Journal of Banking and Finance 162 (2024) 107127

Contents lists available at ScienceDirect

Journal of Banking and Finance

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

# Discount rates and cash flows: A local projection approach ☆

Matthijs Lof[a,*], Henri Nyberg[b]

[a] Aalto University School of Business, P.O.Box 21220, FI-00076 Aalto, Finland
[b] University of Turku, Department of Mathematics and Statistics, FI-20014, University of Turku, Finland

ARTICLE INFO

*JEL classification:*
C32
G12

*Keywords:*
Dividend yield
Decomposition
Dividend growth
Local projections

ABSTRACT

We develop flexible local projections to quantify the relative contributions of expected discount rates and cash flows to the variation of dividend yields. Local projections enable the incorporation of large information sets, the use of monthly data along with annual data, and the consideration of time variation in the dividend yield decomposition. By expanding the set of state variables and allowing for time-varying parameters, our results show that the variation of expected discount rates remains the primary contributor to market volatility, whereas the contribution of expected cash flows is considerably smaller.

## 1. Introduction

The value of a stock should equal expected discounted cash flows. Understanding the relative contribution of expected discount rates (returns) and cash flows (dividends) to the volatility of equity markets is one of the central topics in asset pricing research. A voluminous literature demonstrates that expected dividends contribute only marginally, if at all, to the volatility of prices (see, e.g., the early evidence in Shiller, 1981; LeRoy and Porter, 1981; Campbell and Shiller, 1988b; and Cochrane, 1992, 2008). In response to these findings, the focus of asset pricing research in recent decades has primarily been on the analysis of discount rate variation (see, e.g., Cochrane, 2011, 2017).

To analyze the discount rate vs. cash flow conundrum, a typical starting point is the Campbell and Shiller (1988b) log-linear present value model, which decomposes the dividend yield into expected discount rates and expected cash flow growth. The empirical implementation in Campbell and Shiller (1988b), Cochrane (2008) and many other studies utilizes a vector autoregressive (VAR) representation describing the dynamics of (log) market returns, dividend yields, dividend growth, and possibly additional variables. The estimated VAR coefficients are subsequently used to infer long-run expectations of discount rates and cash flow growth. Specifically, Cochrane (2008) uses the lagged dividend yield as the only state variable to predict future returns and dividend growth rates and finds, since the dividend yield is a poor predictor of future dividend growth, a negligible contribution of expected dividend growth to price volatility.

In this study, we also build upon the log-linear present value model but introduce an alternative methodology to empirically quantify the relative contributions of expected discount rates and cash flows in a more general environment than in the past approaches. Our approach is in spirit similar to Campbell and Shiller (1988b) and Cochrane (2008) as we use regression-based techniques to infer cash flow and discount rate expectations. However, instead of inferring implied long-run expectations from a VAR with one state variable, we obtain the required predictions for the discounted (cumulative) expected returns and dividend growth rates using (forecast) horizon-specific single-equation re-

☆ We thank Geert Bekaert (the editor), three anonymous referees, Zhi Da, Jari-Pekka Heinonen, Òscar Jordà, Samu Kärkkäinen, Hugues Langlois, Riccardo Sabbatucci, Matti Suominen, and conference and seminar participants at the American Economic Association 2020 (San Diego), EEA-ESEM 2019 (Manchester), EUROFIDAI Paris December Finance Meeting 2019, European Finance Association 2020 (Helsinki), Finnish Economic Association 2019 (Oulu), Nordic Econometric Meeting 2019 (Stockholm), Society for Nonlinear Dynamics and Econometrics 2020 (Zagreb), Aalto University, University of Helsinki, University of Nottingham, and the University of Turku for useful comments and discussions, as well as Emil Lehti for excellent research assistance. We gratefully acknowledge the Finnish Foundation for Share Promotion and (Nyberg) the Emil Aaltonen Foundation (grant 180287), the Academy of Finland (grant 321968), OP Group Research Foundation and the Foundation for Economic Education (Liikesivistysrahasto, grant 220246) for financial support. Nyberg also acknowledges the resources of the Tampere University (Economics) when working with the first version of this manuscript in 2019.

* Corresponding author.
*E-mail addresses:* matthijs.lof@aalto.fi (M. Lof), henri.nyberg@utu.fi (H. Nyberg).

https://doi.org/10.1016/j.jbankfin.2024.107127
Received 16 June 2021; Accepted 1 March 2024
Available online 7 March 2024

gressions, which we refer to as 'local projections' (LPs), following Jordà (2005).

Our local projection approach has several advantages over the VAR approach. First, LPs are less restrictive than VARs, thereby reducing model misspecification concerns. Second, LPs enable the incorporation of potentially large sets of economic and financial state variables beyond the traditional dividend yield. Third, LPs can be estimated with higher-frequency data, such as monthly data, while the seasonality of dividend data has restricted prior studies using the VAR-based approaches to rely solely on annual data (see, e.g., the survey by Koijen and Van Nieuwerburgh, 2011). Due to the enlarged sample sizes resulting from the use of monthly data, we can estimate the local projections recursively, enabling us to uncover time variation in the predictability of dividends and returns and ultimately in the dividend yield decomposition. We briefly outline these advantages below.

Jordà (2005) proposes local projections as an alternative to VARs for computing impulse response functions in structural macroeconometrics.[1] We apply local projections in a different context: to infer long-run expectations of discount rates and cash flow growth in order to obtain a dividend yield decomposition. Instead of the conventional approach of extrapolating an estimated one-period VAR model over multiple periods, the idea of local projections is to construct predictions at each horizon of interest separately. As argued by Jordà (2005) and Schorfheide (2005), this approach is more robust to potential misspecification than the VAR approach, which is built upon the strong assumption that the underlying VAR representation is correctly specified.[2] In practice, the estimated VAR is likely to be misspecified, like any econometric model, providing at best an approximation to the true correct asset pricing process. As Jordà (2005) puts it, misspecification errors are 'compounded with the forecast horizon' with a VAR, whereas horizon-specific local projections are optimized to minimize misspecification error at each horizon separately, not requiring an exact specification of the true multivariate dynamic system.[3]

An example of a situation where extrapolating short-run predictions does not provide optimal long-run predictions is dividend smoothing. Short-run dividend smoothing adversely affects the predictability of dividends in the short run. As Chen et al. (2012) demonstrate, this lack of short-run predictability induces a negative bias to the VAR-implied contribution of cash flow news: they show by simulation that even if dividends are predictable in the long run, this predictability is for the most part not uncovered by a VAR in the presence of short-run dividend smoothing.[4] Since cumulative dividend growth rates over longer horizons (say, 10 or 15 years) are less affected by dividend smoothing policies, our local projections largely circumvent the concerns raised by Chen et al. (2012), by making direct (as opposed to VAR-implied) predictions of long-run dividend growth.

Local projections also allow for the inclusion of a potentially large set of state variables. We extend the information set beyond lagged dividend yields, returns and dividends, by considering a larger set of financial and macroeconomic predictors of dividends and returns that have been identified in the earlier literature.[5] Demonstrating the flexibility of local projections, we can choose different sets of state variables to model expected returns and expected dividend growth separately. Since the list of potential state variables is long, we also use LASSO (Least Absolute Shrinkage and Selection Operator) to select state variables. LASSO is a machine learning method popularized by Tibshirani (1996) that performs variable selection and parameter estimation simultaneously to guard against overfitting, potentially enhances the prediction accuracy, and facilitates interpretability of the resulting econometric specification. While recent return predictability studies (e.g., Rapach et al., 2010; Gu et al., 2020) have used LASSO and other related machine learning methods to predict market returns, this is, to the best of our knowledge, the first study applying such methods to assess the relative importance of discount rate and cash flow expectations to market volatility.

VAR-based volatility decompositions of the dividend yield are typically based on annual data due to pervasive seasonal patterns in monthly dividends (e.g., Koijen and Van Nieuwerburgh, 2011). On the contrary, the aforementioned studies on the predictability of returns and dividend growth often analyze monthly data. Local projections enable us to establish an approximate volatility decomposition that incorporates dividend growth also at monthly frequency, despite the seasonality of monthly dividends data, by including monthly updated observations of annualized (i.e., 12-month) cumulative dividend growth rates. The use of monthly data increases the number of observations considerably, which facilitates meaningful examination of possible time variation in the discount rate and cash flow contributions over time.

In our empirical analysis, we first apply local projections with the lagged dividend yield as a single state variable. Confirming the findings of Cochrane (2008, 2011), we find that dividend growth expectations conditional on the lagged dividend yield are nearly flat, such that the contribution of expected cash flow growth to the dividend yield is small compared to the contribution of expected discount rates. When including more state variables, we find evidence of increased time variation in cash flow growth expectations, consistent with the recent literature on dividend predictability. Whether this increased predictability of dividends translates into a larger component of cash flows to market volatility is nevertheless a different question. Cochrane (2008, 2011) argues that due to the links between returns, dividend growth, and dividend yields, any additional dividend predictability beyond the predictive power of the lagged dividend yield needs to be compensated by additional predictability of returns such that the dividend yield decomposition remains unaffected. Menzly et al. (2004), Lettau and Ludvigson (2005) and Golez (2014), however, argue that the positive correlation between expected cash flow growth and expected returns has a negative impact on the dividend yield's ability to predict expected cash flow growth, thereby underestimating the contribution of cash flow expectations to dividend yield volatility. Given these considerations, we

[1] See, e.g., Gorodnichenko and Lee (2020), Plagborg-Møller and Wolf (2021) and Li et al. (2022), and the references therein. Cochrane and Piazzesi (2002) provide an early example of impulse response functions constructed by direct (i.e. LP) regressions.

[2] The robustness of local projections in terms of potential misspecification, such as in the presence of potential breakpoints (see Chevillon, 2016), is supported by comparisons between 'direct' and 'iterative' multiperiod forecasting methods (see, e.g., Marcellino et al., 2006; and Chevillon, 2007).

[3] Although LPs are less restrictive than VARs, the LP approach is not assumption-free. Both approaches are based on the implicit assumption that agents know the correct data generating process and form model-consistent, or objective, expectations. The LP approach is therefore distinct from, e.g., De la O and Myers (2021), who use survey data to proxy for subjective cash flow growth and discount rate expectations. In contrast to objective expectation approaches, they find a larger share of market variation to be attributed to cash flow expectations. Moreover, our LP-based decomposition, like the VAR approach, relies on the Campbell-Shiller (1988b) log-linearization, which is known to generate significant approximation errors when the underlying process is highly nonlinear (e.g., Pohl et al., 2018).

[4] Consistent with the adverse impact of dividend smoothing on dividend growth predictability, recent studies identify a larger cash flow component relying on alternative measures of cash flows, including net payout (Larrain and Yogo, 2008), earnings (Chen et al., 2012), and direct cash flow forecasts by analysts (Chen et al., 2013).

[5] Return predictors identified in the literature include, i.a., valuation ratios (Campbell and Shiller, 1988a; Fama and French, 1988; Lamont, 1998), interest rates spreads (Fama and French, 1989; Ang and Bekaert, 2007), stock market volatility (Guo, 2006; Martin, 2017), output gap (Cooper and Priestley, 2009), the consumption-wealth ratio (Lettau and Ludvigson, 2001, 2005) and risk appetite (Bekaert et al., 2022). Dividend growth predictors include, i.a., consumption ratios (Lettau and Ludvigson, 2005), dividend and earnings yields (Ang and Bekaert, 2007; Møller and Sander, 2017) and the long-run relationship between prices, earnings and dividends (Garrett and Priestley, 2012).

evaluate the predictability of dividends and returns (specifically: the *variance* of expected returns and expected dividend growth) separately from the contribution of expected returns to the dividend yield variance (specifically: the *covariance* of expected returns and expected dividend growth with the dividend yield).

Our results largely provide support for the 'offsetting hypothesis' by Cochrane (2008, 2011). Despite the increased volatility of cash flow growth expectations, we do not find a dramatic increase in the relative contribution of expected dividend growth to the dividend yield decomposition. Similar to the VAR (e.g., Cochrane, 2008) and recent latent variable approaches (van Binsbergen and Koijen, 2010; Zhu, 2015; Choi et al., 2017), our LP approach confirms that the discount rate channel remains the primary contributor to dividend yield volatility. We emphasize that our results are specific to our US sample. When considering different time periods (e.g. Chen, 2009; Golez and Koudijs, 2018), or international evidence (Rangvid et al., 2014) it is possible that the results are different.

The remainder of this paper is organized as follows. Section 2 presents the methodology behind our dividend yield decomposition based on local projections. Section 3 describes the data. Section 4 presents empirical results with the dividend yield as the single state variable, both from using local projections with time-invariant (constant) and time-varying parameters. The results of using various extended selections of state variables are presented in Section 5. We provide a discussion of our findings in Section 6 and Section 7 concludes. Appendix A provides detailed descriptions of previously used methods to consider dividend yield volatility decomposition. Supplementary results are documented in our Internet Appendix.

## 2. Methodology

### 2.1. Present-value framework

Our starting point is the log-linearized present value model by Campbell and Shiller (1988b), who show that the return on holding an asset for one period ($R_{t+1} = \frac{P_{t+1}+D_{t+1}}{P_t}$) can be approximated by a linear equation:

$$r_{t+1} \approx \kappa - \rho dp_{t+1} + dp_t + \Delta d_{t+1}, \tag{1}$$

where $r_t \equiv \log\left(R_t\right)$, $dp_t \equiv \log\left(\frac{D_t}{P_t}\right)$, and $\Delta d_t \equiv \log\left(\frac{D_t}{D_{t-1}}\right)$. In (1), all variables are typically interpreted as deviations from means, such that the constant term $\kappa$ can be omitted:

$$r_{t+1} \approx -\rho dp_{t+1} + dp_t + \Delta d_{t+1}, \tag{2}$$

where $\rho$ is required to be below, but close to 1. Empirically, $\rho$ is typically estimated as

$$\hat{\rho} = \frac{e^{\bar{dp}}}{1+e^{\bar{dp}}}, \tag{3}$$

where $\bar{dp}$ is the sample average of the log dividend yield $dp_t$. Rearranging (2) and iterating forward results in the dividend yield expressed in terms of discounted future returns, dividend growth rates, and dividend yields:

$$dp_t \approx \sum_{j=1}^{k} \rho^{j-1} r_{t+j} - \sum_{j=1}^{k} \rho^{j-1} \Delta d_{t+j} + \rho^k dp_{t+k}. \tag{4}$$

The identity (4) should hold ex-post as well as ex-ante conditional on any information set $\Omega_t$ (see, e.g., Campbell and Shiller, 1988b; Campbell, 1991; and Cochrane, 2008). Therefore, taking expectations of (4), conditional on the information set $\Omega_t$ available at time $t$ (i.e., $E_t(\cdot) \equiv E(\cdot|\Omega_t)$), results in

$$\begin{aligned} dp_t &\approx E_t \sum_{j=1}^{k} \rho^{j-1} r_{t+j} - E_t \sum_{j=1}^{k} \rho^{j-1} \Delta d_{t+j} + E_t \rho^k dp_{t+k} \\ &\equiv \delta_t^{(r,k)} - \delta_t^{(d,k)} + \delta_t^{(dp,k)}. \end{aligned} \tag{5}$$

The finite-horizon expression (5) implies that the dividend yield contains three components: *(i)* discounted expected returns up to $k$ periods $\delta_t^{(r,k)}$, *(ii)* discounted expected dividend growth rates up to $k$ periods $\delta_t^{(d,k)}$, and *(iii)* the discounted expected dividend yield in $k$ periods, $\delta_t^{(dp,k)}$, which in turn implies expectations of both returns and dividends over horizons longer than $k$ periods.

Multiplying (5) by $dp_t$ and taking expectations leads to an expression

$$\begin{aligned} \mathrm{Var}(dp_t) &\approx \mathrm{Cov}\Big(dp_t, E_t \sum_{j=1}^{k} \rho^{j-1} r_{t+j}\Big) - \mathrm{Cov}\Big(dp_t, E_t \sum_{j=1}^{k} \rho^{j-1} \Delta d_{t+j}\Big) + \mathrm{Cov}\Big(dp_t, \rho^k E_t dp_{t+k}\Big) \\ &\equiv \mathrm{Cov}\Big(dp_t, \delta_t^{(r,k)}\Big) - \mathrm{Cov}\Big(dp_t, \delta_t^{(d,k)}\Big) + \mathrm{Cov}\Big(dp_t, \delta_t^{(dp,k)}\Big). \end{aligned} \tag{6}$$

This expression is the basis for the dividend yield decomposition that is of main interest in this paper. The key objective is to evaluate the magnitudes of these covariance components at different horizons $k$, focusing primarily on the cash flow and return components. In the existing literature, researchers often focus on infinite horizons ($k \longrightarrow \infty$), combined with the assumption that rational bubbles cannot exist (i.e., the transversality condition $\lim_{k\to\infty} E_t\left[\rho^k dp_{t+k}\right] = 0$ holds). Under these conditions, the identity (5) converges to:

$$\begin{aligned} dp_t &\approx E_t \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} - E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} \\ &\equiv \delta_t^{(r,\infty)} - \delta_t^{(d,\infty)}, \end{aligned} \tag{7}$$

and the third covariance component in (6) vanishes. The dividend yield $dp_t$ thus reflects expected discounted returns and dividend growth rates, both up to infinite horizons. This representation yields the important insight by Cochrane (2008) that observing variation in the dividend yield implies that either future returns or dividends, or both, must be predictable.

