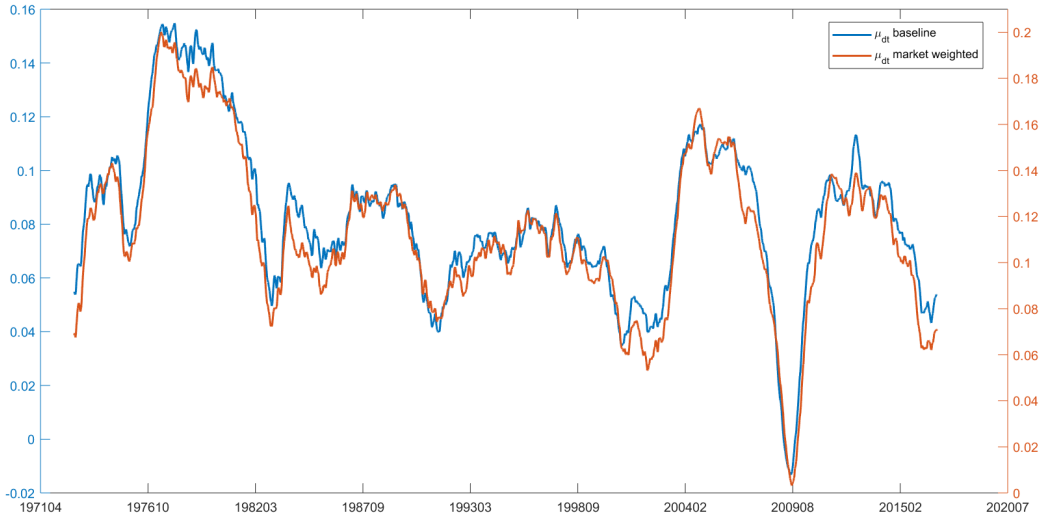

Figure 4: This figure plots estimates of the persistent dividend growth measure μ_{dt} extracted from our model using dollar-weighted (baseline) and market-weighted (market weighted) dividend growth data.

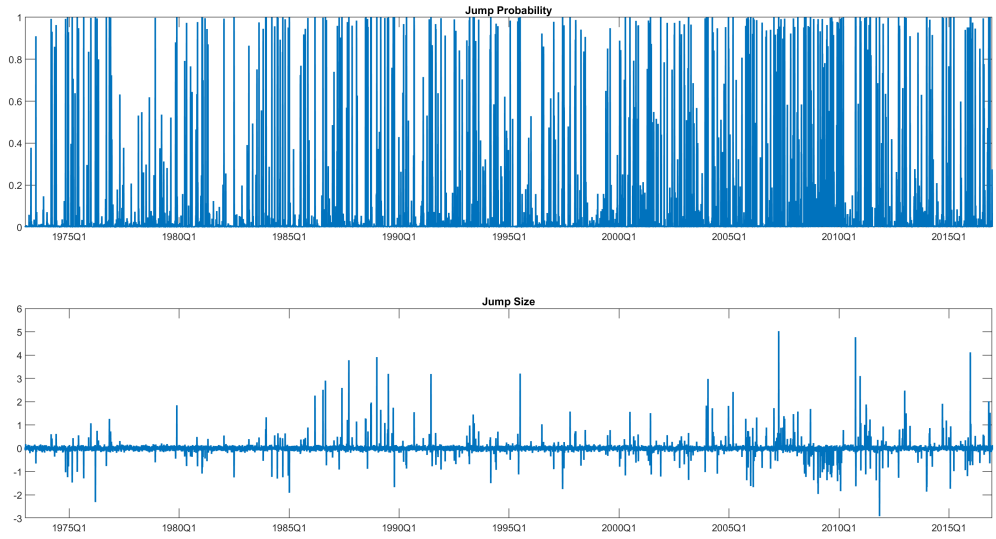

Figure 5: Jumps in the daily dividend growth series. The top panel plots the probability of a jump in the daily dividend growth series while the bottom panel plots the magnitude of such jumps. Both plots use daily dividend data over the sample 1973-2016.

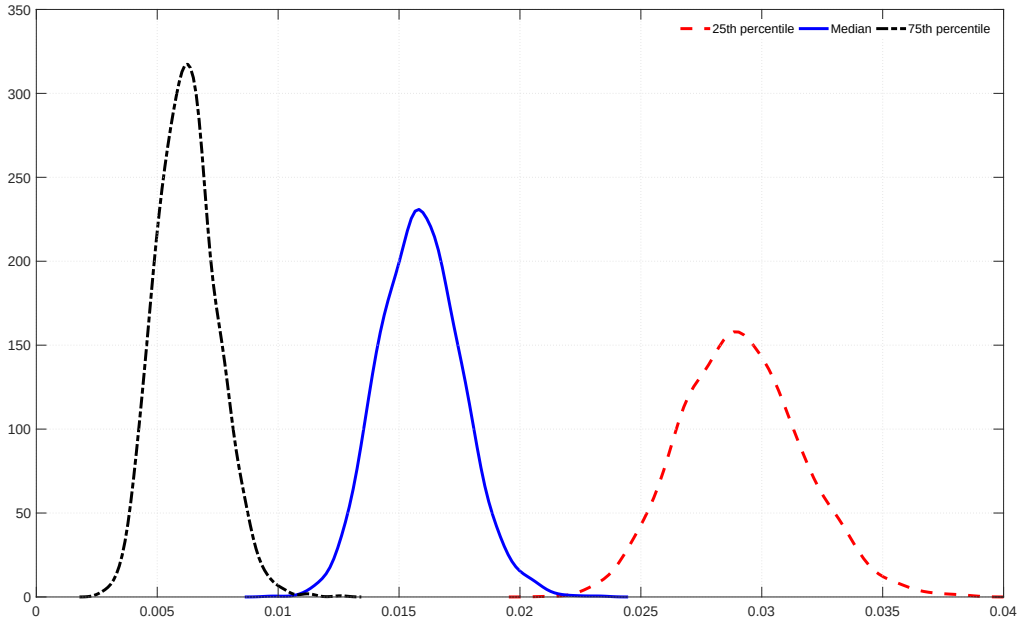

Figure 6: Jump intensities and the number of firms announcing dividend news. This figure shows the sensitivity of the dividend growth jump probability to the number of firms announcing dividends on a given day, N_{dt}, chosen to match the 25th, median and 75th percentiles of the distribution of the daily number of firms announcing dividends. On days with a large number of announcing firms (black, dashed curve), the jump intensity distribution is centered around 0.005, corresponding to a jump on average every 200 days. On days with a typical (median) number of announcing firms (blue curve), the jump intensity is centered around 0.016, implying a jump roughly every 60 days. Finally, on days with a small number of announcing firms (red, dotted curve), the probability of a jump is 0.03, corresponding to a jump on average every 35 days.

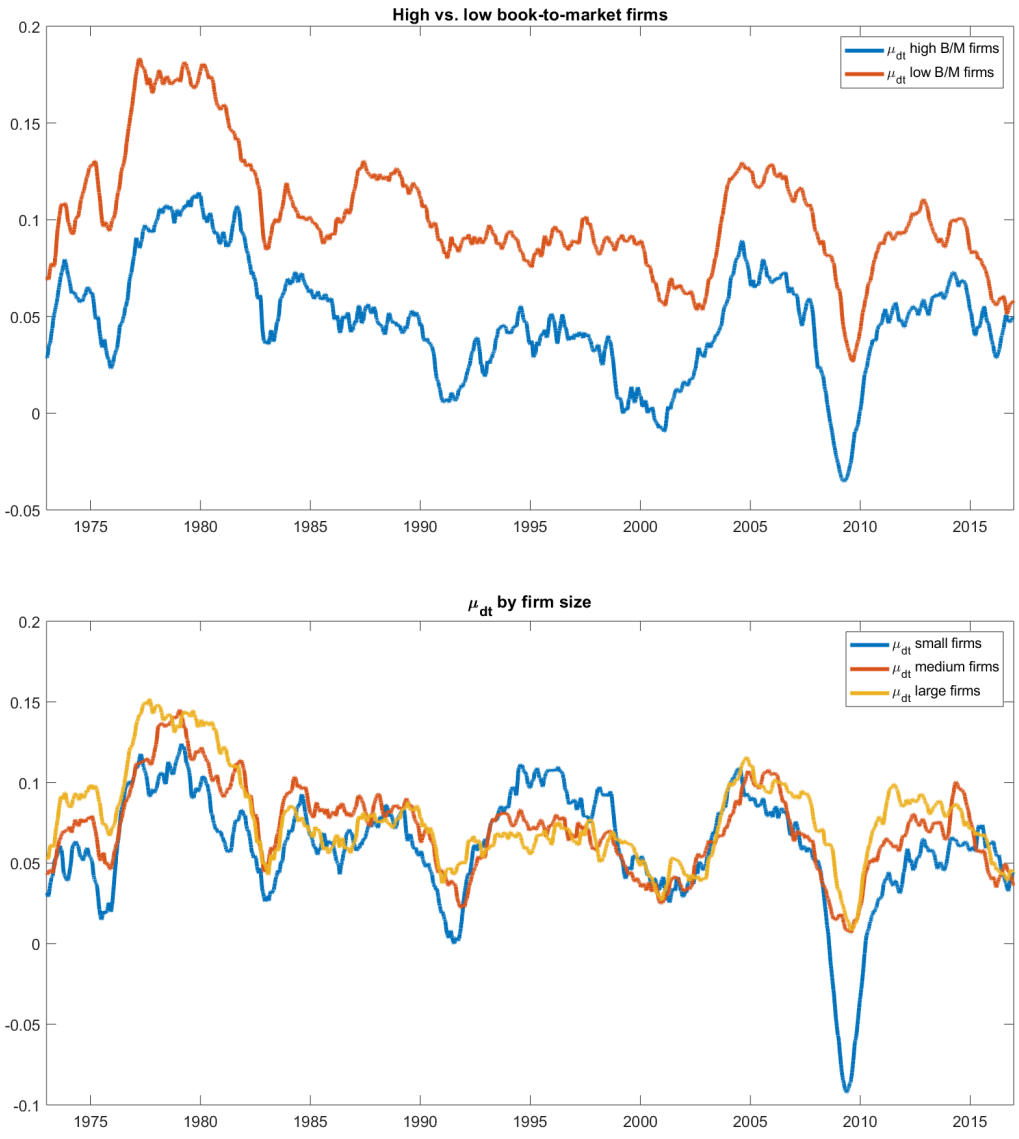

Figure 7: Heterogeneity in μ_{dt} and firm characteristics. The top panel plots estimates of the persistent dividend growth component, μ_{dt}, computed separately for samples of firms with either high or low book-to-market ratios. The bottom panel plots estimates of the persistent dividend growth component, μ_{dt}, computed separately for small, medium and large firms. All estimations use daily data over the sample period 1973-2016.

