Discount Rates and Cash Flows: A Local Projection Approach *

Matthijs Lof †   Henri Nyberg ‡

September 18, 2019

Abstract

We develop a volatility decomposition derived from flexible and robust 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 to consider time variation in the volatility decomposition. While the variation of expected discount rates remains the dominant contributor to market volatility, we find that the contribution of expected cash flows is non-negligible when moving beyond the standard model with the dividend yield as the single state variable.

Keywords: volatility decomposition, dividend growth, local projections, LASSO
JEL classification: C32, G12

* We thank Zhi Da, Jari-Pekka Heinonen, Òscar Jordà, Samu Kärkkäinen, Matti Suominen, and conference and seminar participants at the EEA-ESEM 2019 (Manchester), Finnish Economic Association 2019 (Oulu), Nordic Econometric Meeting 2019 (Stockholm), Aalto University, University of Helsinki, University of Nottingham, and the University of Turku for useful discussions. We gratefully acknowledge the Finnish Foundation for Share Promotion and (Nyberg) the Emil Aaltonen Foundation and the Academy of Finland (grant 321968) for financial support.

† Aalto University School of Business. P.O.Box 21220, FI-00076 Aalto, Finland. matthijs.lof@aalto.fi. +358 504014448.

‡ University of Turku, Department of Mathematic and Statistics; and Tampere University, Faculty of Management and Business. FI-20014 University of Turku. Finland. henri.nyberg@utu.fi. +358294504322

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 classic 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 for recent surveys).

To analyze the discount rate vs. cash flow conundrum, the 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, (log) 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 the VAR, we obtain the required predictions for the discounted (cumulative) expected returns and dividend growth rates as well as the discounted dividend yield using (forecast) horizon-specific single-equation regressions, which we refer to, following Jordà (2005), as 'local projections'.

These flexible and information-rich local projections, which facilitate modern data-driven machine (statistical) learning methods as a part of the analysis, enable us to reconcile the mounting evidence of both return and dividend predictability and the prior VAR-based volatility decompositions implying that dividend expectations are flat (i.e., they do not contribute to market volatility). That is, we are able to integrate vast past research effort emphasizing the role of various other state variables besides the dividend yield as predictors of returns at different horizons.[1] Similarly for dividend growth, Lettau and Ludvigson (2005) emphasize the role of consumption ratios at longer horizons, while Ang and Bekaert (2007) and Møller and Sander (2017) find the earnings yield as a useful predictor of future cash flows. Moreover, Engsted and Pedersen (2010) and Rangvid, Schmeling and Schrimpf (2014) find international evidence in favor of predictability of dividend growth by the dividend yield and other variables. In line with these findings, our main empirical result is that the estimated contribution of long-run cash flow expectations to market volatility is considerably larger than estimates based on the VAR approach (e.g. Cochrane, 2008) and recent latent variable approaches (see van Binsbergen and Koijen, 2010; Zhu, 2015; and Choi, Kim and Park, 2017).

In his seminal work, Jordà (2005) proposes local projections as an alternative to VARs for computing macroeconomic impulse response functions. We apply local projections in a different context: to infer long-run expectations of discount rates and cash flow growth of interest for the dividend yield volatility 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. Following Jordà (2005), this is more robust to misspecification than the VAR approach, which is built upon the strong assumption that the underlying VAR representation is correctly specified. In practice, however, the estimated VAR, like any econometric model, is likely to be misspecified, providing at best an approximation to the true correct asset pricing process. As Jordà (2005) puts it, misspecification errors are hence '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, and are thus generally more robust to misspecification.[2]

[1] These variables include, e.g., valuation ratios (Campbell and Shiller, 1988a; Fama and French, 1988; Lamont, 1998), interest rates and interest rate spreads (Ang and Bekaert, 2007; Fama and French, 1989), stock market volatility (Guo, 2006; Martin, 2017) and the consumption-wealth ratio (Lettau and Ludvigson, 2001; 2005), to name just a few. See Koijen and Van Nieuwerburgh (2011) and Rapach and Zhou (2013) for recent surveys. Several studies have nevertheless also questioned the predictive power of the dividend yield and other valuation ratios for returns, even at long horizons (see, e.g., Ang and Bekaert, 2007; Welch and Goyal, 2008; and Boudoukh, Richardson and Whitelaw, 2008).

[2] Local projections have recently gained popularity as a tool for structural inference in macroeconomic applications. See, e.g., Owyang, Ramey, and Zubairy, 2013; Ramey, 2016; Gorodnichenko and Lee, 2017; and Ramey and

A clear example of a situation where extrapolating short-run predictions does not provide optimal long-run predictions is the corporate policy of dividend smoothing. Short-run dividend smoothing adversely affects the predictability of dividends in the short run. As Chen, Da, and Priestley (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. Since 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, Da, and Priestley (2012), by making direct (as opposed to VAR-implied) predictions of long-run dividend growth.

In addition to minimizing misspecification concerns, raised partly by the above vastly varying findings in return and dividend growth predictability research, the use of local projections has a number of clear-cut advantages compared to the VAR approach. First, local projections enable the incorporation of potentially large sets of economic and financial state variables. We apply advanced model averaging and statistical learning-based methods to facilitate large information sets, while avoiding dimensionality concerns. Second, in contrast to the existing VAR and latent variable based decompositions with a fixed infinite horizon, our volatility decomposition can be estimated at short, intermediate and long-run horizons, which allows for selecting the set of predictive variables at each horizon of interest separately. Third, we show that local projections 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). Fourth, due to the enlarged sample sizes resulting from the use of monthly data, we are able to evaluate possibly important time variation of the volatility decomposition by estimating the local projections recursively with time-varying estimation windows. These advantages, with the resulting main empirical implications, are briefly outlined below.

In our empirical analysis, we start by applying local projections with the lagged dividend yield as a single state variable and find that the contribution to market volatility of expected cash flow growth is indeed marginal compared to expected discount rates, which is consistent

Zubairy, 2018. Cochrane and Piazzesi (2002) provide an early example of impulse response functions constructed by direct regressions. The robustness of local projections in terms of potential misspecification is also supported by comparisons between 'direct' and 'iterative' multiperiod forecasting methods (see, e.g., Marcellino, Stock, and Watson, 2006; and Chevillon, 2007). To the best our knowledge, this is the first study to systematically integrate local projections in an empirical asset pricing application.

with the findings by Cochrane (2008, 2011). We then first extend the information set by including lagged (cumulative) returns and dividend growth rates along with the lagged dividend yield, and find increased predictability of both returns and dividend growth. Importantly, the incremental predictive power obtained from increasing the information set is particularly significant when predicting dividend growth. As a result, we find that the contribution of expected cash flow (dividend) growth increases considerably. In other words, even though discount rate variation remains the primary component of market volatility, we find the role of expected dividends to be far from negligible when incorporating predictive information beyond the lagged dividend yield.

We proceed by extending the information set to include a broader set of potential state variables. As the aforementioned list of potential state variables is long, we offer two specific data-driven solutions facilitated by local projections in a data-rich environment: Model averaging and LASSO (Least Absolute Shrinkage and Selection Operator). Both approaches can be seen as shrinkage (i.e., penalization-based) methods to obtain guard against overfitting. Rapach, Strauss, and Zhou (2010) are among the first ones to document strong evidence on the superiority of model averaging in (out-of-sample) return forecasting by taking a simple average of predictive regression models containing a single predictor (see also Timmermann, 2006; and Rapach and Zhou, 2013). This ultimately stems from the highly uncertain, complex and potentially continuously evolving underlying data generating processes for expected returns, dividend growth rates and the dividend yield, which are difficult to approximate with a single and relatively parsimonious model such as a VAR. Model averaging reduces this uncertainty and instability risk associated with reliance on a single model and importantly circumvents overfitting despite the use of large predictor datasets.

The LASSO is a machine learning method popularized by Tibshirani (1996) that performs variable selection (from a potentially large set of predictive variables) and parameter estimation simultaneously to enhance the prediction accuracy and interpretability of the econometric model it produces. LASSO is generally more robust than alternative approaches to variable selection and parameter estimation such as backward or forward stepwise regressions. To the best of our knowledge, this is the first study integrating the LASSO, or any modern machine (statistical) learning-based method, within the context of assessing the relative importance of discount rate and cash flow expectations to market volatility. Overall, the resulting volatility decompositions from model averages and LASSO estimation yield very similar conclusions as the above-mentioned case of three state variable (i.e., lagged returns, dividend growth and dividend yield): expected discounted rates dominate but the cash flow component is also very much present.

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 typically analyze monthly data. As pointed out, local projections enable us to establish an approximative model to incorporate dividend growth also at monthly frequency, despite the seasonality of monthly dividends data, by including monthly updated 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. The empirical results of the full sample (time-invariant) and in particular the time-varying decompositions at the monthly frequency provide additional robustness to the finding of a nonzero impact of expected cash flows. In the time-varying decomposition, expected cash flows are at times even the dominant component, relative to the discount rate component.

The remainder of this paper is organized as follows. Section 2 presents the methodological advancement around local projections and over the previous volatility decomposition studies. These include model averaging and LASSO-based local projections and the approximate monthly approach. Section 3 presents the main empirical results. We provide a discussion of our findings in Section 4 and Section 5 concludes. Appendix A provides detailed descriptions of prior volatility decomposition methods. Supplementary empirical 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:

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

where $r_t \equiv \log(R_t)$, $dp_t \equiv \log\left(\frac{D_t}{P_t}\right)$, 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 κ can be omitted from the model:

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

where ρ is required to be below, but close to 1. Empirically, ρ is typically estimated as

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

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:

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

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

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

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. Quantifying the relative magnitudes of these three components at different horizons k is the key objective of this paper.

In the existing literature, researchers almost solely 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:

$$\begin{aligned} dp_t &= E_t \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} \quad - \quad E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} \\ &\equiv \qquad \delta_t^{(r,\infty)} \qquad - \qquad \delta_t^{(d,\infty)}. \end{aligned} \tag{6}$$

The level of 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, as emphasized by Cochrane (2008), that observing variation in the dividend yield implies that either future returns or dividends, or both, are 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 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, however, also has several disadvantages. First, it assumes that the estimated VAR is the correct data generating process for all three components in (5) at all horizons k, while in reality the VAR parameters are noisy due to estimation and possible misspecification errors. Iterating the VAR-based predictions forward therefore likely leads to poor estimates of the (long-run) expected components in (5). The direct regressions (local projections) that we propose in the next section are specified for each horizon separately, thereby minimizing misspecification bias. Second, VARs have limited capacity to incorporate large sets of state variables since the number of parameters increases quadratically in the number of variables. Our local projections allow for the implied expected returns and expected dividend growth to depend on different state variables, selected even locally for each horizon of interest, and from a large set of potential state variables. Third, VARs are restrictive also in a sense that they do not allow the 'mixed-frequency' matching between annual and monthly data. Within the local projection approach, we are able to include monthly data, even if dividends are measured over rolling twelve-month windows. The use of monthly data increases the number of observations, which is particularly useful when investigating potential time variation in the parameters.