Prior studies, such as Campbell and Shiller (1988b) and Cochrane (2008), apply a vector autoregression (VAR) to evaluate the relative contributions of the infinite-horizon components (7) and the covariance components in (6). These VAR-based approaches are briefly outlined in Appendix A. The essential idea is to obtain long-run discounted expectations on future returns ($\delta_t^{(r,\infty)}$) and dividend growth rates ($\delta_t^{(d,\infty)}$) by iterating forward the predictions of a one-period VAR. Assuming that one wants to evaluate the expected components only at an infinite horizon ($k \longrightarrow \infty$), the linear structure of the VAR has the advantage of allowing for closed-form solutions of the infinite horizon predictions (see Campbell and Shiller (1988b), Cochrane (2008) and Appendix A for details).

The VAR approach also has several disadvantages, as discussed in the Introduction. First, it assumes that the estimated VAR is the correct data generating process for all three components (5) and at all horizons $k$. Second, VARs have limited capacity to incorporate large sets of state variables since the number of parameters increases quadratically in the number of variables. Third, VARs are restrictive also in a sense that they do not allow the 'mixed-frequency' matching between annual and monthly data. In this paper, we apply local projections to tackle these concerns.

### 2.2. Local projections-based dividend yield decomposition: general setup

In this section, we introduce a dividend yield decomposition that is built upon flexible local projections (LPs) to evaluate the relative magnitudes of the contributions of expected returns (discount rates) and expected growth in dividends (cash flows) to the variation of the dividend yield. The use of local projections in structural macroeconometric inference originates from the work of Jordà (2005). At the heart of this approach lie forecast horizon-specific predictions for each of the three components of interest in the identity (5). These predictions form the basis for a flexible dividend yield volatility decomposition (6) where the forecast horizon $k$ can freely vary between short, intermediate and long-term horizons. That is, we construct (linear) local projections for the $k$-period ahead cumulative returns, cumulative dividend growth rates, and the $k$-period ahead dividend yield as dependent variables:

$$\begin{aligned}\sum_{j=1}^{k}\rho^{j-1}r_{t+j} &= \alpha^{(r,k)}+\boldsymbol{x}_t^{(r,k)}\boldsymbol{\beta}^{(r,k)}+\varepsilon_{t+k}^{(r,k)}\\ \sum_{j=1}^{k}\rho^{j-1}\Delta d_{t+j} &= \alpha^{(d,k)}+\boldsymbol{x}_t^{(d,k)}\boldsymbol{\beta}^{(d,k)}+\varepsilon_{t+k}^{(d,k)}\\ \rho^k dp_{t+k} &= \alpha^{(dp,k)}+\boldsymbol{x}_t^{(dp,k)}\boldsymbol{\beta}^{(dp,k)}+\varepsilon_{t+k}^{(dp,k)},\end{aligned} \tag{8}$$

where $\boldsymbol{x}_t^{(a,k)}$ and $\varepsilon_{t+k}^{(a,k)}$, $a\in\{r,d,dp\}$, are the vectors of state variables and zero-mean error terms, respectively. The state variables $\boldsymbol{x}_t^{(a,k)}$ may differ across horizons ($k$) and for each equation of interest '($a$)'. Due to the linear structure of (8), each equation can be consistently estimated by ordinary least squares (OLS) under general conditions. For clarity, in the case of multiple state variables, $\boldsymbol{x}_t^{(a,k)}$ and $\boldsymbol{\beta}^{(a,k)}$ refer to row and column vectors, respectively.

The conditional expectations (or fitted values) of the left-hand-side (LHS) variables in (8), conditional on the information set at time $t$ and the estimated parameters, are the empirical counterparts of $\delta_t^{(r,k)}$, $\delta_t^{(d,k)}$, and $\delta_t^{(dp,k)}$ in (5):

$$\hat{\delta}_t^{(a,k)}=\hat{\alpha}^{(a,k)}+\boldsymbol{x}_t^{(a,k)}\hat{\boldsymbol{\beta}}^{(a,k)},\qquad a\in\{r,d,dp\}. \tag{9}$$

Due to the flexible structure of LPs, the resulting estimates $\hat{\delta}_t^{(a,k)}$ are expected to be more informative and less prone to model misspecification than the estimates obtained from the VAR approaches discussed in Section 2.1 and Appendix A. Importantly, even if the VAR is in fact the correct data generating process, the LPs containing the same state variables are asymptotically equivalent to the VAR predictions, whereas the reverse does not apply (Jordà, 2005). Therefore, in large samples, nothing is lost in terms of estimating $\delta_t^{(a,k)}$ when using LPs instead of the correctly-specified VAR-based approaches.

In addition to the VAR-based approach, Cochrane (2008, 2011) and Maio and Santa-Clara (2015) also consider the dividend yield decomposition using 'direct regressions' with the dividend yield as the only state variable. This is also our starting point and can be seen as a restricted baseline case of (8), where $\boldsymbol{x}_t^{(a,k)}=dp_t$ for all $k$ and for $a\in\{r,d,dp\}$. In Section 4, we estimate the components $\delta_t^{(r,k)}$, $\delta_t^{(d,k)}$, and $\delta_t^{(dp,k)}$ by fitting the regressions (8) using the dividend yield $dp_t$ as the single state variable. There is, however, no a priori reason to assume that the dividend yield is the only state variable of long-run dividends and returns. Lettau and Ludvigson (2005) demonstrate that, even if identity (7) holds, expected returns and dividends may share a common component that is independent of the dividend yield, implying that additional variables beyond the dividend yield may be useful predictors of long-run returns and dividend growth. In Section 5, we consider multiple state variables selected from a larger set of financial and macroeconomic variables.

After estimating the local projections with a given set of state variables, the empirical counterparts of $\delta_t^{(r,k)}$, $\delta_t^{(d,k)}$, and $\delta_t^{(dp,k)}$ for different horizons $k$, not just $k\longrightarrow\infty$ as in the VAR approach, provide the relative covariance components:

$$\hat{\gamma}(r,k)=\frac{\mathrm{Cov}\left(\widehat{dp}_t^{(k)},\hat{\delta}_t^{(r,k)}\right)}{\mathrm{Var}(\widehat{dp}_t^{(k)})},\quad -\hat{\gamma}(d,k)=-\frac{\mathrm{Cov}\left(\widehat{dp}_t^{(k)},\hat{\delta}_t^{(d,k)}\right)}{\mathrm{Var}(\widehat{dp}_t^{(k)})},\quad \hat{\gamma}(dp,k)=\frac{\mathrm{Cov}\left(\widehat{dp}_t^{(k)},\hat{\delta}_t^{(dp,k)}\right)}{\mathrm{Var}(\widehat{dp}_t^{(k)})}, \tag{10}$$

where especially the first two components are of main interest. In these expressions, we use the 'implied' dividend-price ratio $\widehat{dp}_t^{(k)}=\hat{\delta}_t^{(r,k)}-\hat{\delta}_t^{(d,k)}+\hat{\delta}_t^{(dp,k)}$ to ensure that the three components (10) sum up to one. Together with (6) the relative contributions in (10) form our dividend yield decomposition.

Following prior literature (e.g., Campbell and Shiller, 1988b; van Binsbergen and Koijen, 2010), an alternative view to assess the relative contributions (10) is obtained by taking variances of the present-value identity (5). This leads to a 'variance decomposition'

$$\begin{aligned}\mathrm{Var}\left(dp_t\right) = {} & \mathrm{Var}\left(\delta_t^{(r,k)}\right)+\mathrm{Var}\left(\delta_t^{(d,k)}\right)+\mathrm{Var}\left(\delta_t^{(dp,k)}\right)\\ & -2\Big[\mathrm{Cov}\left(\delta_t^{(r,k)},\delta_t^{(d,k)}\right)-\mathrm{Cov}\left(\delta_t^{(r,k)},\delta_t^{(dp,k)}\right)\\ & +\mathrm{Cov}\left(\delta_t^{(d,k)},\delta_t^{(dp,k)}\right)\Big].\end{aligned} \tag{11}$$

We consider the following volatility components

$$\hat{\sigma}(r,k)=\mathrm{Std}\left(\hat{\delta}_t^{(r,k)}\right),\quad \hat{\sigma}(d,k)=\mathrm{Std}\left(\hat{\delta}_t^{(d,k)}\right),\quad \hat{\sigma}(dp,k)=\mathrm{Std}\left(\hat{\delta}_t^{(dp,k)}\right), \tag{12}$$

which measure the degree of variation in expectations and correspond to the (square root of the) first three elements of (11). However, since expected returns and dividend growth rates are highly correlated (see Lettau and Ludvigson, 2005; and Golez, 2014), the fourth term in (11), $\mathrm{Cov}\left(\delta_t^{(r,k)},\delta_t^{(d,k)}\right)$ is non-negligible such that the volatility components (12) do not contain the same information as the covariance-based decomposition (10). Throughout this paper, we therefore measure variation in the expectations of discount rates, cash flow growth and future dividend yields (12), using different state variables, but we are primarily interested in to what extent these expectations *covary* with the dividend yield, implying their contribution to the dividend yield decomposition (10).

Unlike the volatility decomposition (12), our covariance-based decomposition (10) incorporates not only the magnitude but also the signs of the covariances. Based on (5), it is expected that discounted expected returns (dividend growth rates) are positively (negatively) correlated with the dividend yield. However, as documented in the empirical Sections below, in some of our VAR and LP specifications we find, in fact, a positive correlation between discounted expected dividend growth rates $\delta_t^{(d,k)}$ and the dividend yield, implying a negative impact of cash flow expectations on dividend yield volatility.

### 2.3. *Monthly local projections*

Prior studies on volatility and dividend yield decompositions are implemented with annual data (see, e.g., Campbell and Shiller, 1988b; Cochrane, 1992, 2008, 2011; van Binsbergen and Koijen, 2010; Zhu, 2015; and Choi et al., 2017). Since dividend payments are highly seasonal, dividend growth rates are often considered informative only on an annualized basis.[6] One of our contributions is that local projections allow for the use of monthly (or even higher-frequency) data explicitly when modeling the component $\delta_t^{(d,k)}$ associated with the expected dividend growth in (5), which is the typical frequency in most predictability studies.

The essential challenge at the monthly frequency is the measurement of dividends. Even if dividend data are available at the monthly

[6] Closely related return decomposition studies do use monthly data, but compared to (8), these studies only model expected returns explicitly and treat the contribution of expected dividend growth as a residual term (see, e.g., Campbell, 1991; Campbell and Ammer, 1993). Chen and Zhao (2009) criticize the residual approach for leading in general to unrobust results. See also the recent advancement by Pettenuzzo et al. (2020) on modeling dynamics in aggregate cash flows extracted from daily firm-level dividend announcements.

frequency, these figures have strong seasonalities, resulting in highly erratic behavior of the monthly dividend yield ($dp_t$) and dividend growth rate ($\Delta d_t$). We are able to tackle this challenge with the local projection approach by using monthly updated observations of 12-month cumulative dividends. In the following, we treat the monthly dividend growth rates as unobserved, while cumulative 12-month dividend growth rates are observable each month. Monthly dividend yields refer, as common in the literature, to the cumulative 12-month dividend yield. Finally, the return series are simply monthly returns. Section 3 provides further details on the variable definitions.

In the absence of reliable observable monthly dividend growth rates, instead of accumulating dividend growth rates month by month, we accumulate the growth rates in batches of 12 months:

$$\sum_{j=1}^{k} \rho^{j-1} \Delta d_{t+j} \approx \sum_{j=1}^{k} \rho_j^* \Delta^* d_{t+j}^* = \rho_{12}^* \Delta^* d_{t+12}^* + \rho_{24}^* \Delta^* d_{t+24}^* + \cdots, \tag{13}$$

where $\rho_j^* \equiv \rho^{j-1}$ if $j \in \{12, 24, 36, \ldots\}$ and 0 otherwise, $d_t^*$ is the *cumulative* dividend over months $t-11$ up to month $t$ and $\Delta^* d_{t+j}^* \equiv d_{t+j}^* - d_{t+j-12}$ is the annual growth rate in these 12-month cumulative dividends. Thus, to obtain the monthly dividend yield decomposition, we replace the (unobserved) monthly dividend growth component in identity (5) and the local projections (8) by the right hand side of (13), which is based on monthly observations of growth rates in 12-month cumulative dividends. Although the approximation is not exact, we find that $\sum_{j=1}^{k} \rho_j^* \Delta^* d_{t+j}^*$ closely approximates $\sum_{j=1}^{k} \rho^{j-1} \Delta d_{t+j}$, especially at longer horizons $k$. In the Internet Appendix Section VII, we present a simulation exercise to illustrate the approximation and find a correlation $\text{Cor}\left(\sum_{j=1}^{k} \rho^{j-1} \Delta d_{t+j}, \sum_{j=1}^{k} \rho_j^* \Delta^* d_{t+j}^*\right)$ of 0.98 for a horizon of $k = 180$ months, which is the main horizon of interest in this paper.

The approximation (13) implies that the (monthly) forecast horizon $k$ needs to be a multiple of 12 (i.e. $k \in \{12, 24, 36, \ldots\}$). In other words, despite basing the analysis on monthly data, we report the monthly decompositions for annualized horizons.

From an empirical point of view, the introduction of a monthly data dividend yield decomposition, resulting from the development of the local projection framework, implies substantially more observations and therefore increased statistical power, which provides additional robustness of our annual decompositions and opens up possible model extensions. In particular, the increased number of observations allows us to study time variation in our dividend yield decomposition.

### 2.4. Time-varying parameters

The larger sample size resulting from the use of monthly data, enabled by our local projections approach, allows the underlying parameters of our dividend yield decomposition to be time-varying. In the traditional VAR approach with annual data, this would be infeasible in practice because of the inevitably limited sample size.

The motivation for studying time-variation is grounded in recent and mounting empirical and theoretical evidence suggesting that return predictability is time-varying (e.g., Timmermann, 2008; Rapach et al., 2010; Henkel et al., 2011; Dangl and Halling, 2012; Zhu, 2015; Zhu and Zhu, 2013; Cochrane, 2017; and Farmer et al., 2023), which may originate from various economic reasons, including business cycle fluctuations, time-varying risk aversion, and rare disasters. As summarized by Timmermann (2008), investors' search for successful predictive models is expected to cause the data generating process to change over time, which means that time-invariant parameter models can at best hope to uncover only some local predictability. Recent findings suggest that predictive power often concentrates on bad times in financial markets (see Henkel et al., 2011; Zhu and Zhu, 2013; Cujean and Hasler, 2017). Zhu (2015) also finds that time-varying predictable patterns in returns and dividend growth rate is a tug-of-war: when returns are predictable, dividend growth is not, and vice versa. Furthermore, Choi et al. (2017) find that incorporating regime shifts into the present-value framework of van Binsbergen and Koijen (2010) strengthens the importance of dividend growth variation in explaining both the price-dividend ratio and unexpected stock returns in the post-1951 sample period.