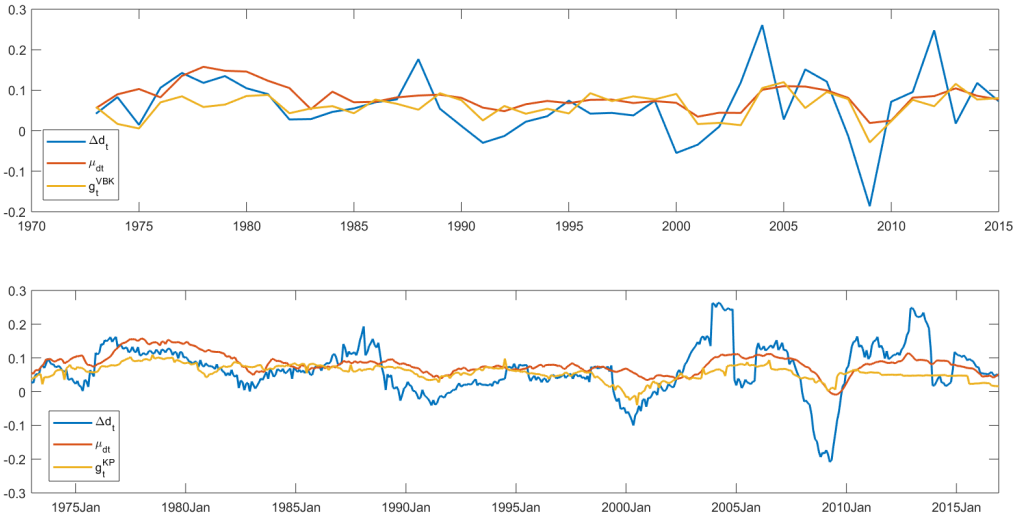

Figure 8: Actual versus predicted dividend growth under alternative modeling approaches. The top panel plots the actual dividend growth, Δd_t, against the persistent dividend growth component extracted from our model, μ_{dt}, and the measure proposed by van Binsbergen and Koijen (2010), g_t^{VBK}. The latter assumes cash reinvested dividend growth. The bottom panel plots actual dividend growth against our persistent dividend growth component and the measure of Kelly and Pruitt (2013), g_t^{KP}. In both cases we have extended the sample period originally used by the papers after replicating their results.

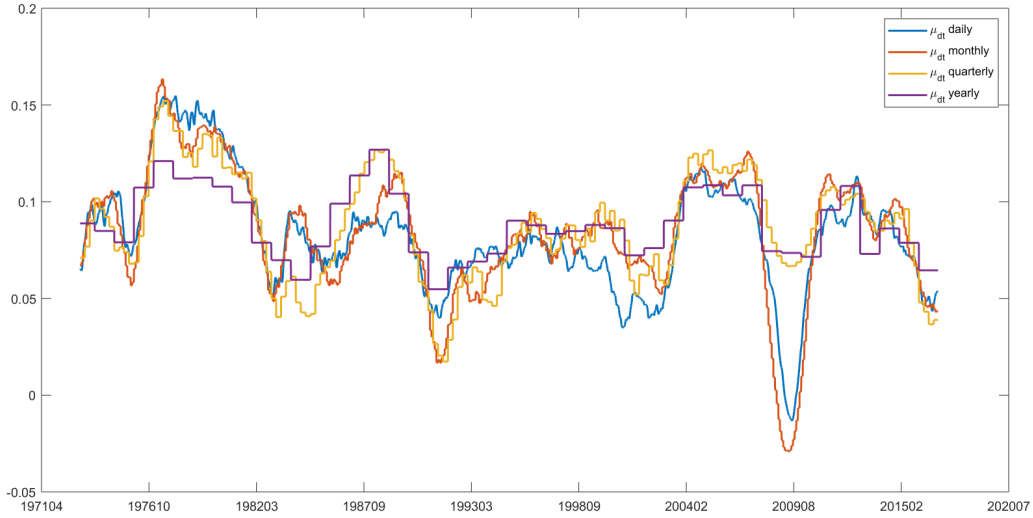

Figure 9: Comparison of estimates of the persistent dividend growth component, μ_{dt}, extracted from data at the daily, monthly, quarterly, and annual frequencies. This figure plots time-series of the μ_{dt} estimated at the daily, monthly, quarterly and annual frequencies.

Internet Appendix A MCMC Algorithm

In this Appendix, we provide the analytical derivations needed to compute the posterior distribution of all parameters and latent states of the model we describe in Section 3 of the paper.

A.1 The Model

We start by rewriting both the model as well as the priors distributions for all model parameters. Starting with the observation equation and time-varying mean and volatility processes, we have

$$\Delta d_{t+1} = \mu_{dt+1} + \xi_{dt+1} J_{dt+1} + \varepsilon_{dt+1}, \tag{A.1}$$

$$\mu_{dt+1} = \mu_d + \phi_\mu \left(\mu_{dt} - \mu_d\right) + \sigma_\mu \varepsilon_{\mu t+1}, \tag{A.2}$$

and

$$h_{dt+1} = \mu_h + \phi_h \left(h_{dt} - \mu_h\right) + \sigma_h \varepsilon_{ht+1} \tag{A.3}$$

where $\varepsilon_{dt+1} \sim \mathcal{N}\left(0, e^{h_{dt+1}}\right)$, $\varepsilon_{\mu t+1} \sim \mathcal{N}\left(0,1\right)$, and $\varepsilon_{ht+1} \sim \mathcal{N}\left(0,1\right)$ independent among each other and across time. The jump process intensity and size follow

$$Pr(J_{dt+1} = 1) = \Phi(\boldsymbol{\lambda}' \boldsymbol{X}^J_{t+1}) \tag{A.4}$$

and

$$\xi_{dt+1} \sim \mathcal{N}(0, \sigma^2_\xi) \tag{A.5}$$

with $\boldsymbol{X}^J_{t+1}$ exogenous. Finally, the initial conditions for μ_{dt} and h_{dt} are as follows:

$$\mu_{d1} \sim \mathcal{N}\left(\mu_d, \frac{\sigma^2_\mu}{1-\phi^2_\mu}\right) \tag{A.6}$$

and

$$h_{d1} \sim \mathcal{N}\left(\mu_h, \frac{\sigma_h^2}{1-\phi_h^2}\right). \tag{A.7}$$

A.2 Priors

The model in (A.1)-(A.7) includes eight parameters, namely μ_d, ϕ_μ, σ_μ^2, μ_h, ϕ_h, σ_h^2, $\boldsymbol{\lambda}$, and σ_ξ^2. We specify the following prior distributions:

$$\mu_d \sim \mathcal{N}(\underline{\mu}_{d0}, \underline{V}_{\mu_d}), \quad \phi_\mu \sim \mathcal{N}(\underline{\phi}_{\mu 0}, \underline{V}_{\phi_\mu}) I(|\phi_\mu| < 1), \quad \sigma_\mu^2 \sim \mathcal{IG}(\underline{\nu}_\mu, \underline{S}_\mu) \tag{A.8}$$

$$\mu_h \sim \mathcal{N}(\underline{\mu}_{h0}, \underline{V}_{\mu_h}), \quad \phi_h \sim \mathcal{N}(\underline{\phi}_{h_0}, \underline{V}_{\phi_h}) I(|\phi_h| < 1), \quad \sigma_h^2 \sim \mathcal{IG}(\underline{\nu}_h, \underline{S}_h) \tag{A.9}$$

$$\boldsymbol{\lambda} \sim \mathcal{N}(\underline{\boldsymbol{\mu}}_{\boldsymbol{\lambda}}, \underline{\boldsymbol{V}}_{\boldsymbol{\lambda}}) \tag{A.10}$$

$$\sigma_\xi^2 \sim \mathcal{IG}(\underline{\nu}_\xi, \underline{S}_\xi) \tag{A.11}$$

where $\mathcal{N}$ denotes a normal distribution while $\mathcal{IG}$ stands for the Inverted-Gamma distribution.

Next, we briefly describe the choices of prior hyperparameters for the dividend growth model, and note that we work, whenever possible, with loose and mildly informative priors.[40] Starting with μ_d and μ_h, we set $\underline{\mu}_{d0} = \underline{\mu}_{h0} = 0$ and specify their variance as $\underline{V}_{\mu_d} = \underline{V}_{\mu_h} = 10$. Next, we set $\underline{\phi}_{\mu 0} = \underline{\phi}_{h0} = 0.98$ and $\underline{V}_{\phi_\mu} = \underline{V}_{\phi_h} = 0.1^2$, which implies a prior belief that the latent processes for μ_{dt+1} and h_{dt+1} will be very persistent. Further, we set $\underline{S}_\mu = 0.01^2$ and $\underline{S}_h = 0.1^2$, restricting the changes to the process for μ_{dt+1} and h_{dt+1} to be 0.01 and 0.1 on average, but we also set their degrees of freedom $\underline{\nu}_\mu$ and $\underline{\nu}_h$ to 2, which provides the least informative proper priors on these parameters. For the timing and intensity of the jumps, we specify a normal prior for the parameters governing the timing of the jumps, i.e. $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)' \sim \mathcal{N}\left(\underline{\boldsymbol{\mu}}_{\boldsymbol{\lambda}}, \underline{\boldsymbol{V}}_{\boldsymbol{\lambda}}\right)$, where $\underline{\boldsymbol{\mu}}_{\boldsymbol{\lambda}} = (0,0)'$ and $\underline{\boldsymbol{V}}_{\boldsymbol{\lambda}} = diag\,(10,10)$. Finally, for the magnitude of the jumps, we set $\underline{\nu}_\xi = 1,000$ and $\underline{S}_\xi = 3^2 \times \left(\underline{\nu}_\xi - 1\right)$, which implies a prior

[40] Note that we impose the stationarity conditions $|\phi_\mu| < 1$ and $|\phi_h| < 1$ directly on the priors.

belief that on average the jump magnitude is equal to 3 and a weight given to this prior that is approximately 10% of the weight put on the data.

A.3 Posteriors

We now describe how to obtain posterior estimates for all model parameters (μ_d, ϕ_μ, σ_μ^2, μ_h, ϕ_h, σ_h^2, $\boldsymbol{\lambda}$, σ_ξ^2), as well as latent state vectors $\boldsymbol{\mu}_d = \{\mu_{dt}\}_{t=1}^T$, $\boldsymbol{h}_d = \{h_{dt}\}_{t=1}^T$, $\boldsymbol{J}_d = \{J_{dt}\}_{t=1}^T$, and $\boldsymbol{\xi}_d = \{\xi_{dt}\}_{t=1}^T$. While the joint posterior distribution of all model parameters and latent state variables is non-standard, we can employ a Gibbs sampler algorithm augmented with a number of Metropolis-Hastings steps to draw recursively from the conditional posteriors of all model parameter and state variables. In particular, we sample from the joint posterior distribution in five different blocks, namely:

1. $\boldsymbol{\mu}_d | \, \boldsymbol{h}_d, \boldsymbol{\xi}_d, \boldsymbol{J}_d, \mu_d, \phi_\mu, \sigma_\mu^2, \mathcal{D}^T$
2. $\boldsymbol{J}_d | \, \boldsymbol{\mu}_d, \boldsymbol{\xi}_d, \boldsymbol{h}_d, \mathcal{D}^T$
3. $\boldsymbol{\xi}_d | \, \boldsymbol{\mu}_d, \boldsymbol{J}_d, \boldsymbol{h}_d, \sigma_\xi^2, \mathcal{D}^T$
4. $\boldsymbol{h}_d | \, \boldsymbol{\mu}_d, \boldsymbol{\xi}_d, \boldsymbol{J}_d, \mu_h, \phi_h, \sigma_h^2, \beta, \mathcal{D}^T$
5. $\mu_d, \phi_\mu, \sigma_\mu^2, \mu_h, \phi_h, \sigma_h^2, \boldsymbol{\lambda}, \sigma_\xi^2, \, | \, \boldsymbol{\mu}_d, \boldsymbol{h}_d, \boldsymbol{\xi}_d, \boldsymbol{J}_d, \mathcal{D}^T$

The last block is further broken down into eight separate sub-blocks, one for each element of the parameter vector. We now describe in details all steps of the Gibbs sampler algorithm.