2.2 Local projections and volatility decomposition

In this section, we introduce a novel volatility decomposition that is built upon flexible local projections 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 (hereafter often LPs) in structural econometric inference originates from the work of Jordà (2005). At the heart of this approach lie forecast horizon-specific predictions of the three components of interest in the identity (5). These predictions form the basis for a flexible dividend yield volatility decomposition 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:

$$\begin{aligned}
\sum_{j=1}^{k} \rho^{j-1} r_{t+j} &= \alpha^{(r,k)} + \boldsymbol{x}_t^{(r,k)} \boldsymbol{\beta}^{(r,k)} + \varepsilon_{t+k}^{(r,k)} \\
\sum_{j=1}^{k} \rho^{j-1} \Delta d_{t+j} &= \alpha^{(d,k)} + \boldsymbol{x}_t^{(d,k)} \boldsymbol{\beta}^{(d,k)} + \varepsilon_{t+k}^{(d,k)} \\
\rho^k dp_{t+k} &= \alpha^{(dp,k)} + \boldsymbol{x}_t^{(dp,k)} \boldsymbol{\beta}^{(dp,k)} + \varepsilon_{t+k}^{(dp,k)},
\end{aligned} \tag{7}$$

where $\boldsymbol{x}_t^{(a,k)}$ and $\varepsilon_{t+k}^{(a,k)}$, $a \in \{r, d, dp\}$, are the vectors of state variables (or predictors) and zero-mean error terms, respectively. Therefore, due to the linear structure of (7), each equation can be consistently estimated by ordinary least squares (OLS) under general conditions. The conditional expectations (or fitted values) of the left-hand-side (LHS) variables in (7), 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):

$$\widehat{\delta}_t^{(a,k)} = \widehat{\alpha}^{(a,k)} + \boldsymbol{x}_t^{(a,k)} \widehat{\boldsymbol{\beta}}^{(a,k)} \qquad a \in \{r, d, dp\}. \tag{8}$$

Due to the flexible structure of LPs, the resulting estimates $\widehat{\delta}_t^{(a,k)}$ from (7)–(8) are expected to be more informative and less prone to misspecification error than the estimates obtained from the VAR approaches discussed in Section 2.1 and Appendix A. Importantly, if it in fact turns out that the VAR is 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 the construction of $\delta_t^{(a,k)}$ when using the LPs instead of the VAR-based approaches.

In addition to the VAR-based approach, Cochrane (2008, 2011) also obtains the volatility decomposition using 'direct regressions' with the dividend yield as the only state variable. This can be seen as a restricted case of the local projections (7), where $\boldsymbol{x}_t^{(a,k)} = dp_t$ for all k and for $a \in \{r, d, dp\}$. Cochrane's contribution is the starting point of our analysis: we estimate the components $\delta_t^{(r,k)}$, $\delta_t^{(d,k)}$, and $\delta_t^{(dp,k)}$ by fitting the regressions (7) using the dividend yield dp_t as the single state variable.

There is no a priori reason to assume that the dividend yield should be the only relevant predictor of long-run dividends and returns. Lettau and Ludvigson (2005) demonstrate that, even if identity (6) 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 of use in predicting long run returns and dividend growth. We therefore proceed by estimating the same set of regressions using not only lagged dividend yields, but also lagged cumulative returns and lagged dividend growth rates as state variables. That is:

$$\boldsymbol{x}_t^{(a,k)} = \left(\sum_{j=1}^{k} \rho^{j-1} r_{t+j-k}, \sum_{j=1}^{k} \rho^{j-1} \Delta d_{t+j-k}, dp_t \right) \qquad a \in \{r, d, dp\}, \tag{9}$$

indicating that the predictive power of lags of all left-hand-side variables of the system (7) are now utilized. 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. In addition to these three state variables, we also explore a larger set of financial and economic variables. To incorporate large sets of potential variables, we estimate the LPs (7) using a LASSO approach and model averaging, on which we provide more details in Section 2.3.

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 $\delta_t^{(dp,k)}$ provide the necessary ingredients to form our dividend yield volatility decomposition. Taking variances of the present-value identity (5) gives:

$$\begin{aligned} \mathrm{Var}\left(dp_t\right) \quad = \quad & \mathrm{Var}\left(\delta_t^{(r,k)}\right) + \mathrm{Var}\left(\delta_t^{(d,k)}\right) + \mathrm{Var}\left(\delta_t^{(dp,k)}\right) \\ & -2\left[\mathrm{Cov}\left(\delta_t^{(r,k)}, \delta_t^{(d,k)}\right) - \mathrm{Cov}\left(\delta_t^{(r,k)}, \delta_t^{(dp,k)}\right) + \mathrm{Cov}\left(\delta_t^{(d,k)}, \delta_t^{(dp,k)}\right)\right]. \end{aligned} \tag{10}$$

The estimated relative contributions of the first three components to the variance of the dividend yield are:

$$\frac{\mathrm{Var}(\widehat{\delta_t}^{(r,k)})}{\mathrm{Var}(dp_t)}, \quad \frac{\mathrm{Var}(\widehat{\delta_t}^{(d,k)})}{\mathrm{Var}(dp_t)}, \quad \frac{\mathrm{Var}(\widehat{\delta_t}^{(dp,k)})}{\mathrm{Var}(dp_t)}. \tag{11}$$

As equation (10) suggest, the relative variance contributions (11) do not generally sum up to one due to neglected covariance terms in (10). Following the prior literature (e.g., Cochrane, 2008, 2011; van Binsbergen and Koijen, 2010; Zhu, 2015; and Choi, Kim, and Park, 2017), we are primarily interested in the relative variance terms (11) rather than the covariance terms. For ease of interpretation and to allow explicit comparison with the VAR-based decompositions as considered by Campbell and Shiller (1988b) and Cochrane (2008) (see details in Appendix A), we report the square roots of the variance ratios (11). This results in our volatility decomposition:

$$\widehat{\sigma}(r,k) = \frac{\mathrm{Std}(\widehat{\delta_t}^{(r,k)})}{\mathrm{Std}(dp_t)}, \quad \widehat{\sigma}(d,k) = \frac{\mathrm{Std}(\widehat{\delta_t}^{(d,k)})}{\mathrm{Std}(dp_t)}, \quad \widehat{\sigma}(dp,k) = \frac{\mathrm{Std}(\widehat{\delta_t}^{(dp,k)})}{\mathrm{Std}(dp_t)}, \tag{12}$$

where k is the forecast horizon of interest as in (5). In the remainder of this paper, we use these three measures ($\widehat{\sigma}(r,k), \widehat{\sigma}(d,k)$, and $\widehat{\sigma}(dp,k)$) to quantify the relative contributions to the volatility of the dividend yield generated by expected discount rates, expected cash flow growth, and the expected future dividend yield, respectively.

Ultimately, we are primarily interested in measuring the relative importance of expected discount rate and cash flow variation:

$$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)} = \frac{\mathrm{Std}(\widehat{\delta_t}^{(d,k)})}{\mathrm{Std}(\widehat{\delta_t}^{(r,k)})}. \tag{13}$$

If the state variables in (7) do not hold explanatory power on dividend growth, the fitted values $\widehat{\delta_t}^{(d,k)}$ will be essentially constant over time, such that the ratio (13) is close to zero. When moving from one to multiple state variables, it is expected that the volatility of both $\widehat{\delta_t}^{(d,k)}$ and $\widehat{\delta_t}^{(r,k)}$ will increase, because of the improved fit of the local projections. However, the relative sensitivity of the volatility of expected dividend growth and expected returns to different information sets, and therefore the behavior of the ratio (13), remains an open question that we aim to answer in this paper.

2.3 Data-rich local projections: Model averaging 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 forecast horizons. To accommodate the integration of a large set of predictors while keeping concerns on possible overfitting to a minimum, we apply two common data-driven methodologies from the machine and statistical learning literature: model averaging and the LASSO (Least Absolute Shrinkage and Selection Operator).

In this study, similar to the frequentist model averaging approach by Rapach, Strauss, and Zhou (2010), the model averages of expected returns, dividend growth rates and dividend yields are constructed at each period as the equal-weighted averages of the predictions obtained from the local projections (7) with different sets of state variables. Formally this can be written as:

$$\widehat{\delta}_t^{(a,k)} = \frac{1}{n_a}\sum_{j=1}^{n_a}\left(\widehat{\alpha}_j^{(a,k)} + \boldsymbol{x}_{jt}^{(a,k)}\widehat{\boldsymbol{\beta}}_j^{(a,k)}\right), \quad a \in \{r, d, dp\}, \tag{14}$$

where k is the horizon, $\widehat{\alpha}_j^{(a,k)}$ and $\widehat{\boldsymbol{\beta}}_j^{(a,k)}$ are the OLS estimates from the predictor-specific LPs generated by the state variables included in $\boldsymbol{x}_{jt}^{(i,k)}$. The number of relevant candidate state variables is denoted by n_a and described more detail in Section 3.1. In our empirical analysis, $\boldsymbol{x}_{jt}^{(i,k)}$ includes a fixed set of pre-selected variables (e.g., the dividend yield) supplemented by one additional predictor (indexed by j) at the time. Giving equal weight to the n_a predictions is in accordance with prior studies generally showing that equal-weighted model averages (forecast combinations) typically outperform more complicated alternatives (see, e.g., Timmermann, 2006; Rapach, Strauss, and Zhou, 2010; and Baetje, 2018).