As argued in the previous sections, the use of horizon-specific local projections may well reduce model misspecification concerns. Allowing for time variation in the parameter coefficients may importantly further increase the accuracy of the estimated discount rate and cash flow components. Therefore, we extend the local projections (8) by allowing for time-varying parameters (TVP):

$$\begin{aligned}
\sum_{j=1}^{k} \rho_t^{j-1} r_{t+j} &= \alpha_t^{(r,k)} + \boldsymbol{x}_t^{(r,k)} \boldsymbol{\beta}_t^{(r,k)} + \varepsilon_{t+k}^{(r,k)} \\
\sum_{j=1}^{k} \rho_t^{j-1} \Delta d_{t+j} &= \alpha_t^{(d,k)} + \boldsymbol{x}_t^{(d,k)} \boldsymbol{\beta}_t^{(d,k)} + \varepsilon_{t+k}^{(d,k)} \\
\rho_t^k dp_{t+k} &= \alpha_t^{(dp,k)} + \boldsymbol{x}_t^{(dp,k)} \boldsymbol{\beta}_t^{(dp,k)} + \varepsilon_{t+k}^{(dp,k)},
\end{aligned} \tag{14}$$

where all the parameters $\alpha_t^{(a,k)}$ and $\boldsymbol{\beta}_t^{(a,k)}$ ($a \in \{r, d, dp\}$) are now time-varying. Specifically, we allow for time variation by estimating the coefficients in (14) recursively using Exponentially Weighted Least Squares (EWLS), which is a particular case of Weighted Least Squares estimation in which the weight of each observation $i$ in a sample of size $t$ is given by $\left(\sum_{i=0}^{t} \phi^{t-i}\right)^{-1} \phi^{t-i}$. The decay parameter $\phi$ is a number between zero and one. Following the convention in the literature, we calibrate the decay parameter $\phi$ at 0.97, which is suggested by Longerstaey and Spencer (1996) in J.P. Morgan's Riskmetrics report as the optimal exponential decay parameter for modeling volatility using monthly data, but also other parameter values can be entertained accordingly. Applying an expanding window estimation approach combined with exponential weighting ensures that most weight is given to recent observations, while the weights of distant past observations gradually fade.[7]

In addition to the regression parameters, we also allow $\rho$ to vary over time, by applying a similar expanding window scheme combined with exponential weights. That is, instead of estimating $\rho$ over the full sample (cf. Eq. (3)):

$$\hat{\rho}_t = \frac{e^{\widetilde{dp}_t}}{1 + e^{\widetilde{dp}_t}}, \tag{15}$$

where $\widetilde{dp}_t$ is the exponentially weighted moving average (EWMA) of the dividend yield up to period $t$:

$$\widetilde{dp}_t = \left(\sum_{i=0}^{t} \phi^{t-i}\right)^{-1} \sum_{i=0}^{t} \phi^{t-i} dp_i, \tag{16}$$

where, as in the regressions (14), the decay parameter $\phi$ is set at 0.97.

In addition to allowing for time-varying coefficients in prediction as such, Lettau and Van Nieuwerburgh (2008) show that the poor performance of financial ratios as predictors of returns can be improved if the assumption of a fixed and time-invariant steady state mean of the economy is relaxed. That is, adjusting the dividend yield, earnings yield and book-to-market ratio for level shifts increases the predictive performance substantially. Lettau and Van Nieuwerburgh (2008) correct these nonstationarities by estimating the timing of structural break points. However, locating the exact timing of breaks or identifying regime switching patterns (cf. Zhu, 2015; Choi et al., 2017) is generally difficult, in particular in small samples. As an alternative approach, we handle the time variation of the steady-state levels of variables by detrending not only the dividend yield, but also all other variables using the same recursive EWMA filtration as in (16). Specifically, in our time-varying local projections (14), we recursively detrend all variables by subtracting the exponentially weighted moving average at each point

[7] See Hallerbach and Menkveld (2004), Kofman and McGlenchy (2005), and Taylor (2008), for applications of EWLS.

in time, as opposed to demeaning these variables over the full sample. Thus, we implicitly allow the parameter $\kappa$ in the log-linear present value model (1) to vary over time.

Given the time-varying patters, in addition to the unconditional (i.e., full sample) decompositions (10), our exponentially weighted expanding window approach allows us to compute time-varying variances and covariances:

$$\widehat{\text{Var}}_t\left(\widehat{dp}_t^{(k)}\right) = \left(\sum_{i=0}^{t}\phi^{t-i}\right)^{-1}\sum_{i=0}^{t}\phi^{t-i}\left(\widehat{dp}_i^{(k)}\right)^2$$
$$\widehat{\text{Cov}}_t\left(\widehat{dp}_t^{(k)},\hat{\delta}_t^{(a,k)}\right) = \left(\sum_{i=0}^{t}\phi^{t-i}\right)^{-1}\sum_{i=0}^{t}\phi^{t-i}\,\widehat{dp}_i^{(k)}\hat{\delta}_i^{(a,k)}, \quad a\in(r,d,dp), \tag{17}$$

where the elements $\hat{\delta}_t^{(r,k)}$, $\hat{\delta}_t^{(d,k)}$, $\hat{\delta}_t^{(dp,k)}$ and the implied dividend-yield $\widehat{dp}_t^{(k)} = \hat{\delta}_t^{(r,k)} - \hat{\delta}_t^{(d,k)} + \hat{\delta}_t^{(dp,k)}$ have a conditional mean of zero, because they are obtained after demeaning all underlying state-variables recursively (i.e. all variables are demeaned by subtracting the exponentially weighted mean using filter (16)). The ratio of the time-varying covariance to the variance of the dividend yield provides the time-varying contributions of expected discount rates, expected cash flow growth rates and expected dividend yields to the observed dividend yield: $\hat{\gamma}_t(a,k)$, for $a$ in $(r,d,dp)$, as the time-varying counterparts of the static contributions (10).

## 3. Data

Our main variables are market-level log returns, dividend growth rates, and dividend yields, observed over the period 1928–2020. For our monthly analysis, we use the monthly value-weighted market return reported by the Center for Research in Security Prices (CRSP). Following Koijen and Van Nieuwerburgh (2011), among others, we compute monthly dividends by $D_t = (R_t - Rx_t)P_{t-1}$, where $R_t$ refers to the gross CRSP value-weighted market in month $t$, and $Rx_t$ refers to the gross CRSP value-weighted market return over the same period *excluding* dividends. To avoid seasonality concerns, monthly dividends are accumulated over 12 months before computing 12-month (log) growth rates, as discussed in Section 2.3. The annualized log dividend yield in month $t$ is computed as the accumulated dividends over the months $t-11$ to $t$ divided by the price at the end of month $t$.

Monthly dividends are accumulated under the assumption that dividends paid out during the 12-month periods are at the end of each month reinvested in the risk-free rate of return, following Chen (2009), van Binsbergen and Koijen (2010), and others. Alternatively, Cochrane (2008) assumes that dividends are reinvested in the market portfolio. Chen (2009) discusses the implications of these different assumptions on the predictability of dividends and argues that it is difficult to disentangle return predictability from dividend predictability when dividends are reinvested in the market. For robustness, Section II of the Internet Appendix reports our main results computed using market-reinvested and non-reinvested dividends.

For our analysis of annual data, we use cumulative returns over each calendar year (January–December), and use the end of year (December) observations of the 12-month dividend growth rate and dividend yield.

In Section 5, we consider various additional state variables. First, we focus on a selection of state variables that have been identified in the recent literature as strong predictors of the dividend growth rate and also of market returns, listed in Panel A of Table 1. Lettau and Ludvigson (2005) propose the long-run (cointegration) relationship between consumption, asset wealth and labor income ($cay$) and an analogously derived long-run relationship between consumption, dividends and labor income ($cdy$) as (theoretically-motivated) predictors of stock returns and dividend growth, respectively. Various studies have gauged the joint dynamics of earnings and dividends as an indicator of future dividend growth, inspired by Lintner's (1956) dividend model: Garrett and Priestley (2012) propose the estimated long-run (cointegration) relationship between dividends, (market) prices and earnings ($dpe$) as a predictor of dividend growth. Ang and Bekaert (2007) and Møller and Sander (2017) find that the dividend yield ($dp$) and the earnings yield ($ep$) hold jointly predictive power of dividend growth rates. We also include the dividend to earnings ratio ($de$) as a candidate state variable. Campbell and Vuolteenaho (2004) and Garrett and Priestley (2012) use the price-earnings ratio with earnings smoothed over a 10-year period ($pe10$), as well as the value spread among small stocks ($vs$) and the term spread ($tms$) as return predictors. In addition, Cooper and Priestley (2009) have emphasized the importance of the output gap ($gap$), measuring the difference between realized and potential output, as a predictor of stock returns. Detailed descriptions of the data are provided in Section I of the Internet Appendix.

**Table 1**
Additional state variables. Panel A lists selected predictors used in the recent literature to predict dividend growth and market returns (see Campbell and Vuolteenaho, 2004; Lettau and Ludvigson, 2005; Ang and Bekaert, 2007; Cooper and Priestley, 2009; Garrett and Priestley, 2012; and Møller and Sander, 2017). Panel B lists additional state variables obtained from the dataset of Welch and Goyal (2008). Panel C lists option-implied state variables that are available for the 1990–2020 subsample. '(A)' indicates data availability only at annual frequency, while '(M)' indicates variables used only at monthly frequency. Section I of the Internet Appendix provides details on the variables.

| | |
|---|---|
| A: Selected predictors of dividend growth and returns | |
| $cay$ | Cointegration relationship between consumption, asset wealth and labor income |
| $cdy$ | Cointegration relationship between consumption, dividends and labor income |
| $dpe$ | Cointegration relationship between dividends, prices and earnings |
| $ep$ | Log earnings-price ratio |
| $de$ | Log dividend-earnings ratio |
| $pe10$ | Log price-10 year earnings ratio |
| $vs$ | Value spread |
| $tms$ | Term spread: $lty - tbl$ |
| $gap$ | Output gap (extracted based on industrial production) |
| B: Other state variables from the Welch-Goyal dataset | |
| $svar$ | Realized volatility (monthly sum of squared daily returns on the S&P 500) |
| $bm$ | Book-to-market value ratio for the DJIA (Dow Jones Industrial Average). |
| $ntis$ | Net equity expansion |
| $tbl$ | Treasury bill rate (three-month Treasury bill, secondary market) |
| $lty$ | Long-term government bond yield |
| $rfree$ | Risk-free rate |
| $ltr$ | Return on long-term government bonds |
| $corpr$ | Return on long-term corporate bond |
| $dfy$ | Default yield spread |
| $dfr$ | Default return spread: $corpr - ltr$ |
| $infl$ | Inflation (CPI inflation) |
| $eqis$ (A) | Percent equity issuing |
| $ik$ (A) | Investment-to-capital ratio |
| C: Option-implied state variables | |
| $cv$ (M) | Conditional variance of stock returns (Bekaert and Hoerova, 2014) |
| $vp$ (M) | Equity variance premium (Bekaert and Hoerova, 2014) |
| $ra$ (M) | Risk aversion index (Bekaert et al., 2022) |
| $unc$ (M) | Uncertainty index (Bekaert et al., 2022) |

Panel B of Table 1 lists additional state variables that are based on the prior extensive literature on the predictability of market-level equity returns (see the discussion in the Introduction and in the surveys by Welch and Goyal, 2008, and Rapach and Zhou, 2013). These additional variables are obtained from the updated dataset of Welch and Goyal (2008), who provide detailed descriptions of the data and their sources. When considering the state variables in Table 1, we restrict our sample period to 1952–2020 (monthly observations start in March 1952). This choice of starting point of the sample period (1952) is mainly driven by the data availability of $cdy$ and $cay$, which the prior literature has identified as important predictors of both dividends and returns. This starting point coincides with Cochrane (2011) and Campbell and Am-

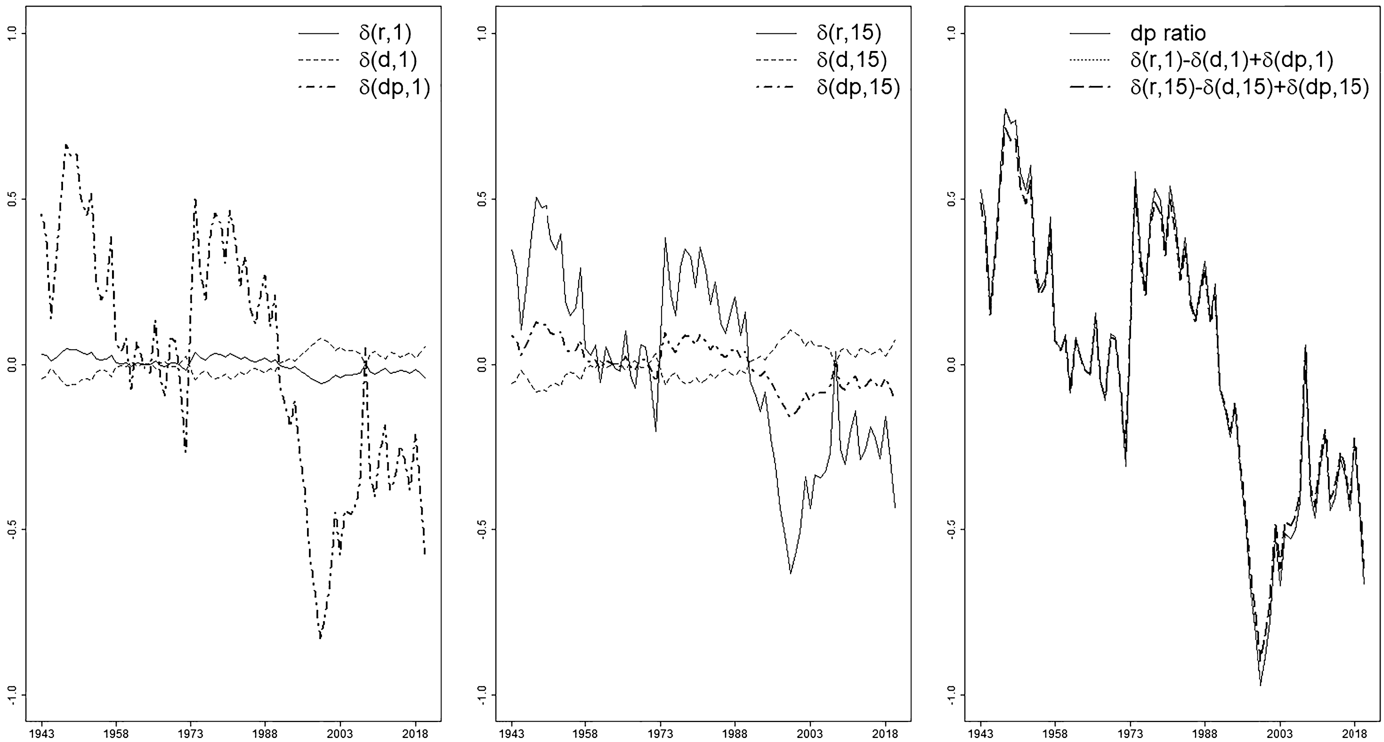


**Fig. 1.** Plots of the components $\widehat{\delta}_t^{(r,k)}$, $\widehat{\delta}_t^{(d,k)}$, and $\widehat{\delta}_t^{(dp,k)}$ (see (5)), estimated from the annual local projections (8), using $dp_t$ as the single state variable, for $k = 1$ year (left panel) and $k = 15$ years (middle panel). The right panel shows the observed dividend yield $dp_t$ and the implied dividend yield $\widehat{dp}_t^{(k)} = \widehat{\delta}_t^{(r,k)} - \widehat{\delta}_t^{(d,k)} + \widehat{\delta}_t^{(dp,k)}$, for $k = 1$ year and $k = 15$ years.

mer (1993) and is very close to the ones in Cochrane (2008), Lettau and Ludvigson (2005) and van Binsbergen and Koijen (2010).

Finally, Panel C of Table 1 lists four recently proposed indicators of market risk. Unlike traditional return predictors, such as the dividend yield, these indicators are less persistent and respond more rapidly to changing market conditions. These four measures all require option prices (among other financial variables), which have been documented to contain incremental forward-looking information over historical returns (see also, e.g., Bollerslev et al., 2009; and Martin, 2017). Bekaert and Hoerova (2014) decompose the squared VIX index into the conditional variance of stock returns ($cv$) and the equity variance premium ($vp$): $vp_t = VIX_t^2 - cv_t$, where $cv_t$ is the predicted realized variance of the S&P 500 index over the next month. Bekaert and Hoerova (2014) find that $vp$ has predictive power for future stock returns, while $cv$ predicts real economic activity and financial turmoil. The two remaining state variables are proposed by Bekaert et al. (2022): $ra$ is a risk-aversion index and $unc$ is an uncertainty index, estimated from a structural model with an instrument set including earnings yield, bond spreads, equity and bond market realized return variances, and equity option-implied risk-neutral variance. Both indices are found to hold predictive power for equity and bond returns for up to 12 months. Due to the dependence on option-market data, these state variables are available only for a more recent subsample. In Section 5.5, we analyze our decomposition based on monthly observations of these state variables over the period 1990–2020.

## 4. Results: one state variable

We start by considering the benchmark case with the dividend yield as a single state variable, following Cochrane (2008) and others. First, in Section 4.1 we estimate the local projections (LPs), with time-invariant (constant) parameters (Eq. (8)), over the full sample period (1928–2020). In Section 4.2, we consider the local projections with time-varying parameters (Eq. (14)).