A.3.1 $\boldsymbol{\mu}_d | \, \boldsymbol{h}_d, \boldsymbol{\xi}_d, \boldsymbol{J}_d, \mu_d, \phi_\mu, \sigma_\mu^2, \mathcal{D}^T$

Start by rewriting the observation equation in (A.1) as follows:

$$\Delta \boldsymbol{d}^\star = \boldsymbol{X}_\mu \boldsymbol{\mu_d} + \boldsymbol{\varepsilon}_d \quad \boldsymbol{\varepsilon}_d \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_d) \tag{A.12}$$

where

$$\Delta \boldsymbol{d}^\star = \begin{bmatrix} \Delta d_1 - \xi_{d1} J_{d1} \\ \vdots \\ \Delta d_T - \xi_{dT} J_{dT} \end{bmatrix}, \tag{A.13}$$

$$\boldsymbol{X}_{\mu} = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix} \quad \boldsymbol{\mu_d} = \begin{bmatrix} \mu_{d1} \\ \vdots \\ \mu_{dT} \end{bmatrix} \quad \boldsymbol{\varepsilon}_d = \begin{bmatrix} \varepsilon_{d1} \\ \vdots \\ \varepsilon_{dT} \end{bmatrix} \tag{A.14}$$

and

$$\boldsymbol{\Sigma}_d = \begin{bmatrix} e^{h_{d1}} & & \\ & \ddots & \\ & & e^{h_{dT}} \end{bmatrix}. \tag{A.15}$$

Next, combine the state equation for $\boldsymbol{\mu_d}$ in (A.2) with the initial condition in (A.6) into:

$$\boldsymbol{H}_{\mu}\boldsymbol{\mu_d} = \widetilde{\boldsymbol{\delta}}_{\mu} + \boldsymbol{\varepsilon}_{\mu} \quad \boldsymbol{\varepsilon}_{\mu} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_{\mu}) \tag{A.16}$$

where

$$\boldsymbol{H}_{\mu} = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ -\phi_{\mu} & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & -\phi_{\mu} & 1 \end{bmatrix}, \quad \widetilde{\boldsymbol{\delta}}_{\mu} = \begin{bmatrix} \mu_d \\ (1-\phi_{\mu})\mu_d \\ \vdots \\ (1-\phi_{\mu})\mu_d \end{bmatrix}, \quad \boldsymbol{\varepsilon}_{\mu} = \begin{bmatrix} \varepsilon_{\mu 1} \\ \epsilon_{\mu 2} \\ \vdots \\ \varepsilon_{\mu T} \end{bmatrix} \tag{A.17}$$

and

$$\boldsymbol{\Sigma}_{\mu} = \begin{bmatrix} \frac{\sigma_{\mu}^2}{(1-\phi_{\mu}^2)} & & & \\ & \sigma_{\mu}^2 & & \\ & & \ddots & \\ & & & \sigma_{\mu}^2 \end{bmatrix} \tag{A.18}$$

It is easy to show that

$$\boldsymbol{\mu}_d = \boldsymbol{\delta}_{\mu} + \boldsymbol{H}_{\mu}^{-1}\boldsymbol{\varepsilon}_{\mu} \tag{A.19}$$

where $\boldsymbol{\delta}_{\mu} = \boldsymbol{H}_{\mu}^{-1}\widetilde{\boldsymbol{\delta}}_{\mu}$. It follows that

$$\boldsymbol{\mu}_d \sim \mathcal{N}\left(\boldsymbol{\delta}_{\mu}, \boldsymbol{H}_{\mu}^{-1}\boldsymbol{\Sigma}_{\mu}\left(\boldsymbol{H}_{\mu}^{-1}\right)'\right) \tag{A.20}$$

or

$$\boldsymbol{\mu}_d \sim \mathcal{N}\left(\boldsymbol{\delta}_\mu, \left(\boldsymbol{H}'_\mu \boldsymbol{\Sigma}_\mu^{-1} \boldsymbol{H}_\mu\right)^{-1}\right) \tag{A.21}$$

Finally, combining (A.12) and (A.21) leads to the following posterior:

$$\boldsymbol{\mu}_d | \, \boldsymbol{h}_d, \boldsymbol{\xi}_d, \boldsymbol{J}_d, \mu_d, \phi_\mu, \sigma_\mu^2, \mathcal{D}^T \sim \mathcal{N}(\overline{\boldsymbol{\mu}}, \overline{\boldsymbol{V}}_\mu) \tag{A.22}$$

where

$$\begin{aligned} \overline{\boldsymbol{V}}_\mu &= \left[\boldsymbol{H}'_\mu \boldsymbol{\Sigma}_\mu^{-1} \boldsymbol{H}_\mu + \boldsymbol{X}'_\mu \boldsymbol{\Sigma}_d^{-1} \boldsymbol{X}_\mu\right]^{-1} \\ \overline{\boldsymbol{\mu}} &= \overline{\boldsymbol{V}}_\mu \left[(\boldsymbol{H}'_\mu \boldsymbol{\Sigma}_\mu^{-1} \boldsymbol{H}_\mu)\boldsymbol{\delta}_\mu + \boldsymbol{X}'_\mu \boldsymbol{\Sigma}_d^{-1} \Delta \boldsymbol{d}^\star\right] \end{aligned} \tag{A.23}$$

A.3.2 $\boldsymbol{J}_d | \, \boldsymbol{\mu}_d, \boldsymbol{\xi}_d, \boldsymbol{h}_d, \mathcal{D}^T$

It is easy to show that for any given $t \in [1, T]$

$$\begin{aligned} \Pr\left(J_{dt} = 1 | \, \mu_{dt}, \xi_{dt}, \boldsymbol{\lambda}, \boldsymbol{X}_t^J, h_{dt}, \mathcal{D}^T\right) \propto\ & p(\Delta d_t | \, \mu_{dt}, \xi_{dt}, J_{dt} = 1, h_{dt}) \\ & \times \Pr(J_{dt} = 1 | \boldsymbol{X}_t^J, \boldsymbol{\lambda}) \end{aligned} \tag{A.24}$$

where

$$\Delta d_t | \mu_{dt}, \xi_{dt}, J_{dt} = 1, h_{dt} \sim \mathcal{N}\left(\Delta d_t | \, \mu_{dt} + \xi_{dt}, e^{h_{dt}}\right) \tag{A.25}$$

and $\Pr\left(J_{dt} = 1 | \boldsymbol{X}_t^J, \boldsymbol{\lambda}\right) = \Phi\left(\boldsymbol{\lambda}' \boldsymbol{X}_t^J\right)$ while

$$\begin{aligned} \Pr\left(J_{dt} = 0 | \, \mu_{dt}, \xi_{dt}, \boldsymbol{\lambda}, \boldsymbol{X}_t^J, h_{dt}, \mathcal{D}^T\right) \propto\ & p\left(\Delta d_t | \, \mu_{dt}, \xi_{dt}, J_{dt} = 0, h_{dt}\right) \\ & \times \Pr\left(J_{dt} = 0 | \, \boldsymbol{X}_t^J, \boldsymbol{\lambda}\right) \end{aligned} \tag{A.26}$$

where

$$\Delta d_t | \, \mu_{dt}, \xi_{dt}, J_{dt} = 0, h_{dt} \sim \mathcal{N}\left(\Delta d_t | \, \mu_{dt}, e^{h_{dt}}\right) \tag{A.27}$$

and $\Pr\left(J_{dt} = 0 | \, \boldsymbol{X}_t^J, \boldsymbol{\lambda}\right) = 1 - \Phi\left(\boldsymbol{\lambda}' \boldsymbol{X}_t^J\right)$.

A.3.3 $\boldsymbol{\xi}_d|\,\boldsymbol{\mu}_d, \boldsymbol{J}_d, \boldsymbol{h}_d, \sigma_\xi^2, \mathcal{D}^T$

Start by noting that when $J_{dt} = 0$, $\xi_{dt}|\,J_{dt} = 0, \mathcal{D}^T \sim \mathcal{N}(0, \sigma_\xi^2)$. In other words, when $J_{dt} = 0$ we rely on ξ_{dt}'s prior distribution in (A.5). In contrast, when $J_{dt} = 1$, it is possible to rewrite the observation equation of the model in (A.1) as

$$\Delta d_t - \mu_{dt} = \xi_{dt} + \varepsilon_{dt}, \quad \varepsilon_{dt} \sim \mathcal{N}(0, e^{h_{dt}}). \tag{A.28}$$

Combining (A.28) with (A.5) leads to:

$$\xi_{dt}|\,\mu_{dt}, J_{dt} = 1, h_{dt}, \sigma_\xi^2, \mathcal{D}^T \sim \mathcal{N}(\overline{\mu}_{\xi_{dt}}, \overline{\sigma}^2_{\xi_{dt}}) \tag{A.29}$$

where

$$\begin{aligned} \overline{\sigma}^2_{\xi_{dt}} &= \left(\sigma_\xi^{-2} + e^{-h_{dt}}\right)^{-1} \\ \overline{\mu}_{\xi_{dt}} &= \overline{\sigma}^2_{\xi_{dt}} \left(e^{-h_{dt}} \left(\Delta d_t - \mu_{dt}\right)\right). \end{aligned} \tag{A.30}$$

A.3.4 $\boldsymbol{h}_d|\,\boldsymbol{\mu}_d, \boldsymbol{\xi}_d, \boldsymbol{J}_d, \mu_h, \phi_h, \sigma_h^2, \mathcal{D}^T$

Start by combining the state equation for h_{dt} in (A.3) with the initial condition for h_{d1} in (A.7) into:

$$\boldsymbol{H}_h \boldsymbol{h_d} = \widetilde{\boldsymbol{\delta}}_h + \boldsymbol{\varepsilon}_h, \quad \boldsymbol{\varepsilon}_h \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_h) \tag{A.31}$$

where

$$\boldsymbol{H}_h = \begin{bmatrix} 1 & 0 & \dots & \dots & \dots & 0 \\ -\phi_h & 1 & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & \dots & 0 & -\phi_h & 1 \end{bmatrix}, \quad \widetilde{\boldsymbol{\delta}}_h = \begin{bmatrix} \mu_h \\ (1-\phi_h)\mu_h \\ \dots \\ (1-\phi_h)\mu_h \end{bmatrix}, \quad \boldsymbol{\varepsilon}_h = \begin{bmatrix} \varepsilon_{h1} \\ \varepsilon_{h2} \\ \dots \\ \varepsilon_{hT} \end{bmatrix} \tag{A.32}$$

and

$$\boldsymbol{\Sigma}_h = \begin{bmatrix} \frac{\sigma_h^2}{(1-\phi_h^2)} & & & \\ & \sigma_h^2 & & \\ & & \ddots & \\ & & & \sigma_h^2 \end{bmatrix} \tag{A.33}$$