The other method that we use to accommodate a data-rich information set is the LASSO. The LASSO estimator, as popularized by Tibshirani (1996), is an alternative to the usual 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. This shrinkage (or 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 $(k \geq 1)$

$$\widehat{\boldsymbol{\varphi}}_{LASSO}^{(a,k)} = \underset{\boldsymbol{\varphi}^{(a,k)}}{\arg\min}\left\{\frac{1}{2T}\sum_{t=1}^{T}\left(LHS(a) - \alpha^{(a,k)} - \boldsymbol{x}_t^{(a,k)}\boldsymbol{\beta}^{(a,k)}\right)^2 + \lambda\sum_{j=1}^{n_a}\left|\beta_j^{(a,k)}\right|\right\}, \tag{15}$$

where $LHS(a)$ is one of the three left hand side variables in (7), T is the number of observations in the estimation sample (depending also on the horizon k), and all the other notations are the same as above in (14). The essential difference to (14) is that now all n_a candidate predictors, for which the shrinkage is set to apply, are initially included simultaneously in $\boldsymbol{x}_t^{(a,k)}$ (i.e.: $\boldsymbol{x}_t^{(a,k)}$ is an $1 \times n_a$ vector) whereas the predictors are considered one by one in the model averaging method. 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 $\sum_{j=1}^{n_a} |\beta_j^{(a,k)}|$.[3]

The amount of shrinkage is controlled by the tuning parameter λ: Increasing λ results in greater shrinkage toward zero in $\widehat{\beta}_j^{(a,k)}$. We follow Medeiros and Mendes (2016) and Medeiros and Vasconcelos (2016), who recommend in a time-series context to determine λ 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 λ, the LASSO estimator shrinks some $\widehat{\beta}_j^{(a,k)}$ exactly to zero (e.g., Hastie, Tibshirani, and Friedman, 2009, Section 3.4), performing thus parameter estimation and model selection at the same time. This is effective and in practice highly useful in our context, producing automatically 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 the LASSO are 'sparse' and circumvent overfitting since only a subset of the full set of potential state variables is involved.[4]

2.4 Monthly local projections

Prior studies on dividend yield volatility 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, Kim, and Park, 2017). Since dividend payments are highly seasonal, dividend growth rates are often considered informative only on an annualized basis.[5] One of our contributions is that local projections allow for the use of monthly (or even higher-

[3] For simplicity, the notation for state variables $\boldsymbol{x}_t^{(a,k)}$ is the same throughout this study. Following common practice, we standardize the predictors within the construction of the LASSO estimator (15) when determining the penalty function, but for the construction of the LPs (Eq. 7–8) the original variables are used and the OLS estimator is just replaced by the LASSO estimator (15).

[4] All computations in this paper are carried out in R. Specifically, the LASSO-based local projections are constructed with the glmnet package and BIC-based tuning parameter λ selection (see, e.g., Medeiros and Vasconcelos, 2016).

[5] Closely related return decomposition studies do use monthly data, but compared to (7), 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; and Chen and Zhao, 2009).

frequency) data explicitly when modelling the component $\delta_t^{(d,k)}$ associated with the expected dividend growth in (5). From an empirical point of view, the introduction of a volatility decomposition based on monthly data implies substantially more observations and hence greater statistical accuracy than the annual case. This provides additional robustness and enables various extensions to the annual decompositions, such as examining potential time variation and time-varying parameters in the discount rate and cash flow dynamics (see Section 3.5).

Instead of the annual frequency, as implicitly assumed in the previous sections, we now let the time index t be monthly. The essential challenge at the monthly frequency is the measurement of dividends. Even though dividend data are available at the monthly frequency, these figures have strong seasonalities, resulting in highly erratic behaviour of the monthly dividend yield (dp_t) and dividend growth rate (Δd_t). Therefore, 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, return series are simply monthly returns. Section 3.1 below 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 by groups of 12 months:

$$\sum_{j=1}^{k} \rho^{j-1} \Delta d_{t+j} \quad \approx \quad \sum_{j=1}^{k} \rho_j^* \Delta^* d_{t+j} = \rho_{12} \Delta^* d_{t+12} + \rho_{24} \Delta^* d_{t+24} + \cdots \, , \tag{16}$$

where $\rho_j^* \equiv \rho^{j-1}$ if $j \in \{12, 24, 36, ...\}$ and 0 otherwise, and $\Delta^* d_{t+j} \equiv d_{t+j} - d_{t+j-12}$ is the 12-month cumulative dividend growth. Since $\Delta^* d_t = \sum_{i=0}^{11} \Delta d_{t-i}$, the approximation is exact if $\rho = 1$ and holds closely when the difference between ρ^{j-1} and ρ^{j-12} is small, which is the case when ρ is close to one, as assumed in (3) and estimated hereafter in this study. Thus, to obtain the monthly volatility decomposition of the dividend yield, we replace the (unobserved) monthly dividend growth component in identity (5) and the local projections (7) by the right hand side of (16), which is based on monthly observations of 12-month cumulative dividend growth rates.

The approximation (16) implies that the (monthly) forecast horizon k needs to be a multiple of 12 (i.e., $k \in \{12, 24, 36, ...\}$). In other words, despite basing the analysis on the monthly frequency data, we report the monthly decompositions only for annualized horizons. Nevertheless, although we are considering annualized horizons, the local projection approach enables us to use updated data in every month, thereby greatly increasing the number of observations and therefore statistical power. The increased number of observations in turn allows us to study time variation of the volatility decomposition.

3 Empirical results

This section first describes the data and then presents the estimated volatility decompositions based on local projections with different sets of state variables.

3.1 Data

For our monthly analysis, our main variables are the monthly value-weighted market returns reported by the Center for Research in Security Prices (CRSP), the dividend-price ratio (dividend yield), and the 12-month dividend growth rate. Following Cochrane (2008), the annualized log dividend yield in each month is computed as follows:

$$dp_t = \log\left(\frac{R_{t-11:t}}{Rx_{t-11:t}} - 1\right) \tag{17}$$

where $R_{t-11:t}$ refers to the cumulative gross CRSP value-weighted market return over the 12-month period ending in month t, and Rx_t refers to the cumulative gross CRSP value-weighted market return over the same period *excluding* dividends.

Following Koijen and van Nieuwerburgh (2011), among others, we compute monthly dividends by $D_t = (R_t - Rx_t)P_{t-1}$. To avoid seasonality concerns, monthly dividends are compounded over 12 months to compute 12-month (log) growth rates, as discussed in Section 2.4. Monthly dividends are compounded 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 re-invested 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.[6]

[6] For robustness, we report in Section IV of the the Internet Appendix our main results computed using market-reinvested and non-reinvested dividends. In addition, we also consider S&P 500 returns instead of CRSP market returns. Overall, these variations of the data lead to qualitatively similar results.

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

Table 1: A list of additional candidate state variables

List of candidate state variables obtained from the datasets of Welch and Goyal (2008) and Rapach and Zhou (2013). Here '(A)' and '(M)' denote annual and monthly data availability only. Default yield spread (DFY) is the difference between BAA- and AAA-rated corporate bond yields. Following Campbell (1991), $RREL$ is the difference between the short-term interest rate and its average in the past 12 months. The monthly CAY series is obtained from linear interpolation of the quarterly CAY series (following, e.g., Guo and Whitelaw, 2006). Specifically, in the first month of each quarter, the CAY is the weighted average of the prior quarter's CAY and the current quarter's CAY with weights $\frac{1}{3}$ and $\frac{2}{3}$, respectively. In the second month of the quarter, the respective weights are $\frac{2}{3}$ and $\frac{1}{3}$. In the last month of the quarter, the monthly CAY is equal to the same-quarter CAY observation.

EP	Log earnings-price ratio (S&P 500 earnings yield)
DE	Log dividend-payout ratio
$SVAR$	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
TMS	Term spread: $LTY - TBL$
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)
CAY	Consumption-wealth ratio
$EQIS$ (A)	Percent equity issuing
IK (A)	Investment-to-capital ratio
$RREL$ (M)	Stochastically detrended risk-free rate

Table 1 lists the additional state variables that we consider in the local projections over and above the three state variables given in (9). The set of additional variables is based on prior literature on the predictability of returns and dividends (see the discussion in the Introduction, as well as, e.g., 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.[7]

In our main analysis, we consider the sample period starting from 1952 until the end of 2017. Monthly observations start in March 1952. This choice of starting point is mainly driven by the data availability of CAY, which the prior literature has found to be an important predictor of both dividends and returns. The starting point of our sample coincides with Cochrane (2011) and Campbell and Ammer (1993) and is very close to the ones in Cochrane (2008), Let-

[7] Monthly and annual data updated up to 2017 can be found at Amit Goyal's website `http://www.hec.unil.ch/agoyal/docs/PredictorData2017.xlsx`.

tau and Ludvigson (2005) and van Binsbergen and Koijen (2010). Section IV of the Internet Appendix reports the results for subsamples and earlier data (1928–1951).

3.2 One state variable

We start by estimating the local projections (7) over the full annual sample, for $k \in \{1, 2, \ldots, 15\}$ years, using the last observed dividend yield as a single predictor (state variable). That is, we select $\boldsymbol{x}_t^{(a,k)} = dp_t$ for $a \in \{r, \Delta d, dp\}$, which is the same single state variable as considered by Cochrane (2008). The maximum horizon of 15 years is the same as the longest horizon applied by Cochrane (2011) in his direct regressions.

Figure 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 (7)), for $k = 1$ and $k = 15$ years. Due to the maximum horizon of $k = 15$ years, the first 15 years of the sample are missing from the figures. The first panel of Figure 1 shows that, at short horizons, only little dividend yield variation can be explained by expected discount rates or cash flows: the time-series $\widehat{\delta_t}^{(r,1)}$ and $\widehat{\delta_t}^{(d,1)}$ are mostly flat, while $\widehat{\delta_t}^{(dp,1)}$ is highly volatile. This means that most of the volatility of the dividend yield is attributed to discount rate and cash flow expectations over horizons beyond one year. 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 rather flat, suggesting that cash flow expectations, even at longer horizons, can explain only a minor part of observed market volatility.