### 4.1. Constant parameters

We start by estimating the LPs (8), with $\boldsymbol{x}_t^{(a,k)} = dp_t$ for $a \in \{r, d, dp\}$, and for $k \in \{1, 2, \ldots, 15\}$ years, using annual data. The maximum horizon of 15 years is the same as the longest horizon considered by Cochrane (2011) in his direct regressions.

Fig. 1 plots the estimated components $\widehat{\delta}_t^{(r,k)}$, $\widehat{\delta}_t^{(d,k)}$, and $\widehat{\delta}_t^{(dp,k)}$ (i.e., the fitted values of the local projections (8)), for $k = 1$ and $k = 15$ years. Due to the maximum horizon of $k = 15$ years, the first 15 years of the sample are omitted from the figures. The first panel of Fig. 1 shows that, at short horizons, only little dividend yield variation can be explained by variation in expected discount rates or cash flows: the time series $\widehat{\delta}_t^{(r,1)}$ and $\widehat{\delta}_t^{(d,1)}$ are mostly flat, while the expected dividend yield $\widehat{\delta}_t^{(dp,1)}$ is highly volatile. At longer horizons ($k = 15$ years), a substantial part of dividend yield variation is captured by expected discount rate variation ($\widehat{\delta}_t^{(r,15)}$). The cash flow component $\widehat{\delta}_t^{(d,15)}$ remains relatively flat, suggesting that cash flow expectations, even at longer horizons, can only explain a minor part of observed market volatility.

Following Eq. (5), the final panel of Fig. 1 plots the observed dividend yield and the implied dividend yield obtained as $\widehat{dp}_t^{(k)} = \widehat{\delta}_t^{(r,k)} - \widehat{\delta}_t^{(d,k)} + \widehat{\delta}_t^{(dp,k)}$ for $k = 1$ and $k = 15$. The plot provides supporting evidence on the accuracy of the LP-based estimates of the components in the present-value relation (5), both at short and long horizons, as the implied dividend yields closely trace the observed yield.

Panel A in Table 2 reports the dividend yield decomposition (10) and volatility components (12) at the annual frequency for selected horizons $k$ (results at non-reported horizons are available upon request). These results reaffirm the pattern in Fig. 1. The contribution of expected returns increases over the horizon $k$, up to a maximum of 0.71 at the 15-year horizon ($\widehat{\gamma}(r, 15)$), implying that about 71% of dividend yield volatility can be attributed to variation of long-run expectations of discount rates when considering the covariance with the dividend yield.

**Table 2**
Variance and dividend yield decompositions: One state variable. This table reports the annual (Panel A) and monthly (Panel B) volatility components (12) and dividend yield decomposition (10), based on the local projections (8), using $dp_t$ as the single state variable for different horizons $k$. The first columns report the volatility components of return, dividend growth and dividend yield predictions and ratio $\frac{\hat{\sigma}(d,k)}{\hat{\sigma}(r,k)}$. The final four columns report the relative covariance contributions of expected discount rates $\hat{\gamma}(r,k)$, cash flows $-\hat{\gamma}(d,k)$ and forward dividend yields $\hat{\gamma}(dp,k)$, as well as the ratio $\frac{-\hat{\gamma}(d,k)}{\hat{\gamma}(r,k)}$. The sample period is 1928–2020.

| $k$ | $\hat{\sigma}(r,k)$ | $\hat{\sigma}(d,k)$ | $\hat{\sigma}(dp,k)$ | $\frac{\hat{\sigma}(d,k)}{\hat{\sigma}(r,k)}$ | $\hat{\gamma}(r,k)$ | $-\hat{\gamma}(d,k)$ | $\hat{\gamma}(dp,k)$ | $\frac{-\hat{\gamma}(d,k)}{\hat{\gamma}(r,k)}$ |
|---|---|---|---|---|---|---|---|---|
| (years) | | | | Panel A: Annual | | | | |
| 1 | 0.06 | 0.08 | 0.86 | 1.35 | 0.06 | 0.08 | 0.86 | 1.35 |
| 2 | 0.15 | 0.11 | 0.74 | 0.76 | 0.15 | 0.11 | 0.74 | 0.76 |
| 3 | 0.20 | 0.11 | 0.69 | 0.56 | 0.20 | 0.11 | 0.69 | 0.56 |
| 5 | 0.32 | 0.08 | 0.58 | 0.25 | 0.33 | 0.08 | 0.59 | 0.25 |
| 10 | 0.50 | 0.08 | 0.39 | 0.15 | 0.52 | 0.08 | 0.40 | 0.15 |
| 15 | 0.65 | 0.11 | 0.16 | 0.16 | 0.71 | 0.12 | 0.18 | 0.16 |
| (months) | | | | Panel B: Monthly | | | | |
| 12 | 0.07 | 0.10 | 0.85 | 1.35 | 0.07 | 0.10 | 0.83 | 1.35 |
| 24 | 0.15 | 0.12 | 0.75 | 0.82 | 0.14 | 0.12 | 0.74 | 0.82 |
| 36 | 0.20 | 0.11 | 0.70 | 0.55 | 0.20 | 0.11 | 0.69 | 0.55 |
| 60 | 0.33 | 0.09 | 0.58 | 0.27 | 0.33 | 0.09 | 0.58 | 0.27 |
| 120 | 0.50 | 0.07 | 0.39 | 0.13 | 0.52 | 0.07 | 0.41 | 0.13 |
| 180 | 0.65 | 0.11 | 0.18 | 0.17 | 0.70 | 0.12 | 0.19 | 0.17 |

The contribution of dividend growth expectations ($-\hat{\gamma}(d,k)$) remains low, although not zero: the ratio $\frac{-\hat{\gamma}(d,k)}{\hat{\gamma}(r,k)}$ is 0.16 at the 15-year horizon ($k = 15$). Interestingly, the relative contribution of the cash flow component is diminishing with the horizon $k$.

The volatility-based components $\hat{\sigma}(a,k)$ (see Eq. (12)) show a similar pattern as the covariance-based dividend yield decomposition (10) on the right. Our results are thus qualitatively consistent with Cochrane (2008), in the sense that discount rate expectations are the main driver of price volatility. The contribution of dividend growth expectations, estimated based on our local projections, is only slightly higher than the contribution of dividend growth implied by Cochrane's (2008) VAR approach, especially when the horizon $k$ is short. As we show in Appendix A, our volatility $\hat{\sigma}(a,k)$ and covariance components $\hat{\gamma}(a,k)$ ratios are conceptually equivalent to the long-run coefficients by Cochrane when $k \longrightarrow \infty$, which can be backed out from our local projections with $k = 1$ and the lagged dividend yield as the only state variable. Indeed, Cochrane (2008, Table 4) reports VAR-implied long-run return and dividend contributions of 1.09 and 0.09, respectively, which is qualitatively similar to our decomposition, with the expected discount rate channel dominating the cash flow channel.

Panel B of Table 2 presents the results estimated with monthly data for the same selected yearly horizons as in Panel A up to 15 years (i.e., $k = 12$ to $k = 180$ months) and utilizing the approximation (13). The relative contributions of discount rates and cash flows are very similar, although not fully equivalent, to those reported in Panel A. Another notable result in Table 2 is that the relative impact of the forward dividend yield is diminishing monotonically with the horizon $k$ as discussed around Eq. (7). As evident from Eq. (5), the expected (forward) dividend yield $\hat{\delta}_t^{(dp,k)}$ reflects expectations of dividend growth and returns over horizons beyond $k$ years. Hence, the high volatility of $\hat{\delta}_t^{(dp,1)}$ implies that most of the volatility of the dividend yield is attributed to discount rate and cash flow expectations over horizons beyond one year, while the relatively low volatility of $\hat{\delta}_t^{(dp,15)}$ suggests a relatively minor contribution of expectations over horizons beyond 15 years.[8]

### 4.2. Time-varying parameters

Next, we move to local projections obtained with time-varying parameters (TVP), as described in Section 2.4. At this stage, we still consider the dividend yield as the only state variable and hence the differences to the results presented above emanate exclusively from the time-varying parameters. The local projections are estimated with monthly data, by EWLS using an expanding window approach with an initial window size of 180 months. Since reliable implementation of the recursive estimation requires a reasonable number of observations, we only present results based on monthly data (for annualized horizons). The possibility to address time-varying parameters is thus a direct consequence of the increased number of observations resulting from the ability to use monthly data in the LP framework (as developed in Section 2.3).

The volatility and dividend yield decompositions, based on the local projections with time-varying parameters, are reported in Table 3 for selected horizons $k$. The main pattern is clear: even without extending the information set beyond the dividend yield, both the volatility of expected returns ($\hat{\sigma}(r,k)$) and the volatility of expected dividend growth ($\hat{\sigma}(d,k)$) are considerably higher than with the constant (time-invariant) parameter LPs. The relative increase in expected dividend volatility is higher, leading to a long-run (180 months) volatility ratio $\frac{\hat{\sigma}(d,k)}{\hat{\sigma}(r,k)}$ of 0.64. Interestingly, despite the higher volatility of expected dividends, the covariance of expected dividends with the dividend yield remains low, such that the relative contribution of cash flows to the dividend yield decomposition $\frac{-\hat{\gamma}(d,k)}{\hat{\gamma}(r,k)}$ is substantially smaller at 0.24, for $k = 180$ months. At intermediate horizons (36–60 months), the contribution of dividend expectations $-\hat{\gamma}(d,k)$ is not only small in magnitude but even slightly negative. This means that at these horizons, the discounted expected dividend growth rates ($\delta_t^{(d,k)}$) correlate positively with the dividend yield, implying a small but negative impact of dividend growth expectations on the variance of the dividend yield.

The contrast between the volatility- and covariance-based decompositions is illustrated in Fig. 2. Panel (a) plots the components $\hat{\delta}_t^{(r,k)}$, $\hat{\delta}_t^{(d,k)}$, and $\hat{\delta}_t^{(dp,k)}$ at the 15-year horizon ($k = 180$ months), obtained

[8] Section II of the Internet Appendix provides various robustness checks, by reproducing Table 2 using a shorter sample starting in 1952, using market-reinvested and non-reinvested dividends, and using real returns and dividend growth. Section IV reports full regression (estimation) results of the annual and monthly local projections (8) for all horizons $k$.

**Table 3**

Variance and dividend yield decompositions: Single state variable and time-varying parameters. This table reports the monthly (unconditional) variance and dividend yield decompositions (12), for various (annualized) horizons of $k$ months, based on the local projections (14) with the last observation of the dividend yield as the single state variable, estimated by Exponentially Weighted Least Squares (EWLS) using an expanding window approach. The initial estimation sample is 180 months. All variables are demeaned recursively using EWMA filtration. The first columns report the volatility components of return, dividend growth and dividend yield predictions and ratio $\frac{\hat{\sigma}(d,k)}{\hat{\sigma}(r,k)}$. The final four columns report the relative covariance contributions of expected discount rates $\hat{\gamma}(r,k)$, cash flows $-\hat{\gamma}(d,k)$ and forward dividend yields $\hat{\gamma}(dp,k)$, as well as the ratio $\frac{-\hat{\gamma}(d,k)}{\hat{\gamma}(r,k)}$. Sample period: 1928–2020.

| $k$ (months) | $\hat{\sigma}(r,k)$ | $\hat{\sigma}(d,k)$ | $\hat{\sigma}(dp,k)$ | $\frac{\hat{\sigma}(d,k)}{\hat{\sigma}(r,k)}$ | $\hat{\gamma}(r,k)$ | $-\hat{\gamma}(d,k)$ | $\hat{\gamma}(dp,k)$ | $\frac{-\hat{\gamma}(d,k)}{\hat{\gamma}(r,k)}$ |
|---|---|---|---|---|---|---|---|---|
| 12 | 0.55 | 0.32 | 0.85 | 0.57 | 0.28 | 0.07 | 0.65 | 0.23 |
| 24 | 0.88 | 0.46 | 0.81 | 0.53 | 0.58 | 0.05 | 0.37 | 0.09 |
| 36 | 0.96 | 0.52 | 0.78 | 0.55 | 0.59 | -0.07 | 0.49 | -0.12 |
| 60 | 1.22 | 0.52 | 1.00 | 0.43 | 0.61 | -0.06 | 0.44 | -0.09 |
| 120 | 0.95 | 0.47 | 0.66 | 0.49 | 0.70 | 0.06 | 0.24 | 0.08 |
| 180 | 0.80 | 0.51 | 0.57 | 0.64 | 0.66 | 0.16 | 0.18 | 0.24 |

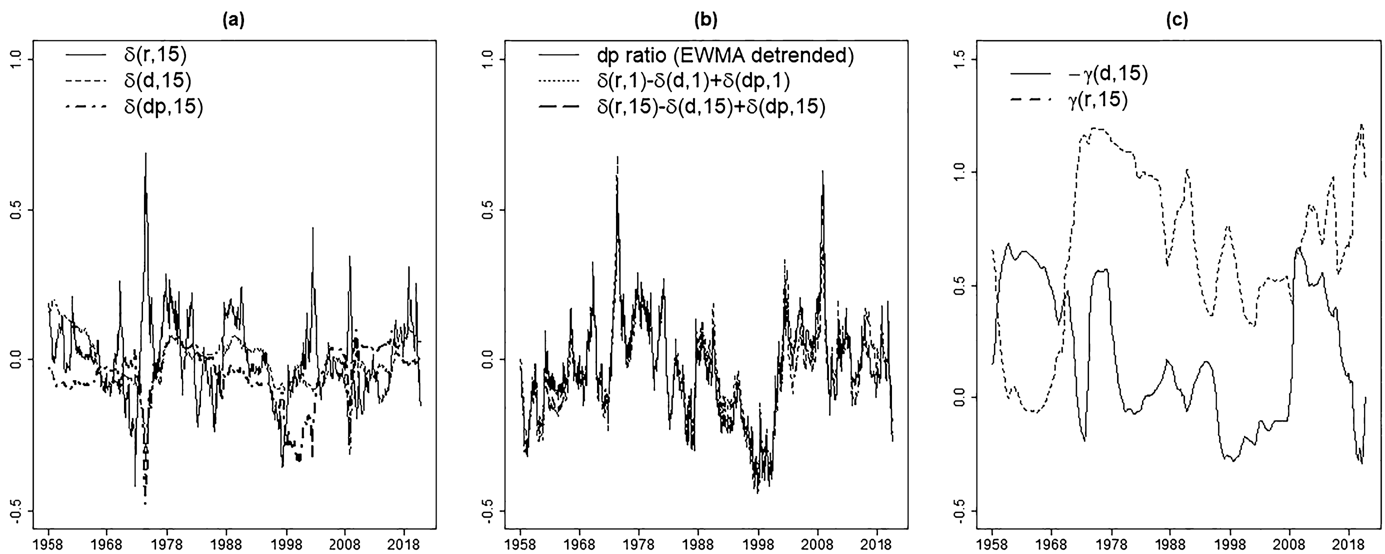


**Fig. 2.** Panel (a) shows plots of the components $\hat{\delta}_t^{(r,k)}$, $\hat{\delta}_t^{(d,k)}$, and $\hat{\delta}_t^{(dp,k)}$ (see (5)), estimated from the monthly time-varying parameter (TVP) local projections (14), using $dp_t$ as the single state variable. Panel (b) shows the observed dividend yield $dp_t$ (detrended by EWMA estimated mean) and the implied dividend yield $\widehat{dp}_t^{(k)} = \hat{\delta}_t^{(r,k)} - \hat{\delta}_t^{(d,k)} + \hat{\delta}_t^{(dp,k)}$, for $k = 12$ (1 year) and $k = 180$ (15 years). Panel (c) depicts the time-varying discount rate and cash flow components corresponding (10) when using the time-varying components (17). If not otherwise mentioned, the horizon $k = 180$ (15 years) in all the plots.

with the time-varying parameter local projections. The period 1928-1957 is omitted from the figure, due to the maximum horizon $k$ of 180 months and an initial estimation sample of an additional 180 months. The figure clearly shows stronger volatility of the expected dividend growth than with the time-invariant local projections (Fig. 1). Panel (b) plots the estimated dividend yields implied by the 12-month and 180-month local projections, as well as the actually observed dividend yield, to demonstrate that the implied dividend yields closely track the actual dividend yield. The pattern of the observed dividend yield differs from Fig. 2, because the dividend yield is now demeaned recursively using Eq. (16).