This leads to

$$\boldsymbol{h}_d \sim \mathcal{N}\left(\boldsymbol{\delta}_h, \left(\boldsymbol{H}_h' \boldsymbol{\Sigma}_h^{-1} \boldsymbol{H}_h\right)^{-1}\right) \tag{A.34}$$

where $\boldsymbol{\delta}_h = \boldsymbol{H}_h^{-1} \widetilde{\boldsymbol{\delta}}_h$. Note next that the observation equation is a non-linear function in $\boldsymbol{h}$, so using (A.12) we first rewrite it as follows:

$$\log\left(\Delta d_t^{\star\star}\right)^2 = h_{dt} + \log \varepsilon_{dt}^2, \quad t = 1, ..., T \tag{A.35}$$

where $\Delta d_t^{\star\star} = \Delta d_t^{\star} - \mu_{dt}$. We follow Kim et al. (1998) and approximate $\log \varepsilon_{dt}^2$ with a mixture of normal distributions,

$$\log \varepsilon_{dt}^2 \approx \sum_{j=1}^{7} q_j \times \mathcal{N}\left(m_j - 1.2704, v_j^2\right) \tag{A.36}$$

where m_j, v_j^2, and q_j are constant specified in Kim et al. (1998). Along with (A.36), we also introduce a vector of state variables $\boldsymbol{s}_d = \{s_{dt}\}_{t=1}^T$ such that $\Pr\left(s_{dt} = j\right) = q_j$, for $j = 1, .., 7$ and $t = 1, ..., T$. Conditional on a particular realization of this vector of state variables, we can rewrite the observation equation in (A.35) in compact form as follows:

$$\log\left(\Delta \boldsymbol{d}^{\star\star}\right)^2 \big| \, \boldsymbol{\mu}_d, \boldsymbol{h}_d, \boldsymbol{J}_d, \boldsymbol{\xi}_d, \boldsymbol{s}_d \sim \mathcal{N}(\boldsymbol{m} + \boldsymbol{h}_d, \boldsymbol{V}). \tag{A.37}$$

where

$$\boldsymbol{m} = \begin{bmatrix} m_{s_{d1}} - 1.2704 \\ m_{s_{d2}} - 1.2704 \\ \vdots \\ m_{s_{dT}} - 1.2704 \end{bmatrix}, \quad \boldsymbol{V} = \begin{bmatrix} v_{s_{d1}}^2 & & & \\ & v_{s_{d2}}^2 & & \\ & & \ddots & \\ & & & v_{s_{dT}}^2 \end{bmatrix}. \tag{A.38}$$

Combining (A.37) with (A.34) leads to the following posterior for $\boldsymbol{h}_d$:

$$\boldsymbol{h}_d \big| \, \boldsymbol{\mu}_d, \boldsymbol{\xi}_d, \boldsymbol{J}_d, \boldsymbol{s}_d, \mathcal{D}^T \sim \mathcal{N}\left(\boldsymbol{K}_h^{-1} \boldsymbol{k}_h, \boldsymbol{K}_h^{-1}\right) \tag{A.39}$$

where

$$\boldsymbol{K}_h = \boldsymbol{H}_h' \boldsymbol{\Sigma}_h^{-1} \boldsymbol{H}_h + \boldsymbol{V} \tag{A.40}$$

$$\boldsymbol{k}_h = \boldsymbol{H}_h' \boldsymbol{\Sigma}_h^{-1} \boldsymbol{H}_h \boldsymbol{\delta}_h + \boldsymbol{V}^{-1} \left(\log \left(\Delta \boldsymbol{d}^{\star\star} \right)^2 - \boldsymbol{m} \right) \tag{A.41}$$

As for drawing the vector of state variables $\boldsymbol{s}_d$, note that

$$\Pr\left(s_{dt} = j \middle| \, \mu_{dt}, \xi_{dt}, J_{dt}, h_{dt}, \mathcal{D}^T \right) = \frac{q_j \times f_{\mathcal{N}} \left(\log \left(\Delta d_t^{**} \right)^2 \middle| \, h_{dt} + m_j - 1.2704, v_j^2 \right)}{\sum_{l=1}^{7} q_l \times f_{\mathcal{N}} \left(\log \left(\Delta d_t^{**} \right)^2 \middle| \, h_{dt} + m_l - 1.2704, v_l^2 \right)} \tag{A.42}$$

where $j = 1, ..., 7$, $t = 1, ..., T$, and $f_{\mathcal{N}}\left(y \middle| \, a, b \right)$ denotes the kernel of a normal distribution with mean a and variance b evaluated at y.

A.3.5 $\mu_d, \phi_\mu, \sigma_\mu^2, \mu_h, \phi_h, \sigma_h^2, \boldsymbol{\lambda}, \sigma_\xi^2, \left| \, \boldsymbol{\mu}_d, \boldsymbol{h}_d, \boldsymbol{\xi}_d, \boldsymbol{J}_d, \mathcal{D}^T \right.$

We break this posterior into eight separate blocks:

- $\mu_d \left| \, \boldsymbol{\mu}_d, \phi_\mu, \sigma_\mu^2, \mathcal{D}^T \right.$:

  Start by combining (A.2) and (A.6) and rewriting them as:

  $$\boldsymbol{Z}_\mu = \boldsymbol{X}_\mu \mu_d + \boldsymbol{\varepsilon}_\mu \quad \boldsymbol{\varepsilon}_\mu \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_\mu) \tag{A.43}$$

  where

  $$\boldsymbol{Z}_\mu = \begin{bmatrix} \mu_{d1} \\ \mu_{d2} - \phi_\mu \mu_{d1} \\ \vdots \\ \mu_{dT} - \phi_\mu \mu_{dT-1} \end{bmatrix}, \quad \boldsymbol{X}_\mu = \begin{bmatrix} 1 \\ (1 - \phi_\mu) \\ \vdots \\ (1 - \phi_\mu) \end{bmatrix}. \tag{A.44}$$

  Combining (A.43) with the prior for μ_d in (A.8) leads to

  $$\mu_d \left| \, \boldsymbol{\mu}_d, \phi_\mu, \sigma_\mu^2, \mathcal{D}^T \right. \sim \mathcal{N}(\overline{\mu}_d, \overline{V}_{\mu_d}) \tag{A.45}$$

where

$$\overline{V}_{\mu_d} = \left[\underline{V}_{\mu_d}^{-1} + \boldsymbol{X}_{\mu}'\boldsymbol{\Sigma}_{\mu}^{-1}\boldsymbol{X}_{\mu}\right]^{-1} \tag{A.46}$$

and

$$\overline{\mu}_d = \overline{V}_{\mu_d}\left[\underline{V}_{\mu_d}^{-1}\underline{\mu}_{d0} + \boldsymbol{X}_{\mu}'\boldsymbol{\Sigma}_{\mu}^{-1}\boldsymbol{Z}_{\mu}\right] \tag{A.47}$$

- $\phi_\mu | \, \boldsymbol{\mu}_d, \mu_d, \sigma_\mu^2, \mathcal{D}^T$:

  Following Kim et al. (1998), we start by obtaining a candidate draw from the following distribution:

$$\phi_\mu^\star \sim \mathcal{N}\left(\overline{\phi}_\mu, \overline{V}_\mu\right) \times I\left(|\phi_\mu| < 1\right) \tag{A.48}$$

  where

$$\overline{V}_{\phi_\mu} = \left(\underline{V}_{\phi_\mu}^{-1} + \frac{\boldsymbol{X}_{\phi_\mu}'\boldsymbol{X}_{\phi_\mu}}{\sigma_\mu^2}\right)^{-1}, \tag{A.49}$$

$$\overline{\phi}_\mu = \overline{V}_{\phi_\mu}\left(\underline{V}_{\phi_\mu}^{-1}\underline{\phi}_{\mu_0} + \frac{\boldsymbol{X}_{\phi_\mu}'\boldsymbol{Z}_{\phi_\mu}}{\sigma_\mu^2}\right) \tag{A.50}$$

  and where

$$\boldsymbol{Z}_{\phi_\mu} = \begin{bmatrix} \mu_{d2} - \mu_d \\ \vdots \\ \mu_{dT} - \mu_d \end{bmatrix}, \quad \boldsymbol{X}_{\phi_\mu} = \begin{bmatrix} \mu_{d1} - \mu_d \\ \vdots \\ \mu_{dT-1} - \mu_d \end{bmatrix}. \tag{A.51}$$

  Next, if the draw is retained (i.e., satisfy the stationarity restriction), we accept ϕ_μ^* with probability $e^{\left(g\left(\phi_\mu^*\right) - g\left(\phi_\mu^{old}\right)\right)}$ where ϕ_μ^{old} is the retained draw from the previous iteration of the Gibbs sampler, and

$$g\left(\phi_\mu\right) = \ln p\left(\phi_\mu\right) - \frac{1}{2}\ln\left(\frac{\sigma_\mu^2}{1-\phi_\mu^2}\right) - \frac{\left(1-\phi_\mu^2\right)}{2\sigma_\mu^2}\left(\mu_{d1} - \mu_d\right)^2 \tag{A.52}$$

with $p(\phi_\mu)$ denoting the prior of ϕ_μ from (A.8).

- $\sigma_\mu^2 \big| \boldsymbol{\mu}_d, \mu_d, \phi_\mu, \mathcal{D}^T$:

  The posterior for σ_μ^2 is readily available, and is given by:

$$\sigma_\mu^2 \big| \boldsymbol{\mu}_d, \mu_d, \phi_\mu, \mathcal{D}^T \sim \mathcal{IG}\left(\underline{\nu}_\mu + \frac{T}{2}, \overline{S}_\mu\right) \tag{A.53}$$

  where

$$\overline{S}_\mu = \underline{S}_\mu + \frac{1}{2}\left[\left(1-\phi_\mu^2\right)\left(\mu_{d1}-\mu_d\right)^2 + \sum_{t=1}^{T-1}\left(\mu_{dt+1}-\mu_d-\phi_\mu\left(\mu_{dt}-\mu_d\right)\right)^2\right] \tag{A.54}$$

- $\mu_h | \boldsymbol{h}_d, \phi_h, \sigma_h^2, \mathcal{D}^T$:

  Start by combining (A.3) and (A.7) into:

$$\boldsymbol{Z}_h = \boldsymbol{X}_h \mu_h + \boldsymbol{\varepsilon}_h \quad \boldsymbol{\varepsilon}_h \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_h) \tag{A.55}$$

  where

$$\boldsymbol{Z}_h = \begin{bmatrix} h_{d1} \\ h_{d2} - \phi_h h_{d1} \\ \vdots \\ h_{dT} - \phi_h h_{dT-1} \end{bmatrix}, \quad \boldsymbol{X}_h = \begin{bmatrix} 1 \\ 1-\phi_h \\ \vdots \\ 1-\phi_h \end{bmatrix}. \tag{A.56}$$

  Next, combine (A.55) with the prior for μ_h in (A.9) to get

$$\mu_h | \boldsymbol{h}_d, \phi_h, \sigma_h^2, \mathcal{D}^T \sim \mathcal{N}\left(\overline{\mu}_h, \overline{V}_{\mu_h}\right) \tag{A.57}$$

  where

$$\overline{V}_{\mu_h} = \left[\underline{V}_{\mu_h}^{-1} + \boldsymbol{X}_h' \boldsymbol{\Sigma}_h^{-1} \boldsymbol{X}_h\right]^{-1} \tag{A.58}$$

and

$$\overline{\mu}_h = \overline{V}_{\mu_h}\left[\underline{V}_{\mu_h}^{-1}\underline{\mu}_{h0} + \boldsymbol{X}_h'\boldsymbol{\Sigma}_h^{-1}\boldsymbol{Z}_h\right] \tag{A.59}$$

- $\phi_h|\boldsymbol{h}_d, \mu_h, \sigma_h^2, \mathcal{D}^T$:

  As with ϕ_μ, we follow Kim et al. (1998) and first obtain a candidate draw from the following distribution:

$$\phi_h^\star \sim \mathcal{N}\left(\overline{\phi}_h, \overline{V}_h\right) \times I\left(|\phi_h| < 1\right) \tag{A.60}$$

  where

$$\overline{V}_{\phi_h} = \left(\underline{V}_{\phi_h}^{-1} + \frac{\boldsymbol{X}_{\phi_h}'\boldsymbol{X}_{\phi_h}}{\sigma_h^2}\right)^{-1}, \tag{A.61}$$

$$\overline{\phi}_h = \overline{V}_{\phi_h}\left(\underline{V}_{\phi_h}^{-1}\underline{\phi}_{h_0} + \frac{\boldsymbol{X}_{\phi_h}'\boldsymbol{Z}_{\phi_h}}{\sigma_h^2}\right) \tag{A.62}$$

  and where

$$\boldsymbol{Z}_{\phi_h} = \begin{bmatrix} h_{d2} - \mu_h \\ \vdots \\ h_{dT} - \mu_h \end{bmatrix}, \quad \boldsymbol{X}_{\phi_h} = \begin{bmatrix} h_{d1} - \mu_h \\ \vdots \\ h_{dT-1} - \mu_h \end{bmatrix} \tag{A.63}$$

  Next, if the draw is retained (i.e., satisfy the stationarity restriction), we accept ϕ_h^* with probability $e^{\left(g\left(\phi_h^*\right) - g\left(\phi_h^{old}\right)\right)}$ where ϕ_h^{old} is the retained draw from the previous iteration of the Gibbs sampler, and

$$g\left(\phi_h\right) = \ln p\left(\phi_h\right) - \frac{1}{2}\ln\left(\frac{\sigma_h^2}{1-\phi_h^2}\right) - \frac{\left(1-\phi_h^2\right)}{2\sigma_h^2}\left(h_{d1} - \mu_h\right)^2 \tag{A.64}$$

  with $p\left(\phi_h\right)$ denoting the prior of ϕ_h.