Following Eq. (5), the final panel of Figure 1 plots the observed dividend yield and the implied dividend yield obtained as $\widehat{\delta}_t^{(r,k)} - \widehat{\delta}_t^{(d,k)} + \widehat{\delta}_t^{(dp,k)}$ for $k = 1$ and $k = 15$. The plot provides supporting evidence on the accuracy of the LP-based estimates of the components in the 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, reporting the volatility decomposition (12) at the annual frequency, shows a highly similar pattern to Figure 1. The volatility contribution of expected returns ($\widehat{\sigma}(r, k)$) increases over the horizon k, up to a maximum of 0.80 at 15-year horizon, implying that 80 percent of dividend yield volatility can be attributed to expected discount rates. The contribution of expected dividends ($\widehat{\sigma}(d, k)$) remains very low, if not exactly zero: the ratio $\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$ peaks at the 15-year horizon at 0.10, but is mostly close to zero. Overall, these results are highly consistent with the results reported by Cochrane (2008), who argues that expected

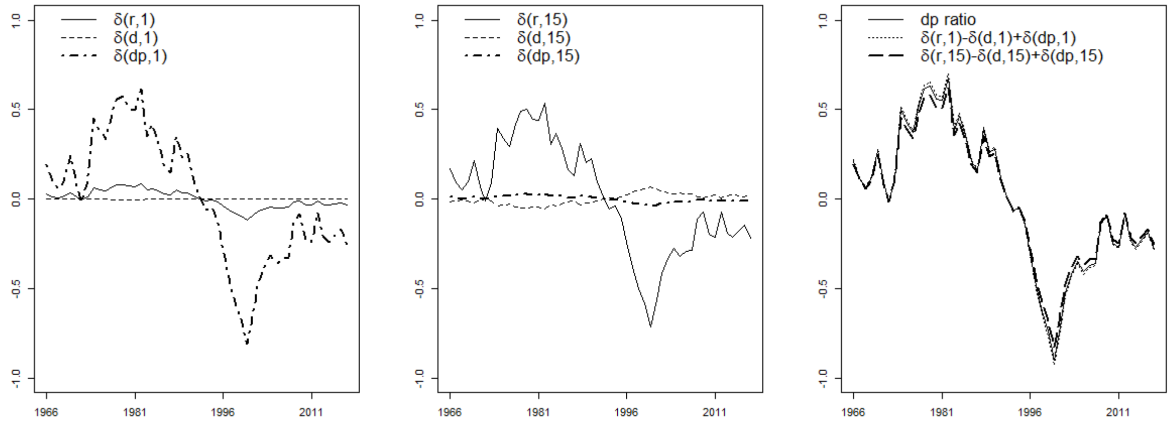

Figure 1: Time-series plots of the the components $\widehat{\delta}_t^{(r,k)}$, $\widehat{\delta}_t^{(d,k)}$, and $\widehat{\delta}_t^{(dp,k)}$ (see (5)), estimated from the annual local projections (7), using dp_t as the single explanatory variable, for $k = 1$ year (left panel) and $k = 15$ years (middle panel). The right panel shows the observed dividend yield dp_t and the implied dividend yield $\widehat{\delta}_t^{(r,k)} - \widehat{\delta}_t^{(d,k)} + \widehat{\delta}_t^{(dp,k)}$, for $k = 1$ year and $k = 15$ years.

cash flow variation hardly contributes to price volatility. In fact, Cohrane's (2008) long-run coefficients can be backed out from our local projections with $k = 1$ and the lagged dividend yield as the only predictor (see Appendix A). We find that the long-run coefficients in his VAR-based approach are $\widehat{b}_r^{lr} = 1.04$ and $\widehat{b}_d^{lr} = 0.03$, confirming that discount rates, in the long run (i.e., the infinite horizon $k \longrightarrow \infty$), are the major drivers of price volatility.[8]

Panel B of Table 2 presents the volatility decomposition (12) estimated with monthly data for horizons of one up to 15 years (i.e., $k = 12$ to $k = 180$ months) and utilizing the approximation (16). 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 ($\widehat{\sigma}(dp, k)$) is diminishing monotonically with the forecast horizon. At the 15-year horizon, the contribution of the expected dividend yield is close to zero. This implies that the lagged dividend yield by itself does not contain any predictive power on the future dividend yield over horizons exceeding 15 years.

3.3 Three state variables

Instead of a single state variable (dividend yield) in Section 3.2, in this section we move to three state variables as specified in (9). That is, lagged cumulative returns and dividend growth rates are included as additional state variables in the estimated local projections.

[8] Full regression (estimation) results of the annual and monthly local projections (7) on all horizons k are reported in Section I and II of the Internet Appendix. This is also the case for the three state variable system to be examined in the next section (Section 3.3).

Table 2: Volatility decomposition: one state variable

This table reports the annual (Panel A) and monthly (Panel B) volatility decomposition of the dividend yield (12), based on the local projections (7), using dp_t as the single explanatory variable for different horizons k. The columns report the relative contributions of expected discount rates $\widehat{\sigma}(r,k)$, cash flow growth $\widehat{\sigma}(d,k)$ and forward dividend yields $\widehat{\sigma}(dp,k)$, as well as the ratio $\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$ given in (13).

A: Annual					B: Monthly				
k (years)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	k (months)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
1	0.13	0.00	0.90	0.03	12	0.12	0.00	0.90	0.02
2	0.23	0.02	0.80	0.07	24	0.21	0.02	0.80	0.09
3	0.28	0.03	0.73	0.10	36	0.28	0.03	0.73	0.11
4	0.33	0.02	0.66	0.06	48	0.33	0.02	0.66	0.07
5	0.40	0.00	0.57	0.00	60	0.40	0.00	0.57	0.01
6	0.47	0.01	0.49	0.03	72	0.45	0.00	0.49	0.01
7	0.53	0.01	0.42	0.03	84	0.51	0.00	0.41	0.01
8	0.59	0.01	0.36	0.01	96	0.58	0.00	0.35	0.00
9	0.65	0.00	0.28	0.00	108	0.64	0.00	0.28	0.00
10	0.70	0.01	0.24	0.01	120	0.68	0.00	0.23	0.01
11	0.73	0.01	0.19	0.01	132	0.70	0.01	0.19	0.01
12	0.73	0.03	0.15	0.04	144	0.71	0.03	0.15	0.04
13	0.74	0.05	0.12	0.07	156	0.71	0.04	0.12	0.06
14	0.76	0.07	0.09	0.09	168	0.73	0.04	0.09	0.06
15	0.79	0.08	0.04	0.10	180	0.77	0.05	0.04	0.06

Moving beyond the one state variable model, in Table 3 we see that the discount rate contribution further increases, but in particular the cash flow contribution is now clearly nonzero. This increase in the volatility of all components is as expected due to additional full-sample predictive power that the lagged cumulative returns and dividend growth rates provide. In Table 2, ratios (13) between the cash flow and discount rate contributions are mostly below 0.1, whereas in Table 3 these ratios are roughly between 0.25–0.5, depending on the horizon k. In other words, the discount rate channel maintains its dominant role but the cash flow contribution is substantially more important when moving ahead from the standard model with the dividend yield as the single state variable.

The additional state variables predict also the future dividend yield. When compared to Table 2, the volatility shares $\widehat{\sigma}(dp,k)$ reported in Table 3 are clearly higher than zero at the long forecast horizons. It is worth noting that this result does not imply rejection of the transversality assumption or the existence of (rational) bubbles: it just suggests that the dividend yield itself is predictable by other factors, even at long but finite horizons.[9]

Figure 2 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

[9] When considering horizons beyond $k = 15$ years, the share $\widehat{\sigma}(dp,k)$ does converge to zero, which is consistent with the transversality assumption. This convergence is nevertheless much slower than when the dividend yield is the single state variable (see Table 2).

Table 3: Volatility decomposition: three state variables

This table reports the annual (Panel A) and monthly (Panel B) volatility decomposition (12) based on the local projections (7) using three state variables (9): $\boldsymbol{x}_t^{(a,k)} = \left(\sum_{j=1}^{k} \rho^{j-1} r_{t+j-k}, \sum_{j=1}^{k} \rho^{j-1} \Delta d_{t+j-k}, dp_t\right)$ for $a \in \{r, d, dp\}$, for different horizons k. The columns report the relative contributions of expected discount rates $\widehat{\sigma}(r,k)$, cash flow growth $\widehat{\sigma}(d,k)$, forward dividend yields $\widehat{\sigma}(dp,k)$, and the ratio $\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$ given in (13).

A: Annual					B: Monthly				
k (years)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	k (months)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
1	0.15	0.12	0.92	0.76	12	0.14	0.16	0.91	1.17
2	0.30	0.08	0.80	0.25	24	0.24	0.10	0.81	0.42
3	0.28	0.17	0.77	0.62	36	0.31	0.14	0.77	0.43
4	0.40	0.27	0.73	0.66	48	0.41	0.25	0.73	0.61
5	0.48	0.23	0.66	0.48	60	0.44	0.22	0.66	0.50
6	0.49	0.19	0.59	0.38	72	0.45	0.16	0.58	0.35
7	0.53	0.15	0.52	0.28	84	0.52	0.15	0.52	0.28
8	0.64	0.17	0.47	0.27	96	0.61	0.17	0.49	0.28
9	0.69	0.23	0.41	0.33	108	0.67	0.26	0.46	0.39
10	0.74	0.32	0.48	0.43	120	0.71	0.32	0.53	0.45
11	0.76	0.34	0.48	0.45	132	0.73	0.35	0.51	0.48
12	0.77	0.36	0.44	0.46	144	0.75	0.35	0.46	0.46
13	0.79	0.32	0.35	0.41	156	0.77	0.32	0.37	0.41
14	0.84	0.27	0.32	0.32	168	0.82	0.28	0.33	0.34
15	0.89	0.25	0.27	0.29	180	0.88	0.26	0.29	0.30

short horizons, the picture is similar to Figure 1, with both expected returns and expected dividends being fairly flat. At longer horizons, the dividend growth and dividend yield components $\widehat{\delta}_t^{(d,15)}$ and $\widehat{\delta}_t^{(dp,15)}$ are now strikingly more volatile than in Figure 1, thereby clearly contributing to the volatility of the dividend yield. The final panel of Figure 2 shows that the approximate present value relation (5) holds accurately also when the implied dividend yield is based on these multivariate local projections.

The interpretation of our decomposition based on three state variables is that variation in expected discount rates is not the sole driver of market volatility. Expectations on future dividends do in fact contribute significantly to the variation of the dividend yield. Conditional on the lagged dividend yield only, expected dividend growth rates are flat (Figure 1). However, after adding lagged returns and dividend growth rates to the information set, these expectations do vary over time. This result corroborates the conclusions by Menzly, Santos, and Veronesi (2004), Lettau and Ludvigson (2005), Ang and Bekaert (2007), van Binsbergen and Koijen (2010), and others, who document dividend growth rate predictability by other factors than the lagged dividend yield. The flexibility of local projections allow us to integrate this predictability into the volatility decomposition.