Panel (c) in Fig. 2 shows that allowing for time-varying parameters, despite considerable variation over time, does not result in a dramatic increase in the relative importance of the long-run dividend expectations. Similar to the time-invariant results, the variation in dividend expectations clearly contributes less to the dividend yield than the variation in return expectations. This is at odds with the volatility components in Panel (a) and Table 3 where the contribution of dividend expectations is substantially higher than as implied by the covariance-based volatility decomposition. In other words, while there is clear evidence of time-varying volatility of dividend growth expectations, these expectations are not always negatively correlated to the dividend yield and therefore do not systematically contribute positively to the dividend yield decomposition over time.

While the discount rate channel is dominant over the full sample, Panel (c) in Fig. 2 does reveal periods in which cash flow expectations contribute more significantly, specifically around 1960 and 2008, where the cash flow contribution approaches or even exceeds the discount rate contribution. On the other hand, there are some periods including the 1990s and the end of the sample, during which the time-varying contribution of cash flows is negative, implying a counterintuitive positive correlation between cash flow expectations $\hat{\delta}_t^{(d,k)}$ and the dividend yield. Earlier studies have found evidence of time variation in the relative contributions of expected returns and dividend growth. Chen (2009) finds a strong reversal in the pattern of return and dividend growth predictability over the period 1872–2005: While dividend growth rates are strongly predictable in the pre-war period, return predictability is more dominant in the post-war period. This result is consistent with the patterns in Fig. 2, showing a relatively high contribution of dividend expectations during the early part of the sam-

**Table 4**
Variance and dividend yield decompositions: Three state variables. This table reports the annual (Panel A) and monthly (Panel B) volatility components (12) and variance decomposition of the dividend yield (10), based on the local projections (8), using three state variables (18): $\boldsymbol{x}_t^{(a,k)} = \left(\sum_{j=1}^{k} \rho^{j-1} r_{t+j-k}, \sum_{j=1}^{k} \rho^{j-1} \Delta d_{t+j-k}, dp_t\right)$ for $a \in \{r, d, dp\}$, and different horizons $k$. The first columns report the volatility components of return, dividend growth and dividend yield predictions and ratio $\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$. The final four columns report the relative covariance contributions of expected discount rates $\widehat{\gamma}(r,k)$, cash flows $-\widehat{\gamma}(d,k)$ and forward dividend yields $\widehat{\gamma}(dp,k)$, as well as the ratio $\frac{-\widehat{\gamma}(d,k)}{\widehat{\gamma}(r,k)}$ Sample period: 1928–2020.

| $k$ | $\widehat{\sigma}(r,k)$ | $\widehat{\sigma}(d,k)$ | $\widehat{\sigma}(dp,k)$ | $\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$ | $\widehat{\gamma}(r,k)$ | $-\widehat{\gamma}(d,k)$ | $\widehat{\gamma}(dp,k)$ | $\frac{-\widehat{\gamma}(d,k)}{\widehat{\gamma}(r,k)}$ |
|---|---|---|---|---|---|---|---|---|
| (years) | | | | Panel A: Annual | | | | |
| 1 | 0.08 | 0.15 | 0.92 | 1.73 | 0.06 | 0.04 | 0.91 | 0.69 |
| 2 | 0.14 | 0.17 | 0.80 | 1.22 | 0.13 | 0.08 | 0.79 | 0.60 |
| 3 | 0.21 | 0.17 | 0.72 | 0.81 | 0.20 | 0.09 | 0.71 | 0.46 |
| 5 | 0.36 | 0.08 | 0.61 | 0.21 | 0.36 | 0.02 | 0.62 | 0.06 |
| 10 | 0.65 | 0.09 | 0.33 | 0.13 | 0.70 | -0.01 | 0.31 | -0.02 |
| 15 | 0.80 | 0.20 | 0.24 | 0.25 | 0.92 | -0.04 | 0.12 | -0.04 |
| (months) | | | | Panel B: Monthly | | | | |
| 12 | 0.08 | 0.21 | 0.91 | 2.75 | 0.07 | 0.04 | 0.89 | 0.63 |
| 24 | 0.15 | 0.23 | 0.81 | 1.52 | 0.14 | 0.07 | 0.79 | 0.47 |
| 36 | 0.21 | 0.16 | 0.73 | 0.75 | 0.21 | 0.06 | 0.73 | 0.29 |
| 60 | 0.38 | 0.04 | 0.61 | 0.11 | 0.37 | 0.00 | 0.62 | 0.00 |
| 120 | 0.66 | 0.09 | 0.33 | 0.14 | 0.74 | -0.07 | 0.33 | -0.10 |
| 180 | 0.79 | 0.19 | 0.21 | 0.25 | 0.94 | -0.06 | 0.12 | -0.07 |

ple period.[9] The high contribution of cash flow expectations during the market downturn in 2008 is emphasized by Campbell et al. (2013), who use a VAR-based return decomposition to demonstrate that this downturn was primarily driven by a decline in expected cash flows, while the earlier market downturn in 2000–2002 is almost entirely attributed to discount rate variation. Garrett and Priestley (2012) however find a more significant role for cash flow expectations during the 2000 market downturn than Campbell et al. (2013). Our findings are largely in accordance with the results by Campbell et al. (2013), with a significant cash flow contribution in 2008 and a near-zero contribution in 2000.

## 5. Multiple state variables

In this section, we extend the results in Section 4 by allowing for multiple state variables beyond the lagged dividend yield. In Section 5.1, we extend the set of state variables to include lagged cumulative returns and lagged dividend growth rates as state variables, while Section 5.2 contains a selection of specific useful predictors (state variables) for returns and dividend growth proposed in the recent literature. In Section 5.3, we apply the time-varying parameter LPs to different selections of state variables, as in Section 4.2. In Section 5.4, we apply the LASSO estimator to an even larger set of state variables and examine whether the main conclusions of Sections 5.1–5.3 are robust to the integration of even larger information sets. Finally, in Section 5.5, we examine a set of fast-moving state variables that are only available for a more recent subsample 1990–2020.

### *5.1. Three state variables*

Instead of a single state variable (dividend yield) in Section 4, in this section we move to three state variables. Lagged cumulative returns and dividend growth rates are included as additional state variables in the estimated local projections, such that the predictive power of the lagged values of all left-hand-side variables of the system (8) is now utilized:

$$\boldsymbol{x}_t^{(a,k)} = \left(\sum_{j=1}^{k} \rho^{j-1} r_{t+j-k}, \sum_{j=1}^{k} \rho^{j-1} \Delta d_{t+j-k}, dp_t\right), \qquad a \in \{r, d, dp\}. \quad (18)$$

In addition to their natural role as predictors due to the three-variable system (8), the lagged values of the LHS variables have been found to improve the theoretical performance of LP regressions (see Montiel Olea and Plagborg-Møller, 2021). Furthermore, using cumulative lagged returns and dividends brings an alternative to the latent variable approach of van Binsbergen and Koijen (2010) and subsequent studies by Zhu (2015) and Choi et al. (2017), aggregating the longer (full) history of state variables. Both of these approaches expand the information set without substantially increasing the number of parameters to be estimated, like in a VAR approach.

Table 4 reports the variance and dividend yield decompositions derived from the local projections with three state variables, using both annual and monthly data over the period 1928–2020. Both the variation of expected returns and expected dividend growth increase compared to the single-state variable case in Table 2. This increase in the volatility of all components is expected to occur because of the additional predictive power provided by lagged cumulative returns and dividend growth rates. However, the relative increase in volatility is greater for cash flow expectations than for discount rate expectations ($\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)} = 0.25$ for $k = 15$ years).

Looking at the covariance contributions (12) in the last four columns of Table 4, the relative contribution of expected cash flow growth at long horizons does not increase compared to the benchmark case of a single state variable (both annual and monthly cases) in Table 2. As already discussed in Section 4.2, the additional predictability of cash flow growth does not materialize in a substantially higher cash flow contribution relative to the discount rate channel. The cash flow contribution $-\widehat{\gamma}(d,k)$ is low across all specifications and not strictly positive.[10]

[9] Section II of the Internet Appendix reports the (constant-parameter) decompositions for the subsample 1952–2020, confirming the finding of a considerably lower contribution of long-run dividend expectations.

[10] In Section II of the Internet Appendix, we reproduce Table 4 for the shorter sample period 1952–2020, resulting in a covariance ratio of 0.04 for $k = 15$ years (0.01 with monthly data). Section III reports the VAR-implied decomposition based on a VAR estimated with annual data, which also reveals a small negative contribution of cash flows to dividend yield volatility.

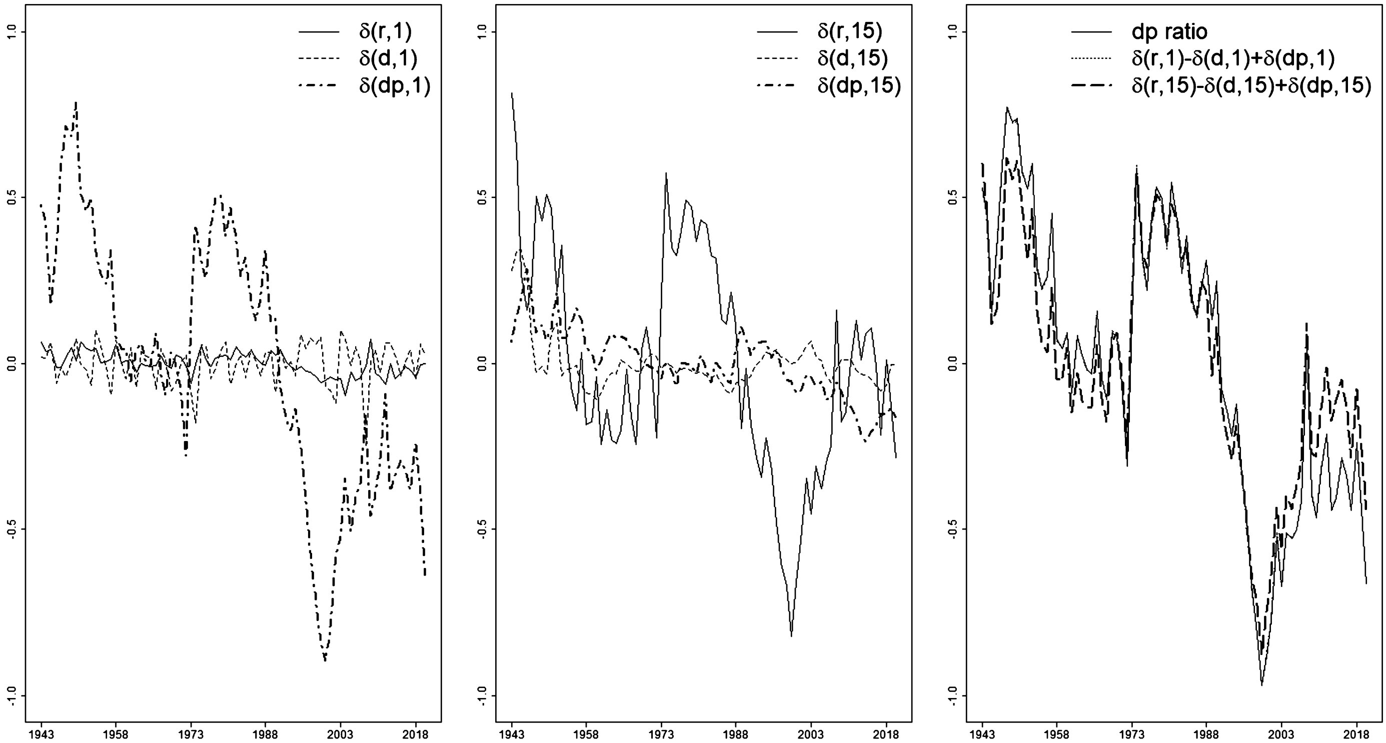


**Fig. 3.** Plots of the components $\widehat{\delta}_t^{(r,k)}$, $\widehat{\delta}_t^{(d,k)}$, and $\widehat{\delta}_t^{(dp,k)}$ (see (5)), estimated from the annual local projections (8), using three fixed state variables: $\boldsymbol{x}_t^{(a,k)} = \left(\sum_{j=1}^{k} \rho^{j-1} r_{t+j-k}, \sum_{j=1}^{k} \rho^{j-1} \Delta d_{t+j-k}, dp_t\right)$ for $a \in \{r, d, dp\}$, with $k = 1$ year (left panel) and $k = 15$ years (middle panel). The right panel shows the observed dividend yield $dp_t$ and the implied dividend yield $\widehat{\delta}_t^{(r,k)} - \widehat{\delta}_t^{(d,k)} + \widehat{\delta}_t^{(dp,k)}$, for $k = 1$ year and $k = 15$ years.

Fig. 3 plots the estimates $\widehat{\delta}_t^{(r,k)}$, $\widehat{\delta}_t^{(d,k)}$ and $\widehat{\delta}_t^{(dp,k)}$ for the case of three state variables. At short horizons, the picture is similar to Fig. 1, with both expected returns and expected dividends being fairly flat. At longer horizons, both the dividend growth and dividend yield components $\widehat{\delta}_t^{(d,15)}$ and $\widehat{\delta}_t^{(dp,15)}$ are now visibly more volatile than in Fig. 1. However, while being highly volatile, the expected cash flow component $\widehat{\delta}_t^{(d,k)}$ is not strongly correlated with the dividend yield and therefore does not contribute considerably to the dividend yield decomposition (10). The final panel of Fig. 3 shows that the approximate present value relation (5) is also accurate when the implied dividend yield is based on these multivariate local projections. The volatility and covariance-based decompositions based on three state variables reaffirm the patterns we found with time-varying parameters in the single state variable case. Expectations on future dividends are more volatile, as various predictability studies show, but this extra variation does not significantly change the decomposition of the dividend yield. This evidence is consistent with of the ‘offsetting’ hypothesis by Cochrane (2011) that incremental predictability of cash flow growth is necessarily offset by predictability of discount rates.

### 5.2. Selected state variables

In this section, we continue with the three state variables (18), supplemented with selected state variables that have been identified as strong predictors of dividend growth and/or returns in the recent literature. We consider separately state variables that are considered return predictors (to estimate $\widehat{\gamma}(r,k)$) and state variables that are considered dividend growth predictors (to estimate $-\widehat{\gamma}(d,k)$).

The first row of Panel A in Table 5 reports the estimated (covariance) contribution of expected returns to the volatility of the dividend yield ($\widehat{\gamma}(r,k)$), based on the local projections (8) using the three state variables (18), over the sample period 1952–2020, for which we have data available on all the variables. In each of the next six rows, the set of state variables is supplemented with one of the following return predictors: the price-earnings ratio with 10-year smoothed earnings ($pe10$), the term spread ($tms$), the value spread ($vs$), the earnings-price ratio ($ep$), the output gap ($gap$), and the cointegration relation between consumption, wealth and labor income ($cay$).[11] The inclusion of these state variables typically increases covariances, particularly with $pe10$ (as pointed out by Campbell and Vuolteenaho, 2004) and $cay$ (Lettau and Ludvigson, 2005). On average, the contribution of 15-year (180-month) return expectations increases slightly from 0.88 to 0.91 in the annual data (0.88 to 0.92 in the monthly data) when including one of the return predictors as an additional state variable.

In Panel B, we repeat this exercise by estimating $-\widehat{\gamma}(d,k)$, the covariance contribution of expected dividends to the dividend yield, with the three state variables (18) supplemented with one of the following dividend predictors: the dividend payout ratio ($de$), the earnings-price ratio ($ep$; Ang and Bekaert, 2007; Møller and Sander, 2017), the cointegration relation between dividends, prices and earnings ($dpe$; Garrett and Priestley, 2012), and the cointegration relation between consumption, dividends and labor income ($cdy$; Lettau and Ludvigson, 2005). Similar to our earlier results, the estimated cash flow component $-\widehat{\gamma}(d,k)$ is generally low and not uniformly positive. Among the different dividend predictors and different horizons, $dpe$ generally contributes the most in terms of increasing the importance of expected cash flows at different horizons.

Overall, the empirical results in Table 5 show again that moving beyond the baseline model, with a single state variable and constant parameters, does not substantially increase the contribution of long-run dividend expectations relative to return expectations: With annual data (1952–2020) the average ratio $\frac{\text{Ave}(d,15)}{\text{Ave}(r,15)}$ based on the model av-

[11] Variable definitions are provided in Section 3 and Section I of the Internet Appendix.