- $\sigma_h^2|\boldsymbol{h}_d, \mu_h, \phi_h\mathcal{D}^T$:

The posterior for σ_h^2 is readily available, and is given by:

$$\sigma_h^2|\boldsymbol{h}, \mu_h, \phi_h, \mathcal{D}^T \sim \mathcal{IG}\left(\underline{\nu}_h + \frac{T}{2}, \overline{S}_h\right) \tag{A.65}$$

where

$$\overline{S}_h = \underline{S}_h + \frac{1}{2}\left[\left(1-\phi_h^2\right)\left(h_{d1}-\mu_h\right)^2 + \sum_{t=1}^{T-1}\left(h_{dt+1}-\mu_h-\phi_h\left(h_{dt}-\mu_h\right)\right)^2\right] \tag{A.66}$$

- $\boldsymbol{\lambda}|\,\boldsymbol{W}, \mathcal{D}^T$ and $\boldsymbol{W}|\,\boldsymbol{\lambda}, \boldsymbol{J}_d, \mathcal{D}^T$:

We follow Albert and Chib (1993) and to simplify the computations introduce the auxiliary latent state variable W_t, $t = 1, ..., T$. We proceed by first rewriting the stochastic process of the jump intenisty in (A.4) as

$$J_{dt+1} = \begin{cases} 1 & \text{if } W_{t+1} > 0 \\ 0 & \text{if } W_{t+1} \leq 0 \end{cases} \tag{A.67}$$

where

$$W_{t+1} = \boldsymbol{\lambda}' \boldsymbol{X}_{t+1}^J + \varepsilon_{Wt+1}, \quad \varepsilon_{Wt+1} \sim \mathcal{N}(0,1) \tag{A.68}$$

or, more compactly,

$$\boldsymbol{W} = \boldsymbol{X}^J \boldsymbol{\lambda} + \boldsymbol{\varepsilon}_W, \quad \boldsymbol{\varepsilon}_W \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_T) \tag{A.69}$$

where

$$\boldsymbol{X}^J = \begin{bmatrix} \boldsymbol{X}_1^{J\prime} \\ \vdots \\ \boldsymbol{X}_T^{J\prime} \end{bmatrix} \text{ and } \boldsymbol{W} = \begin{bmatrix} W_1 \\ \vdots \\ W_T \end{bmatrix}. \tag{A.70}$$

The posterior of $\boldsymbol{\lambda}$ is readily available, and given by

$$\boldsymbol{\lambda}|\,\boldsymbol{W}, \mathcal{D}^T \sim \mathcal{N}\left(\overline{\boldsymbol{\mu}}_\lambda, \overline{\boldsymbol{V}}_\lambda\right) \tag{A.71}$$

where

$$\overline{\boldsymbol{V}}_\lambda = \left[\underline{\boldsymbol{V}}_\lambda^{-1} + \boldsymbol{X}^{J\prime}\boldsymbol{X}^J\right]^{-1} \tag{A.72}$$

and

$$\overline{\boldsymbol{\mu}}_\lambda = \overline{\boldsymbol{V}}_\lambda \left[\underline{\boldsymbol{V}}_\lambda^{-1}\underline{\boldsymbol{\mu}}_\lambda + \boldsymbol{X}^{J\prime}\boldsymbol{W}\right]. \tag{A.73}$$

As for the sequence of latent variables $\{W_t\}_{t=1}^T$, we have that

$$W_t|\,\boldsymbol{\lambda}, J_{dt}, \mathcal{D}^T \sim \begin{cases} TN(\boldsymbol{\lambda}'\boldsymbol{X}_{t+1}^J, 1, 0, \infty) & \text{if } J_{dt} = 1 \\ TN(\boldsymbol{\lambda}'\boldsymbol{X}_{t+1}^J, 1, -\infty, 0) & \text{if } J_{dt} = 0 \end{cases} \tag{A.74}$$

where $TN(\mu, \sigma^2, lb, ub)$ denotes a truncated normal distribution with mean μ, variance σ^2, and lower and upper bounds lb, ub.

- $\sigma_\xi^2|\boldsymbol{\xi}_d, \mathcal{D}^T$:

Finally, the posterior distribution for σ_ξ^2 is readily available, and given by

$$\sigma_\xi^2\big|\,\boldsymbol{\xi}_d, \mathcal{D}^T \sim \mathcal{IG}\left(\underline{\nu}_\xi + \frac{T}{2}, \overline{S}_\xi\right) \tag{A.75}$$

where

$$\overline{S}_\xi = \underline{S}_\xi + \frac{1}{2}\sum_{t=1}^T \xi_{dt}^2. \tag{A.76}$$

Internet Appendix B MCMC Convergence and Efficiency

In this Appendix, we discuss the convergence properties of our MCMC algorithm for the mean-reverting, stochastic volatility model with jumps described in Section 3.1. All results are based on samples of 2,000 retained draws, obtained by sampling a total of 101,000 draws, discarding the first 1,000 draws, and retaining every 20th draw of the post-burn samples.

Table B.1 reports summary statistics of inefficiency factors (IF) for the posterior estimates of all key parameters of the cash flow model. Generally speaking, values of the IFs below 20 are taken as indication that the chain has satisfactory mixing properties. As is clear from the entries in both tables, our algorithm shows excellent mixing properties.

PANEL A: DIVIDENDS			
	IF 4%	*IF 8%*	*IF 15%*
μ_d	0.756	0.607	0.413
ϕ_μ	1.966	2.263	2.161
σ_μ^2	6.586	6.804	4.806
μ_h	0.953	0.908	0.858
ϕ_h	2.337	2.886	3.047
σ_h^2	3.888	4.355	4.615
σ_ξ^2	0.944	0.687	0.543
λ_1	1.204	1.351	0.904
λ_2	2.114	2.693	3.356

PANEL B: DIVIDENDS (from 1927)			
	IF 4%	*IF 8%*	*IF 15%*
μ_d	1.127	1.021	1.001
ϕ_μ	1.849	1.464	1.018
σ_μ^2	9.306	6.666	4.806
μ_h	0.738	0.477	0.353
ϕ_h	5.603	5.813	5.037
σ_h^2	8.144	7.777	6.290
σ_ξ^2	1.822	1.837	1.672
λ_1	3.920	3.786	3.635
λ_2	19.009	18.708	16.129

Table B.1: Inefficiency factors of the model. This table reports the inefficiency factors for the key parameters of the mean-reverting, stochastic volatility model with jumps described in Section 3.1. Panel A reports results for the model using the daily dividend growth series starting in 1973, while Panel B shows estimates using the daily dividend growth series and starting in 1927. For each individual parameter, the inefficiency factor is estimated as $1 + 2\sum_{k=1}^{\infty} \rho_k$ where ρ_k is the kth-order autocorrelation of the chain of retained draws. The estimates use the Newey-West kernel and a bandwidth of 4%, 8%, or 15% of the sample of retained draws. All results are based on a sample of 2,000 retained draws, obtained by sampling a total of 101,000 draws, discarding the first 1,000 and retaining every 20th draw of the post-burn sample.

Internet Appendix C Extensions and robustness tests

This section conducts a set of tests designed to explore some additional implications of our analysis and verify the robustness of our empirical results. First, we analyze whether innovations to our new persistent dividend growth component, μ_{dt}, are related to a range of cross-sectional asset pricing factors that have been studied in the finance literature. Second, we study the relation between stock returns and dividends on dividend payment days, rather than on days where dividends get announced. Third, we study the economic value of return predictability. Fourth, we estimate regressions of different measures of stock market volatility on μ_{dt}. Fifth, we estimate alternative econometric specifications that allow for state dependence in jump probabilities as well as correlated jumps in the mean and volatility processes. Sixth, we estimate our model at the industry level.

C.1 Cross-sectional effects of dividend shocks

Our finding that the behavior of the persistent component of the dividend growth process, μ_{dt}, varies across firms with different sizes and book-to-market ratios suggests that shocks to μ_{dt} may be related to cross-sectional risk factors tied to firm or stock characteristics.

To see if this holds, we next investigate whether μ_{dt} is related to existing risk factors found in the empirical asset pricing literature to capture cross-sectional variation in stock returns. These include the market (MRP), size (SMB), book-to-market (HML), profitability (RMW), quality (CMA) and momentum (UMD) factors. Each of these factors is formed as the return spread on long-short portfolios of firms sorted by firm characteristics. Data are obtained from Ken French's website.

Our cross-sectional analysis should only be viewed as suggestive evidence as we do not perform a full set of empirical tests, nor do we address the caveats highlighted by Lewellen et al. (2010). In addition, several other cross-sectional risk factors have been proposed and could be used in an extended analysis. In the interest of limiting the multiple hypothesis testing problem, we do not explore other factors here.[41]

The first step in our analysis extracts the daily innovation $\varepsilon_{\mu t+1}$ in the persistent cash flow component μ_{dt+1} from equation (4). The second step performs a set of univariate regressions

[41] See Barillas and Shanken (2018) for a discussion of ways to compute model probabilities when the models are driven by a limited subset of risk factors.

of returns on the individual risk factors (F_{t+1}) on an intercept and the innovation $\varepsilon_{\mu t+1}$:

$$F_{t+1} = \alpha + \beta_{\mu_{dt}} \varepsilon_{\mu t+1} + u_{t+1}. \tag{C.1}$$

Results from these regressions are reported in Table C.1. The estimated value of α shows the unconditional mean risk premium of the individual factors. This ranges from highs of 7.4% and 6.3% for the momentum and market factors, respectively, to a low of 1.8% for the size factor (SMB). Our main interest is of course in the estimate of $\beta_{\mu_{dt}}$ which reflects the sensitivity of the daily factor returns with respect to daily shocks to the persistent dividend growth component. Dividend shocks, $\varepsilon_{\mu t+1}$, are positively and significantly correlated with market returns (t-stat of 2.10). Days with above-normal stock market returns are thus associated with days with positive shocks to persistent cash flow growth. Similarly, we find a positive correlation between dividend shocks and returns on the size (SMB) factor, although this is only significant at the 10% level (t-stat of 1.74). Returns on small firms thus tend to be higher on days with positive shocks to cash flow growth and small firms are more sensitive to changes in aggregate cash flows than large firms. The negative and highly significant coefficient (t-stat of -4.65) on the profitability factor (RMW) suggests that returns on firms with weak profitability tend to have higher (relative) returns than firms with robust profitability on days where dividend growth prospects improve.