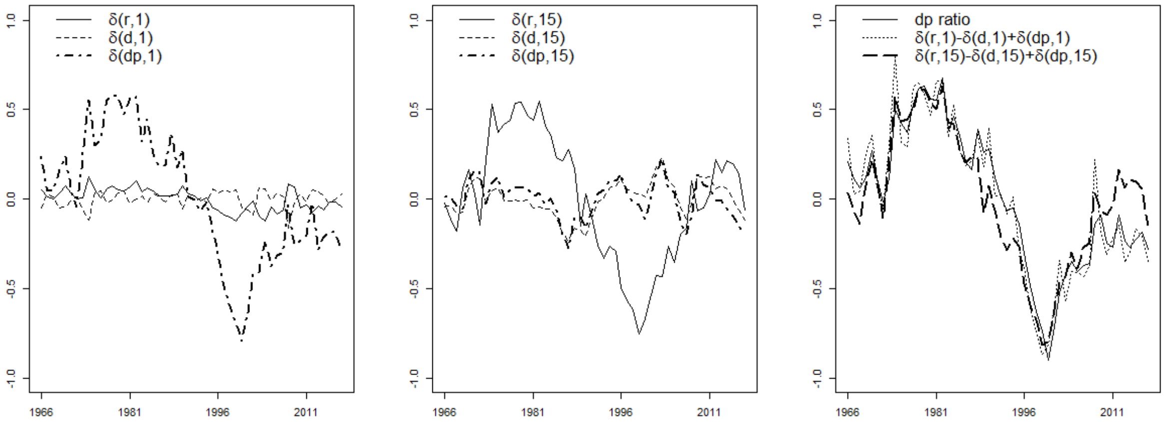

Figure 2: Time-series plots of the components $\widehat{\delta}_t^{(r,k)}$, $\widehat{\delta}_t^{(d,k)}$, and $\widehat{\delta}_t^{(dp,k)}$ (see (5)), estimated from the annual local projections (7), using three fixed state variables: $\boldsymbol{x}_t^{(a,k)} = \left(\sum_{j=1}^{k} \rho^{j-1} r_{t+j-k}, \sum_{j=1}^{k} \rho^{j-1} \Delta d_{t+j-k}, dp_t \right)$ for $a \in \{r, d, dp\}$, 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 $\widehat{\delta}_t^{(r,k)} - \widehat{\delta}_t^{(d,k)} + \widehat{\delta}_t^{(dp,k)}$, for $k = 1$ year and $k = 15$ years.

3.4 LASSO and model averaging results

In this section, we enlarge the information set from three state variables (9) to the more general case where all the potential state variables described in Table 1 are involved in the construction of local projections. That is, we are moving to the data-rich model averaging and LASSO approaches as introduced in Section 2.3 where, importantly, the built-in regularization mechanisms control the potential hazards of overfitting.

Before moving to the empirical results, let us clarify a couple of modelling selections made in this section. According to the regression results in Section 3.3 and also LASSO model selections (especially based on the monthly data), we pre-specify model averages so that the model always includes lagged cumulative returns, dividend growth and the dividend yield (i.e., all three state variables in (9)). The model averages are then constructed as in (14), by including a fourth predictor that varies across specifications j and is one of the variables listed in Table 1. Moreover, in the LASSO estimation we do not apply shrinkage to the dividend yield (i.e., the penalty term in (15) does not include the regression coefficient related to the lagged dividend yield), due to its essential role in the benchmark volatility decompositions (see Section 2 and also Engsted, Pedersen, and Tanggaard, 2012). This implies that dp_t is always included in the resulting local projections.[10]

[10] Internet Appendix Section III 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 4.

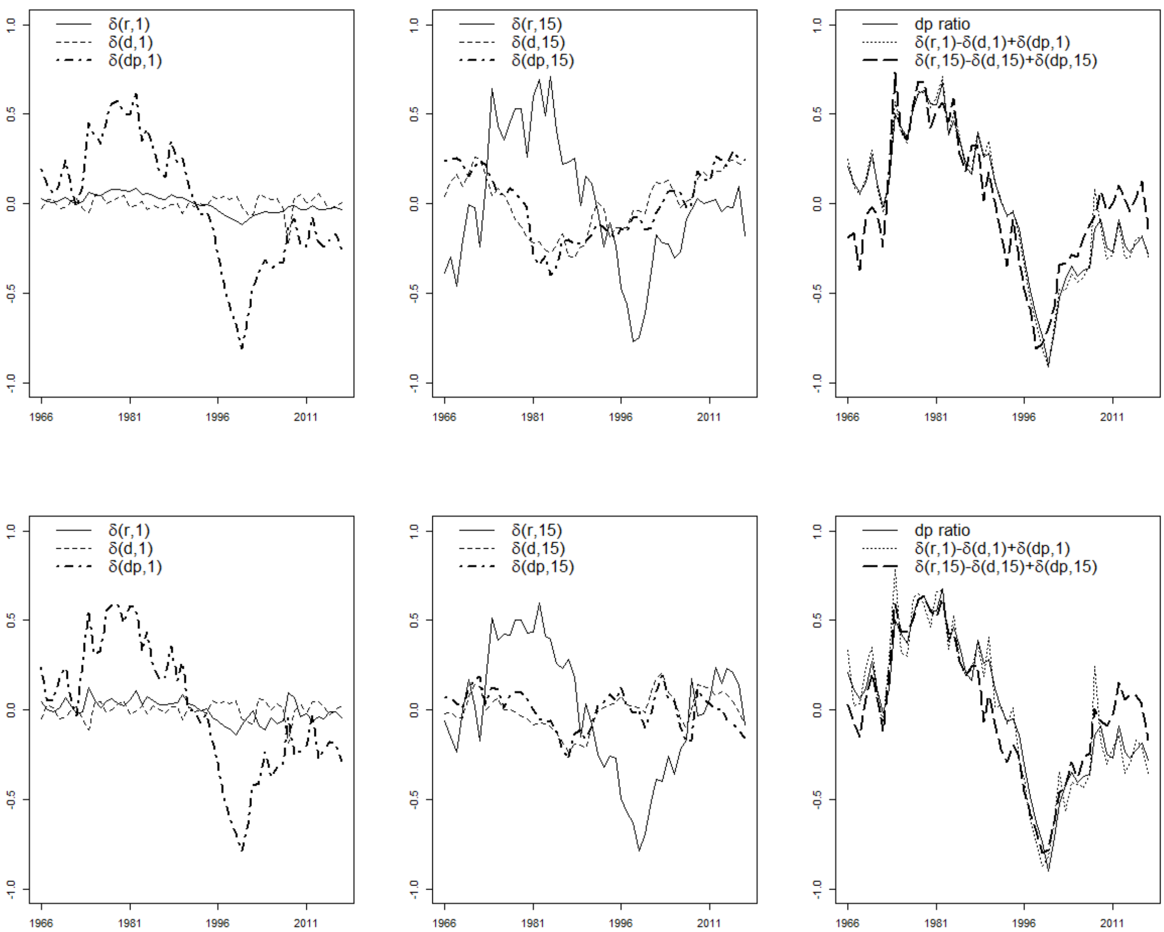

Figure 3: Time-series plots of the the components $\widehat{\delta}_t^{(r,k)}$, $\widehat{\delta}_t^{(d,k)}$, and $\widehat{\delta}_t^{(dp,k)}$ (see (5)), estimated from the annual local projections (7), using the LASSO (upper panel) and model averaging (below panel) approaches, for $k = 1$ year (left panel) and $k = 15$ years (middle panel). The right panel shows the observed dividend yield dp_t and the implied dividend yield $\widehat{\delta}_t^{(r,k)} - \widehat{\delta}_t^{(d,k)} + \widehat{\delta}_t^{(dp,k)}$, for $k = 1$ year and $k = 15$ years.

Table 4 presents the volatility decomposition using local projections estimated by the LASSO and model averaging. Overall, the results are qualitatively in line with those obtained with three state variables. Our main result in Table 3, that expected cash flow variation is by no means negligible relative to expected discount rate variation, is robust to expansion of the predictive information set. Even though long-run expected returns remains the dominant factor, the ratio (13) of dividend yield volatility that can be attributed to expected long-run (15-year) cash flow growth relative to expected discount rates is 0.46 (0.43) when estimated with annual (monthly) data. This is considerably higher than the share of 0.10 (0.06) found with the dividend yield as the single state variable in Table 2.

Table 5 presents detailed LASSO regression results, i.e., the estimated local projection models at different forecast horizons. These results include both the variable selection (i.e., the

Table 4: Volatility decomposition: LASSO and model averaging

This table reports the annual and monthly volatility decomposition of the dividend yield (12), based on the local projections (7), using the LASSO and model averaging approaches for different horizons k. The columns report the relative contributions of expected discount rates $\widehat{\sigma}(r,k)$, cash flow growth $\widehat{\sigma}(d,k)$ and forward dividend yields $\widehat{\sigma}(dp,k)$, as well as the ratio $\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$ given in (13).

A: LASSO - Annual					B: LASSO - Monthly				
k (years)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	k (months)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
1	0.13	0.12	0.90	0.91	12	0.23	0.24	0.94	1.06
2	0.24	0.19	0.81	0.77	24	0.39	0.28	0.86	0.74
3	0.34	0.24	0.82	0.70	36	0.52	0.31	0.82	0.59
4	0.56	0.27	0.78	0.49	48	0.62	0.31	0.80	0.50
5	0.62	0.30	0.73	0.49	60	0.72	0.31	0.76	0.43
6	0.72	0.30	0.70	0.41	72	0.80	0.32	0.79	0.40
7	0.75	0.24	0.67	0.32	84	0.83	0.32	0.76	0.38
8	0.86	0.27	0.66	0.31	96	0.95	0.33	0.74	0.35
9	0.94	0.29	0.66	0.31	108	0.92	0.39	0.75	0.43
10	0.94	0.28	0.62	0.29	120	0.91	0.38	0.70	0.42
11	0.92	0.32	0.67	0.35	132	0.92	0.48	0.68	0.52
12	0.96	0.51	0.54	0.53	144	0.93	0.54	0.65	0.58
13	0.96	0.38	0.54	0.40	156	0.94	0.52	0.61	0.56
14	0.94	0.36	0.54	0.38	168	0.92	0.49	0.54	0.54
15	0.93	0.42	0.50	0.46	180	0.95	0.40	0.55	0.43
C: Model averaging - Annual					D: Model averaging - Monthly				
k (years)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	k (months)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
1	0.15	0.12	0.92	0.78	12	0.14	0.17	0.92	1.20
2	0.30	0.08	0.81	0.27	24	0.24	0.10	0.81	0.44
3	0.28	0.17	0.78	0.62	36	0.31	0.14	0.78	0.45
4	0.41	0.27	0.73	0.66	48	0.41	0.25	0.73	0.61
5	0.49	0.23	0.67	0.48	60	0.44	0.22	0.66	0.50
6	0.50	0.19	0.59	0.38	72	0.46	0.16	0.59	0.35
7	0.54	0.16	0.52	0.29	84	0.53	0.15	0.52	0.28
8	0.65	0.18	0.48	0.27	96	0.62	0.18	0.49	0.28
9	0.70	0.23	0.42	0.33	108	0.68	0.25	0.45	0.37
10	0.75	0.31	0.47	0.42	120	0.73	0.32	0.53	0.44
11	0.76	0.36	0.48	0.47	132	0.74	0.37	0.52	0.50
12	0.77	0.37	0.44	0.48	144	0.76	0.37	0.47	0.49
13	0.80	0.34	0.36	0.42	156	0.77	0.34	0.37	0.44
14	0.85	0.29	0.33	0.34	168	0.83	0.30	0.33	0.36
15	0.89	0.26	0.28	0.29	180	0.87	0.27	0.29	0.31

inclusion and exclusion of state variables), and the specific estimated LASSO coefficients (see (15)). The results in Table 5 are based on annual data: monthly model selection results are reported in Section III of the Internet Appendix. 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 lagged cumulative returns, term spread (TMS), volatility ($SVAR$), investment-to-capital ratio (IK), and consumption-wealth ratio (CAY). As reviewed in the Introduction and Section 3.1, these are largely the variables that are expected to have predictive power (see Fama and French, 1989; Martin, 2017; Lettau and Ludvigson, 2001, 2005). In contrast, earnings yield (EP), book-to-market ratio (BM), and short-term interest rates (TBL and $RFREE$) are typically excluded. For the dividend growth local projections, also CAY appears to be an important predictor, as well as the dividend-payout ratio (DE) and net equity expansion ($NTIS$). For the dividend yield local projections, we get more exclusions (i.e., we select a smaller set of state variables), but also here CAY is a relevant predictor at almost all horizons k.