**Table 5**
Selected state variables. Panel A reports $\hat{\gamma}(r,k)$, for selected annual and monthly horizons $k$, based on the local projections (8) with the three state variables (18), supplemented with one of the state variables listed in the first column, over the sample period 1952–2020. The first row is based on the three state variables (18) alone. The final row of Panel A reports the average $\hat{\gamma}(r,k)$, denoted by Ave$(r,k)$, from the four-state-variable models. Panel B reports $-\hat{\gamma}(d,k)$, based on the local projections (8), and their average Ave$(d,k)$. The final row of the table reports the ratio $\frac{\text{Ave}(d,k)}{\text{Ave}(r,k)}$.

| | Annual | | | | | Monthly | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $x$ | $k=3$ | 6 | 9 | 12 | 15 | $k=36$ | 72 | 108 | 144 | 180 |
| | Panel A: Return predictors | | | | | | | | | |
| . | 0.26 | 0.47 | 0.67 | 0.84 | 0.88 | 0.27 | 0.48 | 0.69 | 0.83 | 0.88 |
| $pe10$ | 0.26 | 0.51 | 0.75 | 0.95 | 0.99 | 0.28 | 0.53 | 0.77 | 0.97 | 0.99 |
| $tms$ | 0.29 | 0.50 | 0.69 | 0.85 | 0.90 | 0.30 | 0.52 | 0.73 | 0.86 | 0.91 |
| $vs$ | 0.29 | 0.47 | 0.68 | 0.86 | 0.90 | 0.27 | 0.48 | 0.69 | 0.83 | 0.88 |
| $ep$ | 0.27 | 0.46 | 0.67 | 0.83 | 0.87 | 0.27 | 0.48 | 0.69 | 0.83 | 0.89 |
| $gap$ | 0.26 | 0.44 | 0.64 | 0.81 | 0.87 | 0.28 | 0.49 | 0.71 | 0.86 | 0.91 |
| $cay$ | 0.32 | 0.60 | 0.84 | 0.98 | 0.94 | 0.33 | 0.63 | 0.88 | 0.99 | 0.95 |
| Ave$(r,k)$ | 0.28 | 0.50 | 0.71 | 0.88 | 0.91 | 0.29 | 0.52 | 0.75 | 0.89 | 0.92 |
| | Panel B: Dividend predictors | | | | | | | | | |
| . | -0.01 | -0.01 | 0.02 | 0.02 | 0.04 | -0.04 | -0.03 | -0.02 | -0.01 | 0.01 |
| $de$ | -0.02 | -0.00 | 0.01 | 0.01 | 0.06 | -0.04 | -0.03 | -0.02 | -0.01 | -0.00 |
| $ep$ | -0.03 | -0.01 | 0.01 | 0.01 | 0.05 | -0.04 | -0.03 | -0.02 | -0.01 | -0.00 |
| $dpe$ | -0.02 | -0.03 | -0.04 | -0.01 | 0.18 | -0.03 | -0.05 | -0.07 | -0.04 | 0.04 |
| $cdy$ | -0.01 | 0.00 | 0.09 | 0.10 | 0.12 | -0.04 | -0.02 | 0.01 | 0.01 | 0.04 |
| Ave$(d,k)$ | -0.02 | -0.01 | 0.02 | 0.03 | 0.10 | -0.04 | -0.03 | -0.02 | -0.02 | 0.02 |
| $\frac{\text{Ave}(d,k)}{\text{Ave}(r,k)}$ | -0.07 | -0.02 | 0.02 | 0.03 | 0.11 | -0.13 | -0.06 | -0.03 | -0.02 | 0.02 |

erage of various specifications with four state variables is estimated at 0.11, compared to 0.16 and -0.04 with one and three state variables, respectively. This shows that the dividend growth predictability by selected predictors (e.g., Lettau and Ludvigson, 2005; Ang and Bekaert, 2007; Garrett and Priestley, 2012; and Møller and Sander, 2017) does not substantially change the dividend yield decomposition.

As a summary of this subsection, we are able to incorporate additional predictive information within our dividend yield decomposition due to the flexibility of local projections, which allow selecting different state variables for estimating the discount rate and cash flow expectations. The results show that dividend growth predictability (as examined in various earlier studies referred above) does not guarantee a *relative* increase in the importance of the cash flow component over the discount rate component. The results largely confirm our conclusions from Sections 4–5.1. That is, due to the correlation of expected returns and dividend growth rates, the increase in dividend predictability (or increased variance of dividend growth expectations) does not affect the relative contributions to the dividend yield decomposition.

### 5.3. Time-varying parameters

In addition to the time-invariant (constant-parameter) LPs (8), we also apply the time-varying parameter LPs (14) to the selections of multiple state variables considered in the previous subsections. In this section, we focus on the conditional (time-varying) long-run contributions (17), obtained with three state variables (Section 5.1) and three state variables supplemented with selected dividend and return predictors (Section 5.2).

The left panel of Fig. 4 depicts the time-varying contribution of expected returns $\hat{\gamma}_t(r,k)$ over time for the different selections of state variables, with $k=180$ (i.e., for a 15-year horizon). The right panel of Fig. 4 shows the corresponding time-varying cash flow contributions $-\hat{\gamma}_t(d,k)$.

The solid lines in both panels, based on one state variable, are the same as plotted in the last panel of Fig. 2. With three state variables, we can use the same sample period to construct the conditional contributions. For the volatility ratio based on selected state variables, we use the average of the estimated discount rate contribution $\hat{\gamma}_t(r,k)$ obtained using each of the six return predictors listed in Table 5, and the average of the estimated discount rate contribution $-\hat{\gamma}_t(d,k)$ for each of the four dividend predictors. Therefore, the thick dashed lines in Fig. 4 show the conditional (time-varying) equivalent of the unconditional average contributions Ave$(r,k)$ and Ave$(d,k)$ in Table 5. As Fig. 4 shows, the conditional contributions based on selected state variables start later due to the availability of several variables only after the year 1952. The first estimates of the conditional contributions based on selected state variables are for 1982, due to the horizon $k$ and an initial estimation window of 180 months each.

The estimated conditional contributions vary considerably over time. The return contributions in the left panel are generally larger than the cash flow contributions in the right panel, reaffirming our earlier finding that the discount rate contribution generally dominates. However, as we noted when discussing the time-varying contributions in Fig. 2, there are clearly periods during which the cash flow contributions are considerably higher, while there are also periods during which the cash flow contribution is negative. Our time-varying volatility decomposition somewhat contradicts the 'tug-of-war' hypothesis by Zhu (2015), in which either dividends or returns are predictable. Overall, Fig. 4 indicates that both the discount rate and cash flow components matter with a varying degree of importance throughout the sample period.

The patterns of the different conditional return contributions in the left panel of Fig. 4 are rather similar, which may result from the dividend yield as a single state variable already capturing a significant part of the return predictability. Adding additional state variables does not have a large impact on the estimated return expectations ($\hat{\delta}_t^{(r,k)}$) and therefore on its (time-varying) contribution to the dividend yield decomposition. For the cash flow expectations, however, changing the set of state variables has a more profound impact on the dynamics of the estimated cash flow contribution, as can be seen from the relative dispersion between the three lines in the right panel of Fig. 4.

### 5.4. Large set of state variables and LASSO

As already surveyed in the Introduction, a vast literature compiles evidence that various variables besides the dividend yield predict stock returns as well as dividend growth rates at different frequencies and

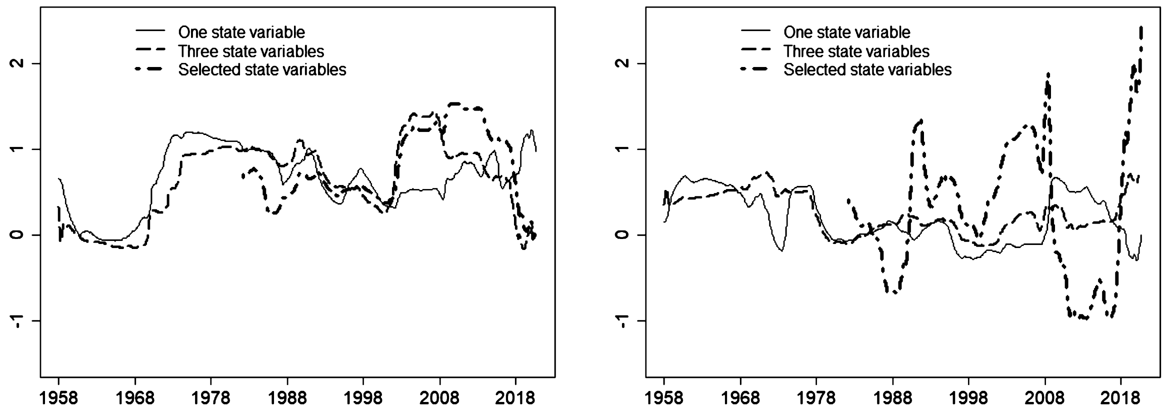


**Fig. 4.** Plots of $\widehat{\gamma}_t(r,k)$ (left) and $-\widehat{\gamma}_t(d,k)$ (right) obtained from (17), with $\widehat{\delta}_t^{(r,k)}$ and $-\widehat{\delta}_t^{(d,k)}$ estimated by EWLS from rolling window local projections (14), with $k = 180$, using three different information sets: *(i)* one state variable, *(ii)* three state variables, and *(iii)* three state variables supplemented with selected dividend and return predictors. The sample period is 1928–2020 for *(i)* and *(ii)* and 1952–2020 for *(iii)*. The figures start 30 years later, due to the horizon $k$ and initial estimation window of 180 months each.

forecast horizons. To accommodate the integration of such large sets of state variables, while keeping concerns on possible overfitting to a minimum, we apply common data-driven methodologies from the machine and statistical learning literature. In this section, we concentrate on LASSO (Least Absolute Shrinkage and Selection Operator) to examine whether the main conclusions in Sections 5.1–5.3 are robust to a larger information set, while Section V of the Internet Appendix contains similar results obtained with the elastic net approach, having potentially some advantages over LASSO for highly correlated state variables.

The LASSO estimator, as popularized by Tibshirani (1996), is an alternative to the OLS estimator where the idea in short is to select the optimal state variables, by shrinking the parameters of irrelevant state variables to zero, without taking a prior standpoint on which variables should be included in LPs. This shrinkage (penalization)-based method allows us to consider a potentially large number of state variables simultaneously. In our context, this means that the LASSO estimator will select the state variables separately for all three components in (5) and for all horizons $k$, without causing excessive computational burden.

The LASSO estimator for parameters $\boldsymbol{\varphi}^{(a,k)} = (\alpha^{(a,k)}, \boldsymbol{\beta}^{(a,k)})$, $a \in \{r, d, dp\}$, is defined as:

$$\widehat{\boldsymbol{\varphi}}_{LASSO}^{(a,k)} = \underset{\boldsymbol{\varphi}^{(a,k)}}{\arg\min} \left\{ \frac{1}{2T} \sum_{t=1}^{T} \left( LHS(a) - \alpha^{(a,k)} - \boldsymbol{x}_t^{(a,k)} \boldsymbol{\beta}^{(a,k)} \right)^2 + \lambda \sum_{j=1}^{n_a} \left| \beta_j^{(a,k)} \right| \right\}, \quad (19)$$

where $LHS(a)$ is one of the three left hand side variables in (8) and $T$ is the number of observations in the estimation sample (depending also on the horizon $k$ in this notation). The number of candidate state variables is denoted by $n_a$. Intuitively, the aim of the LASSO estimator is to find a set of coefficient estimates that lead to the smallest residual sum of squares, subject to the constraint set by the penalty term $\sum_{j=1}^{n_a} |\beta_j^{(a,k)}|$. The amount of shrinkage is controlled by the tuning parameter $\lambda$: Increasing $\lambda$ results in a greater shrinkage towards zero in the coefficients $\widehat{\beta}_j^{(a,k)}$. We follow, e.g., Medeiros and Mendes (2016) and Medeiros and Vasconcelos (2016), who recommend in a time series context to determine $\lambda$ by the Bayesian information criterion (BIC), as opposed to the cross-validation typically used in cross-sectional LASSO analyses. For a sufficiently large value of $\lambda$, the LASSO estimator shrinks some $\widehat{\beta}_j^{(a,k)}$ exactly to zero (e.g., Hastie et al., 2009, Section 3.4), thus performing parameter estimation and model selection at the same time. This is effective and in practice highly useful in our context by automatically producing the required horizon-specific local projections where the optimal state variables are selected depending on the horizon $k$. As a result, the estimated local projections generated from LASSO are 'sparse' and likely circumvent overfitting since only a subset of the full set of potential state variables is involved.[12]

In the LASSO estimation, we do not apply shrinkage to the dividend yield (i.e., the penalty term in (19) does not include the regression coefficient related to the lagged dividend yield), due to its essential role in the benchmark decompositions (see, e.g., Engsted et al., 2012). This implies that $dp_t$ is always included in the resulting local projections.[13]

Table 6 presents the annual and monthly volatility decompositions using local projections estimated by LASSO. Even if LASSO exploits larger information sets, the resulting decompositions are not very different from those obtained with three or selected state variables (Tables 4 and 5). Our main result that, despite the increased predictability of dividend growth rates, the expected cash flow contribution remains low and is occasionally even slightly negative, is robust to expansion of the predictive information set.

Table 7 presents still detailed LASSO regression results, i.e., the estimated local projections at selected forecast horizons that are used for obtaining the annual volatility decompositions reported in Table 6. These results include information on both the variable selection (i.e., the inclusion and exclusion of state variables), and the estimated LASSO coefficients (see (19)). First of all, we can clearly see that the selected LASSO local projections are indeed horizon-specific so that different state variables are valuable for different horizons. For returns, the best predictors in terms of systematic inclusions, along with the dividend yield, are term spread ($tms$), realized volatility ($svar$), investment-to-capital ratio ($ik$), consumer price inflation ($infl$), value spread ($vs$), price-10 year earnings ratio ($pe10$) and the consumption-wealth ratio ($cay$). As reviewed in the Introduction, these are largely the variables that are expected to have predictive power for returns (see, e.g., Fama and French, 1989; Lettau and Ludvigson, 2001, 2005; and Martin, 2017, and as also emphasized in Section 5.2). On the contrary, earnings yield ($ep$), short-term interest rates ($tbl$ and $rfree$) and long-term corporate bond rates ($corpr$) are typically excluded. For the dividend growth lo-

[12] All computations in this paper are carried out in R (RStudio). Specifically, the LASSO-based local projections are constructed with the glmnet package and BIC-based tuning parameter $\lambda$ selection (see, e.g., Medeiros and Vasconcelos (2016) and the modification of ic.glmnet function available in the HDeconometrics package. See https://rdrr.io/github/gabrielrvsc/HDeconometrics/man/ic.glmnet.html. As robustness checks, we also determine the tuning parameter $\lambda$ by cross-validation and apply the elastic net approach, as reported in Section V of the Internet Appendix.

[13] Section V of the Internet Appendix presents annual and monthly results where the dividend yield is also subject to shrinkage. It turns out that the dividend yield is typically included in all local projections and hence the resulting volatility decompositions are very close to those reported in Table 6.