Overall, these results show that our new daily cash flow growth measure can help explain the daily returns not only of the aggregate market, but also of at least some of the risk factors. In particular, our findings suggest that improved cash flow prospects (positive shocks to $\Delta\mu_{dt+1}$) disproportionately benefit small firms and firms with weak profitability.

C.2 Dividend payments versus announced dividends

Our results up to this point show that movements in aggregate stock returns and market volatility are related to dividend news on the announcement date. This relation plausibly reflects how investors re-assess equity prices following cash flow news. We can test this hypothesis by exploiting the fact that we have data on both the date of the dividend announcement and the date where a dividend is paid out, with the payment date typically occurring several days after the announcement date. If the news effect hypothesis is correct, we would expect to find a substantially smaller impact of dividend growth on stock

returns on the payment date as compared to the return effect on the announcement date.

To see if this is the case, we estimate daily (contemporaneous) return regressions with the various dividend components as regressors, but use the dividend payment dates as opposed to the dividend announcement dates in extracting $\Delta\mu_{dt}$. The results, presented in columns 4-6 of Table C.2, show that the coefficient on $\Delta\mu_{dt}$ drops from 2.78 to 0.55, with the t-statistic dropping from 4.21 to 0.96. This is consistent with the cash flow news effect being what matters to movements in aggregate stock market prices, rather than any liquidity effects associated with payment of dividends.

The final column of Table C.2 shows estimates based on the daily dividend growth measure extracted from CRSP, computed as a daily year-on-year growth rate series. Once again, this measure, which uses information on dividend payments as opposed to announced dividends, has no significant explanatory power over aggregate stock returns.

Table C.3 reports results on the predictability of dividend growth for both announced and paid dividends at the lower frequencies (monthly, quarterly, and annual). Results are stronger when we use announced dividends rather than paid dividends.

C.3 Economic value of return predictability

Consider an investor who at time t allocates $\omega_t W_t$ of total wealth to stocks and the remainder, $(1-\omega_t)W_t$ to a risk-free asset, where $W_t = 1$ is the initial wealth while ω_t is the share allocated to stocks. Furthermore, assume that once the investor makes her allocation decision at time t, she waits until time $t+2$ to implement her investment, maintaining this position without rebalancing until time $t+h$ ($h \geq 2$). Skipping one day before trading ensures that results are not affected by slippage.

From Campbell and Viceira (2002) it follows that if the cumulative log excess return between time $t+2$ and $t+h$, $r_{t+2:t+h}$, is conditionally normally distributed with mean $\hat{r}_{t+2:t+h|t}$ and variance $\hat{\sigma}^2_{t+2:t+h|t}$, the optimal weight ($\hat{\omega}_t$) for an investor with CRRA utility and a coefficient of relative risk aversion A can be approximated by

$$\hat{\omega}_t = \left(\frac{1}{A}\right)\left(\frac{\hat{r}_{t+2:t+h|t} + \hat{\sigma}^2_{t+2:t+h|t}/2}{\hat{\sigma}^2_{t+2:t+h|t}}\right). \tag{C.2}$$

We assume that $A = 5$, which is a standard choice from the empirical finance literature.

To measure the economic value of return predictability from $\Delta\mu_{dt}$, consider the following prediction model for the mean of cumulative returns,

$$r_{\tau+2:\tau+h} = \beta_0 + \beta_1 \Delta\mu_{d\tau} + \varepsilon_{\tau+2:\tau+h}, \quad \tau = 1, .., t-h \tag{C.3}$$

and the volatility of cumulative returns:

$$Vol_{\tau+1} = \gamma_0 + \gamma_1 \mu_{d\tau} + \gamma_2 Vol^d_\tau + \gamma_3 Vol^w_\tau + \gamma_4 Vol^m_\tau + \epsilon_{\tau+1}, \tag{C.4}$$

where $Vol_{\tau+1}$ denotes the S&P500 realized volatility at time $\tau+1$ and Vol^d_τ, Vol^w_τ, and Vol^m_τ are the lagged daily, weekly, and monthly volatility averages, respectively, as defined in Corsi (2009). Note that because of the high persistence in $Vol_{\tau+1}$, we use the level of μ_{dt}, rather than its change, as a predictor in the volatility equation. Moreover, (C.4) is easily iterated on to obtain forecasts of $Vol_{\tau+h}$.

Our realized volatility data start in January 2000 and we use a three-year initial estimation window so the forecasting/rebalancing dates begin on January 3, 2003 and include every Friday until one week before the end of our sample, December 16, 2016 which yields a total of $P = 730$ distinct dates.[42] To reduce portfolio turnover, our analysis assumes weekly trades ($h = 5$). Specifically, on any given Friday we first predict the mean and variance of cumulative returns, $r_{t+2:t+h}$, using the model (C.3)-(C.4) that includes our persistent dividend growth measure as a predictor. Next, using these forecasts, we compute the optimal allocation to stocks (C.2) which is held constant between time $t+2$ and $t+h$. As a benchmark, we also compute the portfolio allocation to stocks implied by cumulative return and volatility forecasts for an investor who disregards information on the persistent dividend growth component when forming her forecasts, thus setting $\beta_1 = \gamma_1 = 0$ in (C.3) and (C.4). Finally, following Campbell and Viceira (2002), we approximate the investor's wealth at time $t+h$ by

$$\widehat{\ln(W_{t+h})} = (h-1)\, r_{ft} + \widehat{w}_t r_{t+2:t+h} + \frac{1}{2}\widehat{w}_t\,(1-\widehat{w}_t)\,\widehat{\sigma}^2_{t+2:t+h|t},$$

where r_{ft} denotes the risk-free rate.

[42] For those weeks in which the markets are not open on a Friday, we use as our rebalancing date the last day of that week when the markets are open.

The annualized mean, volatility, and Sharpe ratio based on the return predictions that use information on $\Delta\mu_{dt}$ are 4.98%, 8.32%, and 0.60, respectively. The corresponding numbers for the predictions that do not use such information are 2.26 %, 6.75%, and 0.33. Finally, the annualized certainty equivalent return differential between these two models equals 2.32% which is economically sizeable and so suggests that there are considerable economic gains from using information on the persistent dividend growth component to predict stock returns.

C.4 Stock market volatility and dividend news

Section 6 shows that dividend growth dynamics affect not only the mean of stock returns but also impact the volatility and jump probability of the return process with positive news about the persistent dividend growth component reducing stock market volatility.

To explore the robustness of this finding, we investigate the relation between daily stock market volatility and cash flow news using two different measures of market volatility. First, we use the VIX obtained from options prices which reflects market expectations of short-run (30-day) volatility in stock prices. Second, we use a realized variance (RV) measure of daily stock market volatility based on intra-day movements in the price on the S&P500 index sampled every 5 minutes.[43] Data on the VIX are available starting in 1990, while data on realized volatility begin in 2000.

We first consider the contemporaneous relation between daily stock market volatility and news about the persistent dividend growth component. Panel A in Table C.4 shows that there is a significant and negative correlation between movements in the persistent dividend growth component and stock market volatility measured by either the VIX or the RV, consistent with positive news about long-run dividend growth reducing stock market volatility.

Next, we consider whether dividend growth news helps predict future stock market volatility. Following Paye (2012), we use the level of volatility in our regressions, but account for the high persistence in this variable by including either a single lag or an average of lagged volatility as proposed in the cascade model of Corsi (2009). Specifically,

[43] Our data come from the Oxford-Man Institute of Quantitative Finance, http://realized.oxford-man.ox.ac.uk/data/download.

we use the following two regression specifications for the volatility on day t, VOL_t:

$$Vol_{t+1} = \alpha + \beta_1 Vol_t + \beta_2 \mu_{dt} + \varepsilon_{t+1}, \tag{C.5}$$

$$Vol_{t+1} = \alpha + \beta \mu_{dt} + \beta_d RV_t^d + \beta_w RV_t^w + \beta_m RV_t^m + \varepsilon_{t+1}, \tag{C.6}$$

where RV^d, RV^w and RV^m are daily, weekly, and monthly volatility averages, respectively.

Panel B in Table C.4 shows the results from these regressions using the VIX (left column) or the realized volatility (right column). Regardless of whether we use the specification in (C.5) or (C.6), we find strong evidence of persistence in the volatility process.

Turning to the predictive content of the persistent dividend component, μ_{dt}, over stock market volatility, for both specifications in Panel B we find that the coefficient on μ_{dt} is negative and highly statistically significant with t-statistics of -4.15 and -9.85, respectively. While these t-statistics drop to -2.07 and -2.16 in the cascade model, they remain significant. This confirms that positive news about persistent dividend growth leads to lower stock market volatility, while negative news tends to increase stock market volatility.[44]

C.5 Alternative econometric specifications

To explore how sensitive our estimates of the dividend growth model are to the specifications of the jump and volatility processes, we modify our baseline specification to test for the importance of both state-dependence in the jump probabilities, as in Johannes et al. (1999), and correlated jumps in the mean and volatility, as in Eraker et al. (2003b).

Starting with the state-dependent jump model, we replace (8) with the following specification[45]

$$\Pr\left(J_{dt+1} = 1\right) = \Phi\left(\lambda_1 + \lambda_2 N_{t+1} + \lambda_3 J_{dt} + \lambda_4 \left|\Delta d_t\right|\right). \tag{C.7}$$

Compared to our baseline specification in (8), all posteriors of the coefficients are essentially unchanged and the estimates of λ_3 and λ_4 are not significantly different from zero.

Moving on to correlated jumps, we change our baseline model to match Eraker et al.

[44] We also analyze whether the stochastic volatility and jump components extracted from the jump model have any contemporaneous or predictive effect on the aggregate volatility but find that the effects are negligible and not statistically significant.

[45] We modify our prior on the parameters governing the timing of the jumps as follows: $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)' \sim \mathcal{N}\left(\underline{\boldsymbol{\mu}}_{\boldsymbol{\lambda}}, \underline{\boldsymbol{V}}_{\boldsymbol{\lambda}}\right)$, where $\underline{\boldsymbol{\mu}}_{\boldsymbol{\lambda}} = (0,0,0,0)'$ and $\underline{\boldsymbol{V}}_{\boldsymbol{\lambda}} = diag\,(10,10,10,10)$.

(2003b)'s most general specification with correlated jumps in the mean and volatility. In particular, we modify our model for the log-variance of dividend growth in (10) as follows:

$$h_{dt+1} = \mu_h + \phi_h \left(h_{dt} - \mu_h \right) + J_{dt+1} \xi^h_{dt+1} + \sigma_h \varepsilon_{ht+1}, \tag{C.8}$$

where $\xi^h_{dt+1} \big| \, \xi_{dt+1} \sim \mathcal{N}\left(\rho_J \xi_{dt+1}, \sigma^2_{\xi^h} \right)$.[46] Testing this specification on our data, again we find that none of the estimated coefficients in our baseline specification change materially. We also find that the estimate of the correlation coefficient ρ_J is not statistically different from zero. Most importantly, when we compare the μ_{dt} estimates extracted from either the Johannes et al. (1999) or the Eraker et al. (2003b) model to our baseline specification, the three series are basically indistinguishable with a correlation of 0.9999. Plots of the three μ_{dt} series and a table with parameter estimates are shown in Figure C.1 and Table C.5. We find that none of the estimated coefficients in our baseline specification change materially. Most importantly, when we compare the μ_{dt} estimates extracted from either the Johannes et al. (1999) or the Eraker et al. (2003b) model to our baseline specification, the three estimates of the persistent dividend component are basically indistinguishable with a correlation of 0.9999.[47]

C.6 Estimates for industry portfolios

Table C.6 reports estimates of the dividend growth model for the five Fama-French industries, while Figure C.2 plots their persistent components, μ_{dt}. The basic features of the dividend process remain the same across very different industries, including estimates of ϕ_μ close to unity, indicating a highly persistent component in dividend growth, similar jump sizes, and negative dependence between the jump probability and the number of firms announcing dividends on a given day.