Even if the LASSO and model averaging approaches exploit larger information sets, the resulting volatility decompositions in Table 4 are overall not too different from those obtained with three fixed state variables. Also the time series plots of $\widehat{\delta}^{(r,k)}$, $\widehat{\delta}^{(d,k)}$, and $\widehat{\delta}^{(dp,k)}$ in Figure 3 are largely similar to those in Figure 2, while the final panels in these figures again reaffirm the validation of the approximate present-value relation (5). In addition, the monthly and annual results in Table 4 are generally very similar for both LASSO and model averaging.

In Section III of the Internet Appendix, we still report the monthly LASSO estimates and model selection results. Compared to the annual case (Table 5), there are clearly less exclusions (i.e., more state variables are included in for the monthly frequency), presumably due to the larger number of observations in estimation, but this turns out have only marginal impact on the resulting volatility decomposition, as seen in Table 4. Furthermore, we also extend the presented LASSO procedure with a cross-validation-based determination of the tuning parameter (cf. remarks below Eq. (15)), a post-LASSO estimation step, and the elastic net method. It is important to keep in mind that in addition to selecting state variables, such as presented in Table 5, the LASSO estimator shrinks all the coefficients towards zero. However, in the post-LASSO step the idea is to first select the state variables with the LASSO and perform the final estimation by OLS. The elastic net respectively combines the LASSO with the closely related ridge regression, leading to an important robustness check for the penalty (shrinkage) term as defined in (15). Overall, it turns out that all the main conclusions on the importance of the discount rate and cash flow channels in the volatility decomposition are intact (i.e., the main results are robust to alternative LASSO specifications).

Table 5: LASSO model selection and estimation results – Annual data, 1952–2017

This table reports LASSO model selection results and parameter estimates (see (15)). Empty cells indicate the exclusion of state variables. Candidate state variables are described in Table 1 and r and Δd denote the (discounted) cumulative lagged returns and dividend growth rates (see (3.3)). The results are reported for local projections with three different left hand side (LHS) variables: cumulative returns (top panel); cumulative dividend growth (middle panel); and the dividend yield (bottom panel).

k (years)	const	r	Δd	dp	EP	DE	$SVAR$	BM	$NTIS$	TBL	LTY	$RFREE$	TMS	LTR	DFY	DFR	$CORPR$	$INFL$	CAY	$EQIS$	IK
									LHS: Cumulative return												
1	0.55			0.13																	
2	1.04	-0.10		0.23															1.12		
3	1.36	-0.10		0.30									1.34						3.05		
4	2.42	-0.15	-0.31	0.34	0.23		1.31			-1.83		-0.00	1.21						5.88		
5	2.65	-0.12	-0.34	0.58			1.95						2.19						4.72	-0.18	-0.89
6	2.81	-0.12	-0.21	0.53			1.51						3.60			-0.87		2.04	6.05		-11.60
7	2.79	-0.09		0.51		-0.04	0.55						5.55			-0.93		3.12	5.86		-16.58
8	3.56	-0.19	-0.52	0.56			1.29		-0.27		2.56		3.07			-0.27		2.67	6.03		-23.63
9	4.12	-0.32	-0.85	0.36	0.06		2.97	0.22	-3.99		8.03			0.13	-14.65	-0.53		1.00	4.52	-0.14	-47.89
10	4.23	-0.09	-0.34	0.27	0.40		2.73		-3.17		2.16		4.59	0.38	-9.02			3.98	4.07		-37.37
11	3.45			0.21	0.42		2.44		-1.68		2.92		6.25	0.13		-0.65		3.30	1.83	-0.27	-29.48
12	3.81	-0.23		-0.20	0.66	0.49	4.03		-2.29		7.85		4.38	0.42	-22.59	-1.33		2.03		-0.60	-41.20
13	4.16	-0.34	-0.12	0.54			4.30		-4.12		4.46		5.51		-15.12	-0.97		1.55	1.11		-27.66
14	2.97	-0.51	0.25	0.25		0.20	1.23		-2.50		5.06		7.83	0.31	-24.03	-1.58		3.74	0.81	-0.17	-20.04
15	2.62	-0.33	-0.24	-0.02		0.20	-0.27	0.84	-6.31		5.38		3.45	0.17	-21.18	-0.31		1.41	1.51	-0.10	-34.89
									LHS: Cumulative dividend growth												
1	-0.01	0.16		-0.00		-0.08			0.58						-0.51				-0.22		
2	-0.12		-0.12	-0.05		-0.15		0.28	2.27				0.47			0.20	0.14	-1.20	-0.21	-0.33	-3.12
3	-0.10	-0.09	-0.22	-0.06		-0.09		0.31	1.79	-1.01			1.04		-2.10	0.36	0.19			-0.17	
4	0.01	-0.20	-0.39	-0.09		-0.00		0.26	0.82						-2.46	0.37	0.11				
5	-1.22		-0.44	-0.40	-0.07	0.02		0.87		-0.51		-0.00	0.46		-5.46	0.48	0.09			0.16	-7.65
6	-1.31	0.09	-0.40	-0.49		0.08		0.91		-0.64	-0.63	-0.18				0.80	0.15		-0.85	0.08	-6.48
7	-0.33		-0.25	-0.25				0.35	1.00							0.15			-1.76	0.27	-5.22
8	0.07		-0.36	-0.13		-0.01		0.15	0.90									0.30	-2.70	0.43	
9	0.67		-0.46	-0.02		-0.02		0.06	0.99									0.12	-3.37	0.25	
10	0.83	-0.05	-0.28	-0.01					1.70										-3.03		
11	1.11	-0.10	-0.28	0.02					1.56	-0.62	-0.51	-0.02							-2.61		
12	3.57	-0.15	-0.92	0.02	0.53	0.31	1.31	-0.61			-0.70	-1.12				0.23	0.12	-1.05	-4.68	0.38	6.17
13	1.57		-0.48	0.07		0.04	-0.28		0.91	-1.36		-0.00							-4.07		
14	0.65			-0.08					1.88	-1.26	-0.40								-3.81		
15	0.97		-0.15	-0.11		0.12	-1.09		0.86	-1.30	-1.23	-0.03		0.03				1.13	-4.08		-2.59
									LHS: Dividend yield												
1	-0.25			0.90																	
2	-0.58			0.79															-1.99		
3	-1.58			0.53				0.33											-3.48		
4	-2.25			0.34		0.08		0.48	-0.94										-4.55		2.60
5	-2.47		0.18	0.30				0.31											-5.47	0.25	4.28
6	-1.97			0.40															-7.12	0.39	8.08
7	-2.18			0.31															-7.45	0.45	7.85
8	-2.58			0.22									-0.73						-7.83	0.51	11.72
9	-2.97			0.17							-1.20								-7.89	0.77	20.46
10	-3.00			0.16					1.89		-1.03								-7.61	0.33	23.70
11	-3.90	-0.04		-0.07		0.23	-1.67	0.67	2.63		-5.00		-1.65	-0.12	4.82	0.92		-1.95	-3.61	0.22	34.81
12	-2.55			0.22			-0.73		4.13		-2.99								-5.21		25.32
13	-2.40			0.17			-1.85		4.38		-3.38		-0.47						-4.57		20.22
14	-1.54	0.18	-0.24	0.25	0.20		-1.35		3.93		-8.62		-1.10	-0.20	20.54	0.83			-3.15		25.00
15	-1.56			0.11					1.68		-5.22		-2.13	-0.35			-0.04		-3.38		

3.5 Time-varying parameters

The larger sample size, resulting from the use of monthly data in our local projection approach, facilitates us to allow the underlying parameters of our volatility decompositions 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 potential time-variation of our volatility decomposition is grounded in recent and mounting empirical and theoretical evidence suggesting that return predictability is time-varying (e.g., Timmermann, 2008; Rapach, Strauss, and Zhou, 2010; Henkel, Martin, and Nardari, 2011; Dangl and Halling, 2012; Zhu, 2015; Zhu and Zhu, 2013; Farmer, Schmidt, and Timmermann, 2018; and Cochrane, 2017), 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 single return prediction models can, at best, hope to uncover evidence of local predictability. Recent findings suggest that the predictive power often concentrates during bad times in financial markets (see Henkel, Martin, and Nardari, 2011; Zhu and Zhu, 2013; Cujean and Hasler, 2017). Zhu (2015) also finds that time-varying predictability of return and dividend growth is a tug-of-war: when returns are predictable, dividend growth is not, and vice versa. Furthermore, Choi, Kim, and Park (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.