**Table 6**
Variance and dividend yield decompositions: LASSO. This table reports the annual (Panel A) and monthly (Panel B) volatility components (12) and variance decomposition of the dividend yield (10), based on the local projections (8), using LASSO, for different horizons $k$. The first columns report the volatility components of return, dividend growth and dividend yield predictions and ratio $\frac{\hat{\sigma}(d,k)}{\hat{\sigma}(r,k)}$. The final four columns report the relative covariance contributions of expected discount rates $\hat{\gamma}(r,k)$, cash flows $-\hat{\gamma}(d,k)$ and forward dividend yields $\hat{\gamma}(dp,k)$, as well as the ratio $\frac{-\hat{\gamma}(d,k)}{\hat{\gamma}(r,k)}$. Sample period: 1952–2020.

| $k$ | $\hat{\sigma}(r,k)$ | $\hat{\sigma}(d,k)$ | $\hat{\sigma}(dp,k)$ | $\frac{\hat{\sigma}(d,k)}{\hat{\sigma}(r,k)}$ | $\hat{\gamma}(r,k)$ | $-\hat{\gamma}(d,k)$ | $\hat{\gamma}(dp,k)$ | $\frac{-\hat{\gamma}(d,k)}{\hat{\gamma}(r,k)}$ |
|---|---|---|---|---|---|---|---|---|
| (years) | | | | Panel A: Annual | | | | |
| 1 | 0.15 | 0.10 | 0.86 | 0.65 | 0.13 | 0.02 | 0.84 | 0.17 |
| 2 | 0.24 | 0.21 | 0.77 | 0.88 | 0.23 | 0.02 | 0.75 | 0.08 |
| 3 | 0.31 | 0.23 | 0.75 | 0.74 | 0.26 | 0.02 | 0.72 | 0.07 |
| 5 | 0.51 | 0.29 | 0.64 | 0.56 | 0.42 | 0.03 | 0.54 | 0.08 |
| 10 | 0.95 | 0.34 | 0.71 | 0.35 | 0.84 | 0.06 | 0.11 | 0.07 |
| 15 | 0.96 | 0.45 | 0.55 | 0.47 | 0.94 | 0.18 | -0.12 | 0.19 |
| (months) | | | | Panel B: Monthly | | | | |
| 12 | 0.27 | 0.22 | 0.90 | 0.83 | 0.14 | 0.02 | 0.84 | 0.16 |
| 24 | 0.43 | 0.23 | 0.85 | 0.55 | 0.25 | -0.00 | 0.75 | -0.00 |
| 36 | 0.49 | 0.32 | 0.84 | 0.65 | 0.29 | -0.03 | 0.74 | -0.10 |
| 60 | 0.69 | 0.35 | 0.81 | 0.51 | 0.50 | -0.06 | 0.55 | -0.11 |
| 120 | 0.92 | 0.36 | 0.72 | 0.40 | 0.79 | -0.05 | 0.25 | -0.06 |
| 180 | 1.04 | 0.42 | 0.57 | 0.40 | 1.01 | 0.03 | -0.05 | 0.03 |

**Table 7**
LASSO model selection and estimation results – Annual data. This table reports the results of the LASSO model selection and parameter estimates (see (19)), for selected horizons $k$. Empty cells indicate the exclusion of those state variables. The candidate state variables are described in Table 1 and $r$ and $d$ denote the (discounted) cumulative lagged returns and dividend growth rates (see (18)). The results are reported for local projections with two different left hand side (LHS) variables: cumulative returns (left panel); and cumulative dividend growth (right panel). Sample period: 1952–2020. Full results, including the LPs with the dividend yield as LHS variable, for all horizons $k$, using both annual and monthly data are reported in Internet Appendix Section V.

| | LHS: returns | | | | | LHS: dividend growth | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $k$ (years) | 3 | 6 | 9 | 12 | 15 | 3 | 6 | 9 | 12 | 15 |
| const | 1.09 | 1.58 | 4.31 | 5.59 | 2.79 | 0.05 | -0.24 | 1.12 | 4.43 | 1.46 |
| $r$ | | | | | -0.12 | -0.05 | | -0.10 | -0.08 | |
| $d$ | | -0.16 | -0.78 | | | -0.16 | -0.30 | -0.44 | -0.54 | -0.28 |
| $dp$ | 0.23 | 0.33 | 0.30 | 0.29 | -0.12 | -0.13 | -0.27 | 0.01 | -0.30 | -0.00 |
| $ep$ | | | | | -0.18 | | | | 0.88 | |
| $de$ | | 0.08 | 0.33 | 0.08 | | | -0.14 | | -0.81 | -0.18 |
| $svar$ | | 0.23 | 1.68 | 1.92 | -0.44 | | 0.89 | | -0.02 | |
| $bm$ | | | | -0.58 | 0.78 | 0.31 | 0.59 | | -0.66 | |
| $ntis$ | | | -2.30 | | -4.38 | 1.05 | | 0.32 | 0.67 | 0.30 |
| $tbl$ | | | | | | | -1.03 | | -1.78 | -0.87 |
| $lty$ | | | 3.96 | 6.16 | 2.16 | | | | -3.61 | -1.23 |
| $Rfree$ | | -1.89 | | -3.19 | | -1.76 | | | | |
| $tms$ | | 3.86 | | 2.01 | 1.38 | 1.19 | | | | |
| $ltr$ | | | 0.08 | 0.23 | | | | | | |
| $dfy$ | | | | -8.21 | -14.11 | | | 2.62 | | |
| $dfr$ | | -0.82 | -0.75 | -1.43 | -0.17 | 0.24 | 0.30 | 0.03 | 0.40 | |
| $corpr$ | | | | | | 0.11 | | 0.09 | 0.10 | |
| $infl$ | | 4.00 | 1.01 | 1.72 | 0.13 | | | | | |
| $cay$ | 1.09 | 3.97 | 2.61 | -2.14 | 2.03 | | -1.20 | -3.43 | -4.15 | -4.17 |
| $eqis$ | | | -0.18 | -0.47 | -0.09 | -0.13 | 0.14 | 0.33 | 0.24 | |
| $ik$ | | | -45.73 | -23.24 | -25.73 | | -6.09 | -5.17 | 3.63 | |
| $cdy$ | | | | 3.94 | 1.06 | | 0.28 | | 0.81 | |
| $gap$ | -0.59 | -1.40 | | | | | | | | |
| $vs$ | 0.10 | | 0.03 | 0.03 | 0.08 | 0.07 | | -0.08 | -0.03 | |
| $pe10$ | | | -0.31 | -0.93 | -0.60 | -0.11 | -0.05 | | -0.23 | |
| $dpe$ | | 0.34 | 0.45 | | | -0.19 | -0.36 | -0.30 | -0.84 | |

cal projections, $cay$ also appears to be an important predictor, as well as lagged dividends, value spread ($vs$), net equity expansion ($ntis$) and $dpe$.[14]

[14] Section V of the Internet Appendix reports the full estimation results, including the LPs with the dividend yield as LHS variable at all horizons $k$, estimated with both annual and monthly data.

### 5.5. Option-implied state variables

As a final empirical analysis, we consider four state variables that are recently proposed in the literature: the conditional variance of stock returns ($cv$) and the equity variance premium ($vp$) proposed by Bekaert and Hoerova (2014), and the risk-aversion index ($ra$) and uncertainty index ($unc$) proposed by Bekaert et al. (2022). These four variables are

**Table 8**
Option-implied state variables. This table reports $\hat{\gamma}(r,k)$ and $-\hat{\gamma}(d,k)$ for selected monthly horizons $k$, based on the local projections (8) with the three state variables (18), supplemented with one of the option-implied state variables listed in the first column. The first row is based on the three state variables (18) alone. The final row reports the averages of $\hat{\gamma}(r,k)$ and $-\hat{\gamma}(d,k)$, from the four-state-variable models. Sample period 1990:1–2020:12.

| | $\hat{\gamma}(r,k)$ | | | | | $-\hat{\gamma}(d,k)$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $x$ | $k=12$ | 24 | 36 | 60 | 120 | $k=12$ | 24 | 36 | 60 | 120 |
| . | 0.21 | 0.41 | 0.55 | 0.88 | 1.22 | 0.13 | 0.13 | 0.11 | 0.16 | 0.19 |
| $cv$ | 0.22 | 0.43 | 0.48 | 0.79 | 0.87 | 0.40 | 0.39 | 0.28 | 0.09 | 0.20 |
| $vp$ | 0.30 | 0.54 | 0.57 | 0.78 | 0.93 | 0.43 | 0.39 | 0.28 | 0.13 | 0.16 |
| $ra$ | 0.25 | 0.50 | 0.52 | 0.79 | 0.84 | 0.45 | 0.39 | 0.23 | 0.02 | 0.14 |
| $unc$ | 0.08 | 0.24 | 0.24 | 0.60 | 0.79 | 0.37 | 0.35 | 0.22 | 0.02 | 0.22 |
| Ave | 0.21 | 0.43 | 0.45 | 0.74 | 0.86 | 0.41 | 0.38 | 0.25 | 0.06 | 0.18 |

derived from option market data (and other financial variables), which is why we refer to these as "option-implied state variables". These variables are generally mean-reverting at shorter horizons than the more traditional predictors, such as valuation ratios, and are thus meant to capture higher frequency variation in expected discounted rates and cash flows.

We apply our local projections (8) combining each of these predictors with the three benchmark state variables (18), such that expected discount rates and cash flows have both a short-term and long-term component. Table 8 reports the contribution of expected discount rates ($\hat{\gamma}(r,k)$) and cash flow growth ($-\hat{\gamma}(d,k)$) for selected monthly horizons $k$. Due to the shorter data availability (1990–2020) of these option-implied variables, we focus on monthly data only and exclude the 15-year horizon from our analysis, instead focusing on (monthly) horizons of 1, 2, 3, 5 and 10 years.

Particularly at shorter horizons, the additional state variables increase the contribution of expected dividends. At the 12-month horizon, adding the risk aversion index ($ra$) to the model more than triples $-\hat{\gamma}(d,12)$ from 0.13 to 0.45. For expected returns, the increase is less sizeable: from 0.21 to 0.30 when the variance premium ($vp$) is added to the model. At longer horizons, the incremental impact of these state variables is modest: the covariance ratio $\frac{-\hat{\gamma}(d,120)}{\hat{\gamma}(r,120)}$ at the 10-year horizon is 0.16 for the benchmark model without additional state variables and 0.21 using the average across the four extended specifications. The limited impact at longer horizons is as expected due to the swiftly mean-reverting nature of these specific variables.

## 6. Discussion

The main contribution of this study is the introduction of flexible local projections to examine the relative importance of the cash flow and discount rate components to the variation in the dividend yield. Our main empirical result is that the (time-varying) contribution of expected cash flow growth is not completely negligible but clearly secondary compared to the contribution of expected discount rates. In the baseline case of a single state variable (the dividend yield) and constant parameters, we find that expected dividend growth is relatively flat and does not contribute much to the volatility of the dividend yield. When we expand the information set to contain multiple state variables, we find more evidence of variation of expected cash flows, but this does not show up in the dividend yield decomposition where expected discount rates remain the dominant component. When we estimate the local projections recursively (i.e., allowing for time-varying parameters), we find that expected dividend growth rates are at times more volatile than in other time periods, and even become temporarily the dominant component of market volatility, but discount rate expectations remain the primary driver of volatility for most of the sample period.

Our results provide a new methodological perspective to the puzzling 'stylized fact' that dividend growth is not predictable by the dividend yield in the U.S. during the postwar period. As documented by Engsted and Pedersen (2010), this finding does not hold in general in international equity markets, while Chen (2009) and Golez and Koudijs (2018) do find dividend growth predictability in the U.S. prior to 1945. As emphasized by Menzly et al. (2004), Lettau and Ludvigson (2005), van Binsbergen and Koijen (2010), and Møller and Sander (2017), it is possible for dividend predictors to have offsetting effects on the dividend yield, thereby obfuscating the predictability of dividends by the dividend yield. Indeed, various studies have found evidence of cash flow predictability by other factors than the dividend yield (e.g., Lettau and Ludvigson, 2005; Ang and Bekaert, 2007; Larrain and Yogo, 2008; Garrett and Priestley, 2012; Møller and Sander, 2017). The flexibility of local projections allows this dividend predictability to be utilized when obtaining the dividend yield volatility decomposition by integrating additional state variables beyond the dividend yield in the projections of both returns and dividends. When expanding the set of state variables beyond the dividend yield, the incremental predictive power increases both the volatility of expected dividend growth and of expected returns. However, despite the increased predictability of dividend growth, the contribution of dividend growth to the decomposition of the dividend yield is hardly affected, due to the weak covariance between expected dividend growth and the dividend yield. Surprisingly, the correlation between expected dividend growth and the dividend yield is for some specifications and (short) time periods even slightly positive, resulting in a negative cash flow contribution to the covariance-based decomposition. Therefore, we conclude that predictability of dividend growth is a necessary but not sufficient condition for a high cash flow contribution to the covariance-based dividend yield decomposition.

While the main focus of our analysis is on long-term horizons, our results in Section 5.5 show that recently proposed state variables such as the equity variance premium (Bekaert and Hoerova, 2014) and the risk aversion index (Bekaert et al., 2022) have an impact on the volatility decomposition at 1-2 year horizons. Building upon the local projection approach by combining high and low persistence state variables such that expected returns and dividends have both short-term and long-term elements may be a useful direction for future research.

In addition to including additional state variables, local projections allow us to investigate the time-varying nature of the dividend yield decomposition. Time-varying patterns are surprisingly similar across different choices of state variables. This suggests that there are alternating periods of relatively strong dividend and return predictability (or 'pockets of predictability', cf. Farmer et al., 2023). Within these periods, the choice of state variables appears to be of less importance. The idea of alternating periods of predictability at first glance resembles the 'tug-of-war' hypothesis by Zhu (2015), in which returns and dividends take turns between regimes of high and low predictability. However, our results do not imply this type of binary and exclusive regime-switching process for the dividend yield decomposition. In fact, we find evidence of both dividend and return predictability across the full sample, with the relative importance of discount rate and cash flow variation being relatively smooth and stable over time.

Chen et al. (2012) attribute the apparent lack of dividend growth predictability by dividend yields in the postwar period to dividend smoothing, causing the dividend yield to be uninformative of cash flows in the near future, but not necessarily implying that future cash flows are truly unpredictable. By relying on direct (i.e., LP-based) long-run predictions, as opposed to iterated (i.e., VAR-based) short-run predictions, we circumvent the diminished predictability caused by short-run dividend smoothing and its potential disruptive impact on the dividend yield decomposition.[15] The time-varying nature of dividend growth predictability is in fact consistent with dividend growth being subject to time-varying payout policies, affected by factors including dividend smoothing but also time-varying investor demand for dividends (see, e.g., Baker and Wurgler, 2004; Chen et al., 2012; Larkin et al., 2017).

## 7. Conclusions

We specify horizon-specific local projections to identify the relative contributions of expected discount rates and expected cash flows to the variation of the dividend yield. Building upon the well-known vector autoregressive (VAR) approach, we apply our local projection approach to develop a flexible dividend yield decomposition. In addition to robustness to model misspecification, local projections provide various advantages over the VAR approach. Despite strong seasonalities in dividend payments, we are able to accommodate monthly data in addition to annual data. The enlarged sample size resulting from the use of monthly data allows us to apply recursive estimation to examine time variation in the dividend yield decomposition. Furthermore, in addition to the standard dividend yield, we are able to incorporate several different predictors of both dividend growth and returns.

Our results generally confirm that variation in expected discount rates is the dominant component of observed dividend yield volatility. The cash flow component is also present, but its relative importance does not increase substantially for any choice of state variables. While we find, consistent with earlier literature, evidence of more volatile dividend expectations when we move beyond the lagged dividend yield as the single state variable, this incremental predictability of cash flows does not covary strongly with the dividend yield and therefore does not contribute to its decomposition.

Only when allowing for time variation in the dividend yield decomposition, we find that in certain periods, such as around the 2008 financial crisis, the contribution of expected cash flows to the decomposition of the dividend yield increases temporarily. Nevertheless, over the full sample, the contribution of expected discount rates remains clearly dominant. Various alternative specifications and robustness checks reported throughout this paper and in the Internet Appendix point to the same main conclusion: circumventing the econometric restrictions implied by the VAR approach results in increased predictability of dividend growth, but does not lead to a significantly larger cash flow contribution to the decomposition of dividend yield volatility.

## CRediT authorship contribution statement

**Matthijs Lof:** Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **Henri Nyberg:** Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing.

## Data availability

Data will be made available on request.

[15] Section VI in the Internet Appendix presents a simple simulation exercise to demonstrate that the LP approach is indeed less sensitive to dividend smoothing than the VAR approach.

## Appendix A. VAR-based approaches

In this Appendix, we briefly outline the vector autoregressive (VAR) approaches implemented by Campbell and Shiller (1988b) and Cochrane (2008). In contrast to the local projection approach that we employ, both of these VAR-based approaches are built upon the assumption that the multivariate system containing stock return ($r_t$), dividend growth rate ($\Delta d_t$) and the dividend yield ($dp_t$) follows a VAR representation, from which the long-run contributions of expected dividend growth rates ($E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j}$) and expected discount rates ($E_t \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}$) can be derived. Due to the linear structure of the VAR, Campbell and Shiller (1988b) and Cochrane (2008) derive closed form expressions of these long-run predictions.