[46] We specify the following priors on the additional model parameters: $\rho_J \sim \mathcal{N}\left(0, \sigma^2_{\rho_J}\right)$ and $\sigma^2_{\xi^h} \sim \mathcal{IG}\left(\underline{\nu}^h_\xi, \underline{S}^h_\xi\right)$, and set $\sigma^2_{\rho_J} = 1$, $\underline{\nu}^2_\xi = 2$ and $\underline{S}^h_\xi = 2^2$.

[47] Our analysis is not intended to exhaust all jump and volatility specifications in what is now a very large literature. For example, Bollerslev and Todorov (2011) introduce a flexible non-parametric procedure that can accommodate complex dynamic tail dependencies and stochastic volatility. Working with closing bid and ask quotes for S&P 500 options, Bollerslev et al. (2015) decompose the variance risk premium into diffusive and jump risk components and link predictability of stock market returns to jump tail risk.

PANEL A: Cross-sectional analysis						
	MRP	**SMB**	**HML**	**RMW**	**CMA**	**UMD**
α (annualized)	6.32%	1.84%	5.01%	3.51%	4.27%	7.40%
t-stat	2.51	1.39	3.95	3.71	4.81	4.16
$\beta_{\mu_{dt}}$	1.09	0.43	0.26	-0.84	-0.17	-0.21
t-stat	2.10	1.74	0.10	-4.65	-1.09	-0.58

Table C.1: Return spreads and shocks to the persistent dividend growth component. This table reports the estimated intercept and slope coefficients from regressions of daily returns on spread portfolios tracking a variety of risk factors (MRP, SMB, HML, RMW, CMA and UMD) on a constant and daily shocks to the persistent component in the dividend growth process extracted from our dividend growth rate model. We also report t-statistics computed using Newey-West standard errors. The sample period is 1973-2016.

	Dividend announcement days			**Dividend payment days**			
Δd_t	.00			-.00			
	[1.34]			[-0.37]			
$\Delta\mu_{dt}^{NJ}$		.00			-.00		
		[1.30]			[-0.43]		
$\Delta\mu_{dt}$			2.78***			0.55	
			[4.21]			[0.96]	
$\xi_{dt}J_{dt}$			.00			-.00	
			[1.60]			[-0.00]	
$h_{dt}/2$			-.00			.00	
			[-0.76]			[0.37]	
Δd_t^{CRSP}							.00
							[0.45]
R^2	0.01%	0.01%	0.26%	0.00%	0.00%	0.01%	0.00%
Observations	20,966	20,966	20,966	20,393	20,393	20,393	10,900

Table C.2: Daily regressions of stock returns on dividend news. This table reports estimates from regressions of daily stock market returns on 1) daily growth in aggregate dividends, Δd_t; 2) changes in the persistent dividend growth component, $\Delta\mu_{dt}^{NJ}$, extracted from a dividend growth model without jumps and stochastic volatility; the following components extracted from the dividend growth model that accounts for jumps and stochastic volatility: 3) changes in the persistent component, $\Delta\mu_{dt}$; (iv) jumps, $\xi_{dt}J_{dt}$; (v) stochastic volatility, $h_{dt}/2$. In each case, the dependent variable is the two-day cumulative log stock market return on days t and $t+1$, $r_{t:t+1}$. Columns 1-3 consider stock returns on the days with the dividend news announcements, while columns 4-7 relate stock returns to dividend news on the days where the dividend payments are actually made. The final column reports results from regressing returns on a daily dividend growth series, Δd_t^{CRSP}, computed from the CRSP index. Square brackets report t-statistics using Newey-West standard errors. The sample period is 1927-2016.

$$\Delta d_{t+1}^{CRSP} = \alpha + \rho_i \sum_{i=1}^{3} \Delta d_{t+1-i}^{CRSP} + \beta \mu_{dt}^{i} + \gamma dp_t^{CRSP} + \varepsilon_{t+1}$$

	Monthly			**Quarterly**			**Annual**		
	Announced		*Paid*	*Announced*		*Paid*	*Announced*		*Paid*
μ_{daily}	.17***			.36***			2.50***		
	[5.77]			[4.58]			[3.95]		
$\mu_{monthly}$		.14***	.12***						
		[6.51]	[4.87]						
$\mu_{quarterly}$					.22***	.17***			
					[3.19]	[3.18]			
μ_{yearly}								2.64**	1.72
								[2.58]	[0.59]
dp_t	-.00**	-.00	-.00	-.01	.00	.00	.03	.07	.06
	[-2.49]	[-0.97]	[-0.63]	[-1.29]	[0.20]	[0.21]	[0.50]	[1.12]	[0.87]
Δd_t^{CRSP}	-.05	-.05	-.03	.16**	.22**	.19**	-.64***	-.52***	-.42***
	[-0.97]	[-0.94]	[-0.51]	[2.20]	[2.59]	[2.11]	[-4.58]	[-3.07]	[-2.32]
Δd_{t-1}^{CRSP}	-.07*	-.07**	-.05	.07	.10	.08	-.53***	-.36***	-.24*
	[-1.96]	[-2.01]	[-1.25]	[1.39]	[1.65]	[1.47]	[-3.87]	[-3.01]	[-1.70]
Δd_{t-2}^{CRSP}	.20***	.20***	.22***	.03	.06	.05	-.12	-.08	-.01
	[3.68]	[3.59]	[3.70]	[0.45]	[1.05]	[0.83]	[-0.78]	[-0.52]	[-0.03]
R^2	18.29%	17.80%	15.97%	26.50%	21.76%	23.16%	27.45%	22.14%	16.07%
Vuong test		[-2.70***]			[0.62]			[1.28]	
Observations	525	525	525	169	169	169	40	40	40

Table C.3: Dividend growth regressions. This table shows results from predictive regression of the conventional dividend growth measure extracted from CRSP data, Δd_{t+1}^{CRSP} on the persistent component μ_{dt}^{i} estimated from our daily dividend growth model (at the various frequencies) and the log dividend price ratio, dp_t, at quarterly and annual frequencies. The third column of each block (e.g., monthly, quarterly, annual) shows the results using the actual paid out dividends. We report the Vuong (1989) t-stat comparing the models with announced and paid dividends, which suggest the models differ at the monthly frequency but not at the quarterly/annual ones. Square brackets report t-statistics computed using Newey-West standard errors with three lags. Sample: 1973-2016.

Panel A: Contemporaneous regressions		
	VIX	*SP500 Realized Vol*
μ_{dt}	-20.86***	-14.85***
	[-20.34]	[-14.69]
$\xi_{dt} J_{dt}$	-.08	.00
	[-1.19]	[0.02]
$h_{dt}/2$	2.53***	.76
	[5.07]	[1.40]
R^2	30.57%	21.78%
Observations	6,527	3,977

Panel B: Predictive regressions		
	VIX	*SP500 Realized Vol*
AR(1)	.97***	.74***
	[164.65]	[31.21]
μ_{dt}	-0.48***	-3.80***
	[-4.15]	[-9.85]
R^2	96.28%	64.68%
Observations	6,526	3,953
Corsi (2009) model		
μ_{dt}	-.26**	-.86**
	[-2.07]	[-2.16]
RV_t^d	.85***	.35***
	[28.38]	[7.85]
RV_t^w	.11***	.41***
	[2.75]	[5.75]
RV_t^m	.03	.17***
	[1.64]	[3.21]
R^2	96.38%	70.98%
Observations	6,443	3,542

Table C.4: Relation between the persistent dividend component, VIX, and realized stock market volatility. Panel A in this table reports estimates from daily regressions of the VIX (left column) or the realized volatility based on the S&P500 index (right column) on the contemporaneous value of the persistent dividend growth component μ_{dt} extracted from our components model. Panel B reports similar results, relating the VIX or realized volatility to the lagged value of μ_{dt} as well as a single lag of the dependent variable or multiple lags based on the Corsi (2009) model. The dependent variables in Panel A are standardized. Square brackets show t-statistics using Newey-West standard errors computed using three lags.

Panel A: Baseline Model

	Baseline			State-dependent Jumps			Correlated Jumps		
	Mean	*Std*	*90% Credible Set*	*Mean*	*Std*	*90% Credible Set*	*Mean*	*Std*	*90% Credible Set*
μ_d	0.0802	0.0116	[0.0622,0.0973]	0.0802	0.0121	[0.0615,0.0984]	0.0793	0.0117	[0.0604,0.0958]
ϕ_μ	0.9983	0.0006	[0.9972,0.9993]	0.9983	0.0006	[0.9972,0.9992]	0.9982	0.0006	[0.9971,0.9992]
σ_μ	0.0019	0.0001	[0.0016,0.0021]	0.0019	0.0001	[0.0016,0.0021]	0.0019	0.0001	[0.0016,0.0021]
μ_h	-5.3373	0.0449	[-5.4146,-5.2615]	-5.3362	0.0464	[-5.4123,-5.2590]	-5.3332	0.0463	[-5.4041,-5.2554]
ϕ_h	0.8332	0.0078	[0.8203,0.8462]	0.8324	0.0080	[0.8191,0.8454]	0.8314	0.0079	[0.8185,0.8444]
σ_h	0.7517	0.0279	[0.7056,0.7972]	0.7579	0.0286	[0.7103,0.8056]	0.7476	0.0284	[0.7024,0.7940]
σ_ξ	2.7613	0.0398	[2.6948,2.8291]	2.7629	0.0416	[2.6943,2.8323]	2.7642	0.0422	[2.6957,2.8348]
σ_{ξ^h}							0.8544	0.1286	[0.6658,1.0846]
λ_1	-1.3541	0.0446	[-1.4281,-1.2811]	-1.3243	0.0513	[-1.4087,-1.2404]	-1.3521	0.0453	[-1.4240,-1.2745]
λ_2	-0.0244	0.0022	[-0.0281,-0.0208]	-0.0244	0.0021	[-0.0281,-0.0210]	-0.0245	0.0021	[-0.0280,-0.0212]
λ_3				0.0300	0.1475	[-0.2138,0.2705]			
λ_4				-0.2799	0.2636	[-0.8075,0.0482]			
ρ							0.0008	0.0408	[-0.0679,0.0660]