As argued in the previous sections, the use of horizon-specific local projections reduces model misspecification concerns. However, allowing time-variation in the parameter coefficients may further increase the accuracy of the estimated discount rate and cash flow components. Therefore, we extend the local projections (7) by allowing for time-varying parameters:

$$\begin{aligned}
\sum_{j=1}^{k} \rho_t^{j-1} r_{t+j} &= \alpha_t^{(r,k)} + \boldsymbol{x}_t^{(r,k)} \boldsymbol{\beta}_t^{(r,k)} + \varepsilon_{t+k}^{(r,k)} \\
\sum_{j=1}^{k} \rho_t^{j-1} \Delta d_{t+j} &= \alpha_t^{(d,k)} + \boldsymbol{x}_t^{(d,k)} \boldsymbol{\beta}_t^{(d,k)} + \varepsilon_{t+k}^{(d,k)} \\
\rho_t^{k} dp_{t+k} &= \alpha_t^{(dp,k)} + \boldsymbol{x}_t^{(dp,k)} \boldsymbol{\beta}_t^{(dp,k)} + \varepsilon_{t+k}^{(dp,k)},
\end{aligned} \tag{18}$$

where the parameters $\alpha_t^{(a,k)}$ and $\boldsymbol{\beta}_t^{(a,k)}$ are now time-varying.[11]

Specifically, we allow for time-variation by estimating the coefficients in (18) 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 $\left(\sum_{i=0}^{t}\phi^i\right)^{-1}\phi^i$. The decay parameter ϕ is a number between zero and one. Following the convention in the literature, we calibrate the decay parameter ϕ at 0.97, which is suggested by J.P. Morgan's (1996) Riskmetrics report as the optimal exponential decay parameter for modeling volatility using monthly data. 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.[12]

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

$$\widehat{\rho}_t = \frac{e^{\bar{dp}_t}}{1+e^{\bar{dp}_t}}, \tag{19}$$

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

$$\bar{dp}_t = \left(\sum_{i=0}^{t}\phi^{t-i}\right)^{-1}\sum_{i=0}^{t}\phi^{t-i}dp_t, \tag{20}$$

where, as in the regressions (18), the decay parameter ϕ 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, but also the earnings yield (EP) and book-to-market (BM) ratio, for level shifts increases the predictive performance substantially. Lettau and Nieuwerburgh (2008) correct these nonstationarities by estimating the timing of structural break points. Locating the exact timing of breaks or identifying regime switching patterns (cf. Zhu, 2015; and Choi, Kim and Park, 2017) is in general a difficult task, in particular in small samples. Therefore, we handle time-variation of the steady-state levels of variables

[11] As argued by Granger (2008), any nonlinear model can be approximated by linear models with time-varying parameters. Hence the LPs in (18) are able to accommodate various nonlinear patterns that may have an impact on the volatility decomposition.

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

by detrending not only the dividend yield, but also all other variables using the same recursive EWMA filtration as in (20). Specifically, in our time-varying local projections (18), we recursively detrend all variables by subtracting at each point in time the exponentially weighted moving average, as opposed to demeaning these variables over the full sample, as we have done so far. Thereby, we implicitly also allow the parameter κ in the log-linear present value model (1) to vary over time.

Table 6: Monthly volatility decompositions with time-varying parameters

This table reports the monthly (unconditional) volatility decomposition of the dividend yield (12), based on the local projections (18) estimated by Exponentially Weighted Least Squares (EWLS) using an expanding window approach. The local projections contain one (dividend yield) and three state variables (9) along with the LASSO and model averaging strategies for various (annualized) horizons of k months. All variables are demeaned recursively using EWMA filtration.

	A: One state variable				B: Three state variables			
k (months)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
12	0.39	0.37	1.06	0.96	0.90	0.57	1.09	0.63
24	1.04	0.60	0.98	0.58	1.49	0.94	1.24	0.63
36	1.40	0.74	1.09	0.53	1.48	1.06	1.28	0.72
48	1.61	0.69	1.30	0.43	1.89	0.76	1.50	0.40
60	1.52	0.64	1.52	0.42	1.47	1.05	1.21	0.72
72	1.00	0.54	1.34	0.54	2.12	0.71	1.47	0.34
84	0.59	0.57	1.00	0.97	1.55	0.59	1.24	0.38
96	0.74	0.65	0.72	0.89	0.87	0.78	0.59	0.89
108	1.23	0.54	0.89	0.44	1.30	0.62	0.84	0.48
120	1.12	0.64	0.90	0.57	1.48	0.67	1.01	0.46
132	1.22	0.85	0.84	0.70	1.49	0.91	0.86	0.61
144	1.04	0.93	0.78	0.89	1.59	0.82	0.78	0.51
156	0.86	0.86	0.80	1.00	1.00	0.87	0.95	0.87
168	0.95	0.67	0.98	0.71	1.56	0.60	1.22	0.38
180	0.72	0.56	0.75	0.78	1.25	0.65	0.80	0.52
	C: LASSO				D: Model averaging			
k (months)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
12	0.61	0.45	0.97	0.74	0.88	0.56	1.10	0.64
24	0.82	0.59	0.87	0.72	1.40	0.80	1.23	0.57
36	0.79	0.61	0.91	0.77	1.32	0.97	1.28	0.74
48	1.01	0.57	0.91	0.56	1.72	0.71	1.39	0.41
60	1.57	0.52	1.25	0.33	1.43	0.99	1.21	0.69
72	1.42	0.50	1.36	0.35	2.07	0.70	1.46	0.34
84	0.89	0.47	0.98	0.53	1.55	0.55	1.25	0.36
96	1.04	0.59	1.00	0.57	0.90	0.75	0.54	0.83
108	1.13	0.59	1.16	0.52	1.25	0.59	0.83	0.47
120	1.28	1.04	0.71	0.81	1.33	0.63	0.98	0.47
132	1.36	0.96	1.08	0.71	1.28	0.84	0.83	0.66
144	1.56	0.99	0.95	0.63	1.53	0.83	0.79	0.54
156	1.66	0.66	0.82	0.40	1.02	0.87	0.90	0.86
168	1.21	0.61	1.17	0.51	1.36	0.59	1.16	0.44
180	1.36	0.73	0.89	0.54	1.26	0.55	0.75	0.44

Table 6 presents the monthly (unconditional) volatility decomposition (12) over different horizons for all different choices of state variables considered so far: i.e., one state variable, three state variables, LASSO selection and model averaging. Note that for computing the volatility decomposition (12), also the dividend-price ratio dp_t is now recursively demeaned using the EWMA filtration (20). The local projections are estimated with monthly data, by EWLS and using an expanding window approach with an initial window size of 240 months. Since reliable implementation of the recursive estimation requires a reasonable number of observations, we only present results based on monthly data.

The main pattern in Table 6 is very clear: The relative importance of the cash flow variation turns out to be even higher than with the full-sample results, with the long-run (180 months) ratio (13) fluctuating between 0.44 and 0.78, depending on the choice of state variables. In other words, the time-varying local projections provide additional support for our general finding that local predictability of cash flows does contribute significantly to observed dividend yield volatility.

It is specifically worth noting in Table 6 that the reported ratios (13) between time-varying cash flow and discount rate components are in particular high in the case of a single state variable. That is, allowing time-variation and monthly data, due to the introduction of the local projections-based methods, reveals strong local predictability of dividend growth by the dividend yield that is not identified by the conventional and past static full-sample approaches.

In addition to the unconditional (i.e., full sample) volatility decompositions in Table 6, our expanding window approach in particular allows us to consider the time-varying contribution of the different components to the dividend yield volatility over time. Since we are in particular interested in the relative contribution of expected cash flows and expected discount rates, in line with (13), we compute the time-varying ratio of the conditional standard deviations of the cash flow and discount rate components:

$$\frac{\widehat{\sigma}_t\left(\widehat{\delta}_t^{(d,k)}\right)}{\widehat{\sigma}_t\left(\widehat{\delta}_t^{(r,k)}\right)}, \tag{21}$$

where the time-varying standard deviations $\widehat{\sigma}_t(.)$ are once again computed using an expanding window and exponential weighting:

$$\widehat{\sigma}_t^2\left(\delta_t^{(a,k)}\right) = \left(\sum_{i=0}^{t}\phi^{t-i}\right)^{-1}\sum_{i=0}^{t}\phi^{t-i}\left(\widehat{\delta}_t^{(a,k)} - \bar{\delta}_t^{(a,k)}\right)^2. \tag{22}$$

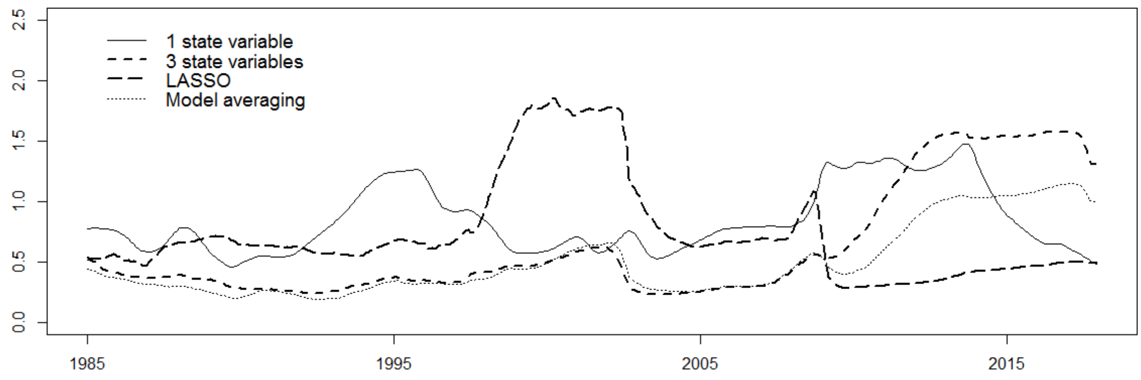

Figure 4: Time-series plots of $\frac{\widehat{\sigma}_t\left(\widehat{\delta}_t^{(d,k)}\right)}{\widehat{\sigma}_t\left(\widehat{\delta}_t^{(r,k)}\right)}$, with $\widehat{\delta}_t^{(d,k)}$ and $\widehat{\delta}_t^{(r,k)}$ estimated by EWLS from rolling window local projections (18), with $k = 180$, using four different information sets: one state variable, three state variables, LASSO selection, and model averaging.

Figure 4 plots the ratio (21) over time, for all four modelling approaches, with $k = 180$ (i.e., for a 15-year horizon). The estimated ratios (21) fluctuate considerably over time. The ratios are mostly below one but clearly higher than zero, fluctuating around 0.5. At times, they do peak above one, indicating that during these periods the contribution of the variation in expected cash flow growth exceeds that of the variation in expected discount rates. All in all, the time-varying dynamics in Figure 4 indicate that both discount rate and cash flow components matter (i.e. both are predictable) with a varying degree of importance. Our time-varying volatility decomposition thus somewhat contradict the 'tug-of-war' hypothesis by Zhu (2015), in which either dividends or returns are predictable. Moreover, we find that the cash flow contribution is systematically more important than reported in the recent regime switching studies by Zhu (2015) and Choi, Kim and Park (2017).

4 Discussion

The main empirical result of this study is that expected dividend growth does contribute to market volatility. The time-varying volatility of the cash flow component $\widehat{\delta}_t^{(d,k)}$ is by no means negligible compared to the volatility of the discount rate component $\widehat{\delta}_t^{(r,k)}$. Only in the static baseline case of a single state variable (the dividend yield) and constant parameters over the full sample (years 1952–2017), we find that expected dividend growth is nearly flat and does not contribute to the volatility of the dividend yield. When we expand the information set to contain multiple state variables, we do find evidence of dividend growth predictability (i.e., variation of expected cash flows), even if expected discount rates remain the primary component in our volatility decomposition. Moreover, when we estimate the local projections recursively (i.e., allowing for time-varying parameters), we find that expected dividend growth rates become substantially more important. During various periods, expected dividend growth becomes temporarily the dominant component of market volatility.