### *A.1. Campbell and Shiller (1988b)*

Starting from the long-run identity (7) with infinite horizon ($k \longrightarrow \infty$), Campbell and Shiller (1988b) attempt to estimate the component associated with expected dividend growth $\left(\delta_t^{(d,\infty)} = E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j}\right)$ by fitting a bivariate VAR to the annual price-dividend ratio and the annual dividend growth rate (both measured in logs):

$$\mathbf{z}_t \equiv \begin{bmatrix} pd_t \\ \triangle d_t \end{bmatrix} = \mathbf{A}\mathbf{z}_{t-1} + \boldsymbol{\varepsilon}_t. \tag{A.1}$$

For ease of exposition, we assume a VAR structure with only one lag (VAR(1)) below, but as Campbell and Shiller (1988b) show, the framework can be straightforwardly adapted to a more general VAR(p) structure. Also note that Campbell and Shiller model the price-dividend ratio, while we and others model the dividend-price ratio. Due to the logarithmic transformation and the linear structure of the models, this choice has no impact on the final results since $pd_t = -dp_t$. The matrix of estimated parameters $\mathbf{A}$ in (A.1) and the calibrated parameter $\rho$ (Eq. (3)) can be used to recover the conditional expectations $E_t \Delta d_{t+j}$, and to compute a time series of the VAR-implied dividend growth variable $\delta_t^{(d,\infty)}$:

$$\delta_t^{(d,\infty)} = E_t \sum_{i=0}^{\infty} \rho^i \triangle d_{t+1+i} = \sum_{i=0}^{\infty} \rho^i \left(\mathbf{e}_2' \mathbf{A}^i \mathbf{z}_t\right) = \mathbf{e}_2' \mathbf{A} \left(\mathbf{I} - \rho \mathbf{A}\right)^{-1} \mathbf{z}_t, \tag{A.2}$$

in which $\mathbf{e}_2$ is a vector of zeros in which the second element is replaced by one. A full derivation is provided by Campbell and Shiller (1988b). The constructed variable $\delta_t^{(d,\infty)}$ can be thought of as a 'theoretical PD ratio' that should closely trace the observed PD ratio, if expected discount rates would be constant (i.e., if all variation in the PD ratio is due to expected cash flow variation). Campbell and Shiller report the ratio $\frac{\mathrm{Std}(\delta_t^{(d,\infty)})}{\mathrm{Std}(pd_t)}$, which is clearly closely related to our measure $\sigma(d,k)$ in (12). The main difference is that, instead of obtaining long-run predictions by iterating forward a one-period VAR, we obtain these predictions with horizon-specific direct regressions (8) at different horizons $k$.

### *A.2. Cochrane (2008)*

Cochrane (2008) fits a first-order VAR system to the annual returns, dividend growth rates and dividend yields:

$$\begin{aligned} r_{t+1} &= c^{(r)} + b^{(r)} dp_t + \varepsilon_{t+1}^{(r)} \\ \Delta d_{t+1} &= c^{(d)} + b^{(d)} dp_t + \varepsilon_{t+1}^{(d)} \\ dp_{t+1} &= c^{(dp)} + b^{(dp)} dp_t + \varepsilon_{t+1}^{(dp)}, \end{aligned} \tag{A.3}$$

where the lagged dividend yield is the only state variable. As Cochrane (2008) shows, the (approximate) log-linear present-value identity (1) implies the following link between the VAR coefficients of (A.3):

$$b^{(r)} = 1 - \rho\, b^{(dp)} + b^{(d)}, \tag{A.4}$$

which also leads to links between the error terms by $\varepsilon^{(r)}_{t+1} = \varepsilon^{(d)}_{t+1} + \rho\, \varepsilon^{(dp)}_{t+1}$. The system of three equations (A.3) is thus overidentified: The regression coefficients and the error term of any of the three equations are implied by the other two.

Dividing the identity (A.4) by $1 - \rho\, b^{(dp)}$ yields the long-run coefficients of returns ($b^{(r,lr)}$) and dividend growth ($b^{(d,lr)}$):

$$b^{(r,lr)} - b^{(d,lr)} = \frac{b^{(r)}}{1 - \rho\, b^{(dp)}} - \frac{b^{(d)}}{1 - \rho\, b^{(dp)}} = 1. \tag{A.5}$$

As Cochrane (2008) demonstrates, the coefficients $b^{(r,lr)}$ and $b^{(d,lr)}$ can be interpreted as the slope coefficients of hypothetically regressing long-run cumulative discounted returns $\left(\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}\right)$ and dividend growth $\left(\sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j}\right)$ on the dividend yield $dp_t$:

$$\begin{aligned} \widehat{b}^{(r,lr)} &= \frac{\mathrm{Cov}\left(\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, dp_t\right)}{\mathrm{Var}(dp_t)} \quad \text{and} \\ \widehat{b}^{(d,lr)} &= \frac{\mathrm{Cov}\left(\sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j}, dp_t\right)}{\mathrm{Var}(dp_t)}. \end{aligned} \tag{A.6}$$

The fitted values of these hypothetical regressions thus correspond to the fitted values of our local projections (9), in the special case of the dividend yield as the only predictor and the infinite horizon ($k \longrightarrow \infty$):

$$\begin{aligned} \widehat{\delta}^{(r,\infty)}_t &= \widehat{c}^{(r,lr)} + \widehat{b}^{(r,lr)} dp_t \\ \widehat{\delta}^{(d,\infty)}_t &= \widehat{c}^{(d,lr)} + \widehat{b}^{(d,lr)} dp_t. \end{aligned} \tag{A.7}$$

From (A.7), it is easy to see that volatility components (12) are related to the implied long-run coefficients by Cochrane (2008):

$$\begin{aligned} \widehat{\sigma}(a,\infty) &= \sqrt{\frac{\mathrm{Var}(\widehat{\delta}^{a,\infty}_t)}{\mathrm{Var}(pd_t)}} = \sqrt{\frac{\mathrm{Var}(\widehat{b}^{(a,lr)} dp_t)}{\mathrm{Var}(dp_t)}} = \left|\widehat{b}^{(a,lr)}\right| \sqrt{\frac{\mathrm{Var}(dp_t)}{\mathrm{Var}(dp_t)}} \\ &= \left|\widehat{b}^{(a,lr)}\right|, \end{aligned} \tag{A.8}$$

for $a \in \{r, d\}$. Moreover, concerning the covariance components (10) (as $\widehat{dp}_t = dp_t$, which holds in this VAR case with $k = \infty$), equations (A.7) imply

$$\mathrm{Cov}\left(dp_t, \widehat{\delta}^{(r,\infty)}_t\right) = \widehat{b}^{(r,lr)} \quad \text{and} \quad \mathrm{Cov}\left(dp_t, \widehat{\delta}^{(d,\infty)}_t\right) = \widehat{b}^{(d,lr)}. \tag{A.9}$$

In other words, in this one-variable case there is one-to-one correspondence between the long-run coefficients in (A.7) and the information in our volatility decomposition (10).

## Appendix B. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jbankfin.2024.107127.

## References


Ang, A., Bekaert, G., 2007. Return predictability: is it there? Rev. Financ. Stud. 20, 651–707.
Baker, M., Wurgler, J., 2004. A catering theory of dividends. J. Finance 59 (3), 1125–1165.
Bekaert, G., Hoerova, M., 2014. The VIX, the variance premium and stock market volatility. J. Econom. 183 (2), 181–192.
Bekaert, G., Engstrom, E.C., Xu, N.R., 2022. The time variation in risk appetite and uncertainty. Manag. Sci. 68 (6), 3975–4004.
Bollerslev, T., Tauchen, G., Zhou, H., 2009. Expected stock returns and variance risk premia. Rev. Financ. Stud. 22 (11), 4463–4492.
Campbell, J.Y., 1991. A variance decomposition for stock returns. Econ. J. 101, 157–179.
Campbell, J.Y., Ammer, J., 1993. What moves the stock and bond markets? A variance decomposition for long-term asset returns. J. Finance 48, 3–37.
Campbell, J.Y., Shiller, R.J., 1988a. Stock prices, earnings, and expected dividends. J. Finance 43, 661–676.
Campbell, J.Y., Shiller, R.J., 1988b. The dividend-price ratio and expectations of future dividends and discount factors. Rev. Financ. Stud. 1, 195–228.
Campbell, J.Y., Vuolteenaho, T., 2004. Bad beta, good beta. Am. Econ. Rev. 94, 1249–1275.
Campbell, J.Y., Giglio, S., Polk, C., 2013. Hard times. Rev. Asset Pricing Stud. 3 (1), 95–132.
Chen, L., 2009. On the reversal of return and dividend growth predictability: a tale of two periods. J. Financ. Econ. 92, 128–151.
Chen, L., Zhao, X., 2009. Return decomposition. Rev. Financ. Stud. 22, 5213–5249.
Chen, L., Da, Z., Priestley, R., 2012. Dividend smoothing and predictability. Manag. Sci. 58 (10), 1834–1853.
Chen, L., Da, Z., Zhao, X., 2013. What drives stock price movements? Rev. Financ. Stud. 26, 841–876.
Chevillon, D., 2007. Direct multi-step estimation and forecasting. J. Econ. Surv. 21, 746–785.
Chevillon, G., 2016. Multistep forecasting in the presence of location shifts. Int. J. Forecast. 32, 121–137.
Choi, K.H., Kim, C-J., Park, C., 2017. Regime shifts in price-dividend ratios and expected stock returns: a present-value approach. J. Money Credit Bank. 49, 417–441.
Cochrane, J.H., 1992. Explaining the variance of price-dividend ratios. Rev. Financ. Stud. 5, 243–280.
Cochrane, J.H., 2008. The dog that did not bark: a defense of return predictability. Rev. Financ. Stud. 21, 1533–1575.
Cochrane, J.H., 2011. Presidential address: discount rates. J. Finance 66, 1047–1108.
Cochrane, J.H., 2017. Macro-finance. Rev. Finance 21, 945–985.
Cochrane, J.H., Piazzesi, M., 2002. The fed and interest rates: a high-frequency identification. Am. Econ. Rev. 92, 90–95.
Cooper, I., Priestley, R., 2009. Time-varying risk premiums and the output gap. Rev. Financ. Stud. 22, 2801–2833.
Cujean, J., Hasler, M., 2017. Why does return predictability concentrate in bad times? J. Finance 72, 2717–2758.
Dangl, T., Halling, M., 2012. Predictive regressions with time-varying coefficients. J. Financ. Econ. 106, 157–181.
De la O, R., Myers, S., 2021. Subjective cash flow and discount rate expectations. J. Finance 76, 1339–1387.
Engsted, T., Pedersen, T.Q., 2010. The dividend-price ratio does predict dividend growth: international evidence. J. Empir. Finance 17, 585–605.
Engsted, T., Pedersen, T.Q., Tanggaard, C., 2012. Pitfalls in VAR based return decomposition: a clarification. J. Bank. Finance 36, 1255–1265.
Fama, E.F., French, K.R., 1988. Dividend yields and expected stock returns. J. Financ. Econ. 22, 3–25.
Fama, E.F., French, K.R., 1989. Business conditions and expected returns on stocks and bonds. J. Financ. Econ. 25, 23–49.
Farmer, L.E., Schmidt, L., Timmermann, A., 2023. Pockets of predictability. J. Finance 78, 1279–1341.
Garrett, I., Priestley, R., 2012. Dividend growth, cash flow, and discount rate news. J. Financ. Quant. Anal. 47 (5), 1003–1028.
Golez, B., 2014. Expected returns and dividend growth rates implied by derivative markets. Rev. Financ. Stud. 27 (3), 790–822.
Golez, B., Koudijs, P., 2018. Four centuries of return predictability. J. Financ. Econ. 127, 248–263.
Gorodnichenko, Y., Lee, B., 2020. Forecast error variance decompositions with local projections. J. Bus. Econ. Stat. 38 (4), 921–933.
Gu, S., Kelly, B., Xiu, D., 2020. Empirical asset pricing via machine learning. Rev. Financ. Stud. 33, 2223–2273.
Guo, H., 2006. On the out-of-sample predictability of stock market returns. J. Bus. 79, 645–670.
Hallerbach, W.G., Menkveld, A.J., 2004. Analysing perceived downside risk: the component value-at-risk framework. Eur. Financ. Manag. 10 (4), 567–591.
Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer-Verlag, New York.
Henkel, S.J., Martin, J., Nardari, F., 2011. Time-varying short-horizon predictability. J. Financ. Econ. 99, 560–580.
Jordà, Ò., 2005. Estimation and inference of impulse responses by local projections. Am. Econ. Rev. 95 (1), 161–182.
Kofman, P., McGlenchy, P., 2005. Structurally sound dynamic index futures hedging. J. Futures Mark. 25 (12), 1173–1202.
Koijen, R.S.J., Van Nieuwerburgh, S., 2011. Predictability of returns and cash flows. Annu. Rev. Financ. Econ. 3, 467–491.
Lamont, O., 1998. Earnings and expected returns. J. Finance 53, 1563–1587.
Larkin, Y., Leary, M.T., Michaely, R., 2017. Do investors value dividend-smoothing stocks differently? Manag. Sci. 63 (12), 4114–4136.
Larrain, B., Yogo, M., 2008. Does firm value move too much to be justified by subsequent changes in cash flow? J. Financ. Econ. 87 (1), 200–226.
LeRoy, S.F., Porter, R.D., 1981. The present-value relation: tests based on implied variance bounds. Econometrica 49 (3), 555–574.

Lettau, M., Ludvigson, S., 2001. Consumption, aggregate wealth, and expected stock returns. J. Finance 56, 815–849.
Lettau, M., Ludvigson, S., 2005. Expected returns and expected dividend growth. J. Financ. Econ. 76, 583–626.
Lettau, M., Van Nieuwerburgh, S., 2008. Reconciling the return predictability evidence. Rev. Financ. Stud. 21, 1607–1652.
Li, D., Plagborg-Møller, M., Wolf, C.K., 2022. Local projections vs. VARs: Lessons from thousands of DGPs. NBER Working Paper 30207.
Lintner, J., 1956. Distribution of incomes of corporations among dividends, retained earnings, and taxes. Am. Econ. Rev. 46, 97–113.
Longerstaey, J., Spencer, M., 1996. Riskmetrics Technical Document. J.P. Morgan/Reuters, New York.
Maio, P., Santa-Clara, P., 2015. Dividend yields, dividend growth, and return predictability in the cross section of stocks. J. Financ. Quant. Anal. 50 (1–2), 33–60.
Marcellino, M., Stock, J.H., Watson, M.W., 2006. A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. J. Econom. 135 (1–2), 499–526.
Martin, I., 2017. What is the expected return on the market? Q. J. Econ. 132, 367–433.
Medeiros, M.C., Mendes, E., 2016. l1-regularization of high-dimensional time-series models with non-Gaussian and heteroskedastic errors. J. Econom. 191, 255–271.
Medeiros, M.C., Vasconcelos, G.F.R., 2016. Forecasting macroeconomic variables in data-rich environments. Econ. Lett. 138, 50–52.
Menzly, L., Santos, T., Veronesi, P., 2004. Understanding predictability. J. Polit. Econ. 112, 1–47.
Møller, S.V., Sander, M., 2017. Dividends, earnings and predictability. J. Bank. Finance 78, 153–163.
Montiel Olea, J.L., Plagborg-Møller, M., 2021. Local projection inference is simpler and more robust than you think. Econometrica 89 (4), 1789–1823.
Pettenuzzo, D., Sabbatucci, R., Timmermann, A., 2020. Cash flow news and stock price dynamics. J. Finance 75 (4), 2221–2270.
Plagborg-Møller, M., Wolf, C.K., 2021. Local projections and VARs estimate the same impulse responses. Econometrica 89 (2), 955–980.
Pohl, W., Schmedders, K., Wilms, O., 2018. Higher order effects in asset pricing models with long-run risks. J. Finance 73 (3), 1061–1111.
Rangvid, J., Schmeling, M., Schrimpf, A., 2014. Dividend predictability around the world. J. Financ. Quant. Anal. 49, 1255–1277.
Rapach, D.E., Zhou, G., 2013. Forecasting stock returns. In: Elliott, G., Timmermann, A. (Eds.), Handbook of Economic Forecasting, vol. 2A. North-Holland, pp. 329–383.
Rapach, D.E., Strauss, J.K., Zhou, G., 2010. Out-of-sample equity premium prediction: combination forecasts and links to the real economy. Rev. Financ. Stud. 23, 821–862.
Schorfheide, F., 2005. VAR forecasting under misspecification. J. Econom. 128, 99–136.
Shiller, R.J., 1981. Do stock prices move too much to be justified by subsequent changes in dividends? Am. Econ. Rev. 71 (3), 421–436.
Taylor, J.W., 2008. Using exponentially weighted quantile regression to estimate value at risk and expected shortfall. J. Financ. Econom. 6 (3), 382–406.
Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Ser. B 58, 267–288.
Timmermann, A., 2008. Elusive return predictability. Int. J. Forecast. 24, 1–18.
van Binsbergen, J.H., Koijen, R.S.J., 2010. Predictive regressions: a present-value approach. J. Finance 65, 1439–1471.
Welch, I., Goyal, A., 2008. A comprehensive look at the empirical performance of equity premium prediction. Rev. Financ. Stud. 21, 1455–1508.
Zhu, X., 2015. Tug-of-war: time-varying predictability of stock returns and dividend growth. Rev. Finance 19, 2317–2358.
Zhu, X., Zhu, J., 2013. Predicting stock returns: a regime-switching combination approach and economic links. J. Bank. Finance 37, 4120–4133.