Panel B: Market Weight Model

	Market weight			State-dependent Jumps			Correlated Jumps		
	Mean	*Std*	*90% Credible Set*	*Mean*	*Std*	*90% Credible Set*	*Mean*	*Std*	*90% Credible Set*
μ_d	0.1085	0.0172	[0.0837,0.1310]	0.1093	0.0160	[0.0842,0.1313]	0.1070	0.0184	[0.0825,0.1301]
ϕ_μ	0.9985	0.0006	[0.9975,0.9994]	0.9985	0.0006	[0.9974,0.9994]	0.9985	0.0006	[0.9975,0.9995]
σ_μ	0.0021	0.0002	[0.0018,0.0024]	0.0021	0.0002	[0.0018,0.0024]	0.0021	0.0002	[0.0018,0.0023]
μ_h	-5.0458	0.0443	[-5.1189,-4.9763]	-5.0450	0.0432	[-5.1155,-4.9752]	-5.0420	0.0439	[-5.1126,-4.9682]
ϕ_h	0.8423	0.0080	[0.8291,0.8560]	0.8423	0.0082	[0.8288,0.8558]	0.8400	0.0081	[0.8267,0.8529]
σ_h	0.6782	0.0264	[0.6338,0.7222]	0.6787	0.0279	[0.6333,0.7246]	0.6729	0.0269	[0.6282,0.7175]
σ_ξ	2.7872	0.0403	[2.7230,2.8545]	2.7888	0.0420	[2.7213,2.8601]	2.7909	0.0412	[2.7255,2.8572]
σ_{ξ^h}							0.8867	0.1189	[0.6947,1.0932]
λ_1	-1.4496	0.0448	[-1.5251,-1.3771]	-1.4402	0.0496	[-1.5236,-1.3606]	-1.4540	0.0421	[-1.5260,-1.3873]
λ_2	-0.0206	0.0021	[-0.0240,-0.0173]	-0.0205	0.0021	[-0.0241,-0.0171]	-0.0205	0.0019	[-0.0236,-0.0172]
λ_3				0.0624	0.1464	[-0.1788,0.2995]			
λ_4				-0.0890	0.1453	[-0.3464,0.1303]			
ρ							-0.0004	0.0395	[-0.0668,0.0632]

Table C.5: Parameter estimates for dividend growth models with different specifications of the jump component. This table shows parameter estimates for the dividend growth models using the jump specifications in Johannes et al. (1999) and Eraker et al. (2003b). The first model replaces our baseline jump probability specification with the following state-dependent specification

$$\Pr\left(J_{dt+1}=1\right)=\Phi\left(\lambda_1+\lambda_2 N_{t+1}+\lambda_3 J_{dt}+\lambda_4\left|\Delta d_t\right|\right).$$

The second model accounts for correlated jumps in the mean and volatility and modifies our original baseline specification to

$$h_{dt+1}=\mu_h+\phi_h\left(h_{dt}-\mu_h\right)+J_{dt+1}\xi^h_{dt+1}+\sigma_h\varepsilon_{ht+1},$$

where

$$\xi^h_{dt+1}\Big|\,\xi_{dt+1}\sim\mathcal{N}\left(\rho_J\xi_{dt+1},\sigma^2_{\xi^h}\right)$$

Columns report the posterior mean, standard deviation and 90% credible sets for the parameter estimates.

Parameter estimates															
	Consumers			**Manufacturing**			**HiTech**			**Healthcare**			**Others**		
	Mean	*Std*	*90% Credible Set*	*Mean*	*Std*	*90% Credible Set*	*Mean*	*Std*	*90% Credible Set*	*Mean*	*Std*	*90% Credible Set*	*Mean*	*Std*	*90% Credible Set*
μ_d	0.078	0.014	[0.058,0.099]	0.066	0.021	[0.033,0.095]	0.069	0.027	[0.032,0.103]	0.089	0.017	[0.065,0.110]	0.098	0.014	[0.077,0.116]
ϕ_μ	0.998	0.001	[0.997,0.999]	0.999	0.000	[0.998,0.999]	0.999	0.001	[0.998,0.999]	0.997	0.002	[0.994,0.999]	0.998	0.001	[0.997,0.999]
σ_μ	0.002	0.000	[0.002,0.002]	0.002	0.000	[0.001,0.002]	0.002	0.000	[0.002,0.003]	0.003	0.000	[0.002,0.003]	0.002	0.000	[0.002,0.003]
μ_h	-4.980	0.046	[-5.051,-4.903]	-5.411	0.050	[-5.493,-5.327]	-4.666	0.058	[-4.762,-4.572]	-4.620	0.073	[-4.738,-4.499]	-4.682	0.051	[-4.766,-4.596]
ϕ_h	0.839	0.008	[0.825,0.852]	0.811	0.008	[0.798,0.824]	0.860	0.008	[0.846,0.873]	0.933	0.010	[0.917,0.949]	0.833	0.008	0.820,0.845]
σ_h	0.688	0.028	[0.644,0.735]	0.878	0.032	[0.825,0.928]	0.685	0.036	[0.626,0.745]	0.308	0.037	[0.248,0.368]	0.852	0.029	[0.804,0.899]
σ_ξ	2.756	0.041	[2.691,2.823]	2.806	0.043	[2.736,2.879]	2.799	0.042	[2.730,2.869]	2.831	0.043	[2.760,2.902]	2.786	0.041	[2.718,2.854]
λ_1	-1.209	0.048	[-1.289,-1.130]	-1.472	0.053	[-1.559,-1.388]	-1.510	0.057	[-1.609,-1.420]	-1.387	0.066	[-1.497,-1.277]	-1.341	0.048	[-1.420,-1.261]
λ_2	-0.108	0.010	[-0.125,-0.092]	-0.064	0.008	[-0.078,-0.051]	-0.092	0.021	[-0.126,-0.058]	-0.122	0.038	[-0.186,-0.059]	-0.058	0.006	[-0.069,-0.049]

Table C.6: Parameter estimates for the dividend growth rate model estimated separately for the five Fama-French industry portfolios. This table shows parameter estimates for a model fitted to the daily dividend growth series of the five Fama-French industry portfolios (consumers, manufacturing, high-tech, health care, and others). The components model underlying these estimates takes the following form:

$$\Delta d_{t+1} = \mu_{dt+1} + \xi_{dt+1} J_{dt+1} + \varepsilon_{dt+1},$$

$$\mu_{dt+1} = \mu_d + \phi_\mu \left(\mu_{dt} - \mu_d\right) + \sigma_\mu \varepsilon_{\mu t+1},$$

$$\varepsilon_{dt+1} \sim \mathcal{N}(0, e^{h_{dt+1}}),$$

$$h_{dt+1} = \mu_h + \phi_h \left(h_{dt} - \mu_h\right) + \sigma_h \varepsilon_{ht+1},$$

$$\Pr\left(J_{dt+1} = 1\right) = \Phi\left(\lambda_1 + \lambda_2 N_{dt+1}\right),$$

$$\xi_{dt+1} \sim \mathcal{N}\left(0, \sigma_\xi^2\right).$$

Here μ_{dt+1} captures the mean of the smooth component of the underlying dividend process, $J_{dt+1} \in \{0, 1\}$ is a jump indicator that equals unity in case of a jump in dividends and otherwise is zero, ξ_{dt+1} measures the jump size, ε_{dt+1} is a temporary cash flow shock, $\varepsilon_{\mu t+1} \sim \mathcal{N}(0, 1)$ is assumed to be uncorrelated at all times with the innovation in the temporary dividend growth component, ε_{dt+1}, and $|\phi_\mu| < 1$. h_{dt+1} denotes the log-variance of ε_{dt+1}, and $\varepsilon_{ht+1} \sim \mathcal{N}(0, 1)$ is uncorrelated at all times with both ε_{dt+1} and $\varepsilon_{\mu t+1}$. N_{dt+1} denotes the number of firms announcing dividends on day $t + 1$, while Φ stands for the CDF of a standard Normal distribution and $\xi_{dt+1} \sim \mathcal{N}\left(0, \sigma_\xi^2\right)$ captures the magnitude of the jumps. The columns report the posterior mean, standard deviation and 90% credible sets for the parameter estimates.

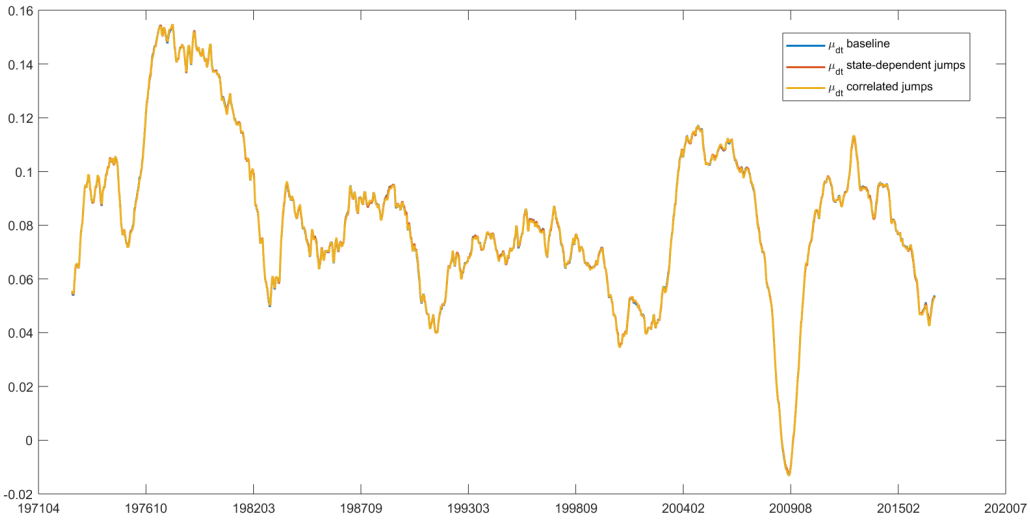

Figure C.1: Estimates of the persistent dividend growth component, μ_{dt}, based on different specifications for the jump component. The figure plots the estimated persistent dividend growth component, μ_{dt}, for (i) our baseline specification; (ii) a model that allows for state-dependent jumps; and (iii) a model that allows for correlated jumps. All estimates use daily dividend announcement data over the period 1973-2016.

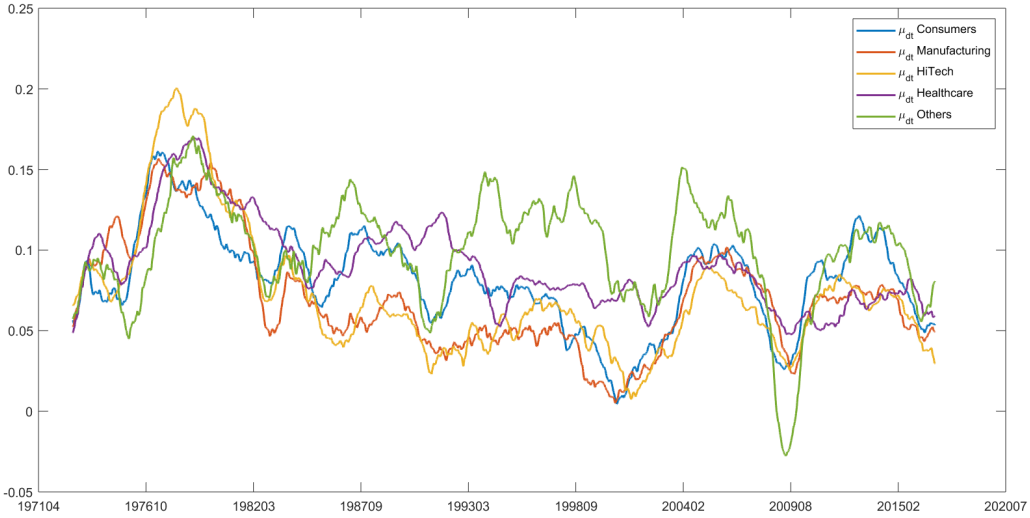

Figure C.2: Comparison of μ_{dt} estimates across different industries. This figure plots time series of μ_{dt} estimated separately for the five Fama-French industry portfolios.