Our results provide a new perspective to the puzzling 'stylized fact' that dividend growth is not predictable by the dividend yield in the US during the postwar period. As documented by Engsted and Pedersen (2010), this finding does not hold in general in international equity markets, while Golez and Koudijs (2018) do find dividend growth predictability in the US prior to 1945. Chen, Da, and Priestley (2012) attribute this apparent lack of dividend growth predictability in the postwar period to dividend smoothing, causing dividend yields to be uninformative of future cash flows, but not necessarily implying that future cash flows are truly unpredictable. Indeed, various studies have found evidence of dividend growth predictability in the US postwar sample (e.g., Lettau and Ludvigson, 2005; Ang and Bekaert, 2007; and Møller and Sander, 2017). Our local projections allow the existing dividend predictability to be recognized in the decomposition of dividend yield volatility, by integrating additional state variables beyond the dividend yield. Moreover, by relying on direct long-run predictions, as opposed to iterated short-run predictions, we circumvent the diminished predictability caused by short-run dividend smoothing and its potential disruptive impact on the volatility decomposition (Chen, Da, and Priestley, 2012).

In terms of return and dividend growth predictability, our results are largely rather consistent with van Binsbergen and Koijen (2010) and the regime switching extensions of their model in Zhu (2015) and Choi, Kim and Park (2017). These approaches are built upon latent variable models with Kalman filtering to extract unobserved components in returns and dividend growth. Van Binsbergen and Koijen (2010) incorporate the full lagged history returns, dividend growth, and dividend yield and find the long-term dependence statistically important in their analysis. This is in spirit similar to our approach where we employ the cumulative lagged returns and dividend growth rates as predictors along with the lagged dividend yield to enlarge the information set. In addition, our approach allows us to extend the information set set even further beyond lagged returns and dividend growth, and to consider monthly data. Zhu (2015) and Choi, Kim and Park (2017) find, as we do with time-invariant and also time-varying parameter local projections, important divergences from the conventional linear

VAR-based single state variable benchmark. Their results show that it is necessary to allow regime switches to increase the importance of the cash flow component, while here linear but otherwise much more flexible and information-rich local projections emphasize the cash flow component even more. Moreover, as discussed at the end of Section 3.5, we do not obtain as strong evidence as Zhu (2015) for the 'tug-of-war' hypothesis, i.e. alternating return and dividend growth predictability.

We find that the time-varying volatility decomposition is sensitive to the choice of state variables (i.e.: the four time-varying ratios plotted in Figure 4 are not highly correlated), which indicates that different factors predict discount rates and cash flows at different points in time. We emphasize that the aim of this paper is not to determine which factors forecast cash flows and returns as such, but rather to evaluate the relative magnitudes of cash flow and return predictability. Our results are clear in the sense that, regardless of the set of state variables, we find that allowing for a time-varying volatility decomposition increases the contribution of dividends relative to discount rates. This is demonstrated by the volatility ratios of cash-flow predictability to return predictability (13), which are for all choices of state variables higher with the time-varying-parameter LPs (Section 3.5) than with the time-invariant (i.e., linear) LPs. This time-varying nature of dividend growth predictability is consistent with dividend growth being subject to time-varying payout policies, affected by factors including dividend smoothing and time-varying investor demand for dividends (e.g., Baker and Wurgler, 2004; Chen, Da, and Priestley, 2012; Larkin, Leary, and Michaely, 2017).

5 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 volatility decomposition. In addition to general flexibility and robustness to model misspecification, local projections allow us to employ LASSO model selection and model averaging, and thereby incorporate large sets of potential state variables. Moreover, despite strong seasonalities in dividend payments, we are able to accommodate monthly data in addition to annual data. The enlarged sample size due to the use of monthly data allows us to apply recursive estimation to examine time variation in the the dividend yield volatility decomposition.

Our results generally confirm that variation in expected discount rates is the dominant component of observed dividend yield volatility. However, the cash flow component is also very much present. Only in the restrictive case of linear (i.e., non-time-varying) local projections with a single state variable (the dividend yield), we find that the contribution of expected cash flows is close to zero. Moving beyond this basic static model, by extending the set of state variables and/or allowing for time-varying parameters, we find that the contribution of expected cash flows is not negligible: the ratio of expected long-run cash flow volatility to expected long-run discount rate volatility ranges between 0.2 and 0.5, depending on the specification. Our time-varying volatility decomposition shows that during certain periods, expected cash flows in fact contribute more to market volatility than expected discount rates.

By incorporating multiple state variables and time-varying parameters within our local projection framework, we believe that we provide a more robust volatility decomposition than prior studies. Indeed, various alternative specifications and robustness checks reported throughout this paper and the Internet Appendix point to the same main conclusion: variation in expected dividend growth contributes significantly to dividend yield volatility.

Appendix A VAR-based approaches

In this Appendix, we briefly outline the commonly used 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 (Δd_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 (6) with infinite horizon ($k \longrightarrow \infty$), Campbell and Shiller (1988b) attempt to estimate the component associated with expected dividend growth $\Big(\delta_t^{(d,\infty)} =$

$E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} \Big)$ by fitting a bivariate VAR to the annual price-dividend ratio and the annual dividend growth rate (both measured in logs):

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

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 ρ (see, e.g., (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 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 (7) at different horizons k, which has several advantages as we discuss in Section 2.

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 $\varepsilon_{t+1}^{(r)} = \varepsilon_{t+1}^{(d)} + \rho\, \varepsilon_{t+1}^{(dp)}$. 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) derives, 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:

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

The fitted values of these hypothetical regressions thus correspond to the fitted values of our local projections (8), in the special case of the dividend yield as the only predictor and the infinite horizon ($k \longrightarrow \infty$):

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

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

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

for $a\in\{r,d\}$.

References

Ang, A., and G. Bekaert (2007). Return predictability: Is it there? *Review of Financial Studies* 20, 651–707.

Baetje, F. (2018). Does a lot help a lot? Forecasting stock returns with pooling strategies in a data-rich environment. *Journal of Forecasting* 37, 37–63.

Baker, M., and J. Wurgler (2004). A catering theory of dividends. *Journal of Finance*, 59(3), 1125-1165.

van Binsbergen, J.H., and R.S.J. Koijen (2010). Predictive regressions: A present-value approach. *Journal of Finance* 65, 1493–1471.

Boudoukh, J., M. Richardson, and R.F. Whitelaw (2008). The myth of long-horizon predictability. *Review of Financial Studies* 21, 1577–1605.

Campbell, J.Y. (1991). A variance decomposition for stock returns. *Economic Journal* 101, 157–179.

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

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

Campbell, J.Y., and R. J. Shiller (1988b). The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors. *Review of Financial Studies* 1, 195–228.

Chen, L. (2009). On the reversal of return and dividend growth predictability: A tale of two periods. *Journal of Financial Economics* 92, 128–151.

Chen, L., Z. Da, and R. Priestley (2012). Dividend smoothing and predictability. *Management Science*, 58(10), 1834-1853.

Chen, L., and X. Zhao (2009). Return decomposition. *Review of Financial Studies* 22, 5213–

5249.

Chevillon, D. (2007). Direct multi-step estimation and forecasting. *Journal of Economic Surveys* 21, 746–785.

Choi, K.H., C-J. Kim, and C. Park (2017). Regime Shifts in Price-Dividend Ratios and Expected Stock Returns: A Present-Value Approach. *Journal of Money, Credit and Banking* 49, 417–441.

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

Cochrane, J.H. (2005). *Asset Pricing. Revised Edition*. Princeton University Press, New Jersey.

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

Cochrane, J.H. (2011). Presidential address: Discount rates. *Journal of Finance* 66, 1047–1108.

Cochrane, J.H. (2017). Macro-Finance. *Review of Finance* 21, 945–985.

Cochrane, J.H., and M. Piazzesi (2002). The Fed and interest rates – A high-frequency identification. *American Economic Review* 92, 90–95.

Cujean, J., and M. Hasler (2017). Why Does Return Predictability Concentrate in Bad Times? *Journal of Finance* 72, 2717–2758.

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

Engsted, T., and T.Q. Pedersen (2010). The dividend-price ratio does predict dividend growth: International evidence. *Journal of Empirical Finance* 17, 585–605.

Engsted, T., T.Q. Pedersen, and C. Tanggaard (2012). Pitfalls in VAR Based Return Decomposition: A clarification. *Journal of Banking and Finance* 36, 1255–1265.

Fama, E. F., and French, K. R. (1988). Dividend yields and expected stock returns. *Journal of Financial Economics*, 22(1), 3–25.

Fama, E.F., and K.R. French (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics* 25, 23–49.

Farmer, L.E, L. Schmidt, and A. Timmermann (2018). Pockets of predictability. Unpublished manuscript, available at SSRN: https://ssrn.com/abstract=3152386.

Golez, B., and P. Koudijs (2018). Four centuries of return predictability. *Journal of Financial Economics* 127, 248–263.

Gorodnichenko, Y., and B. Lee (2017). A Note on Variance Decomposition with Local Projections. *NBER Working Paper* No. 23998.

Granger, C.W.J. (2008). Non-linear models: where do we go next – time-varying parameter models? *Studies in Nonlinear Dynamics & Econometrics* 12, 1–9.

Guo, H. (2006). On the out-of-sample predictability of stock market returns. *Journal of Business* 79, 645–670.

Guo, H., and R. F. Whitelaw (2006). Uncovering the Risk-Return Relation in the Stock Market. *Journal of Finance* 61, 1433–1463.

Hallerbach, W. G., and Menkveld, A. J. (2004). Analysing Perceived Downside Risk: the Component Value-at-Risk Framework. *European Financial Management*, 10(4), 567-591.

Hastie, T., R. Tibshirani, and J. Friedman (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Second Edition, Springer-Verlag, New York.

Henkel, S.J., J. Martin, and F. Nardari (2011). Time-varying short-horizon predictability. *Journal of Financial Economics* 99, 560–580.

Jordà, Ò. (2005). Estimation and Inference of Impulse Responses by Local Projections. *American Economic Review*, 95(1), 161–182.

J. P. Morgan, (1996). *Riskmetrics technical document. (4th ed.)*. New York.

Kofman, P., and McGlenchy, P. (2005). Structurally sound dynamic index futures hedging. *Journal of Futures Markets*, 25(12), 1173-1202.

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

Lamont, O. (1998). Earnings and expected returns. *Journal of Finance* 53, 1563–1587.

Larkin, Y., M. T. Leary, and R. Michaely (2017). Do investors value dividend-smoothing stocks differently?. *Management Science*, 63(12), 4114-4136.

LeRoy, S. F., and R. D. Porter (1981). The present-value relation: Tests based on implied variance bounds. *Econometrica*, 555-574.

Lettau, M., and S. Ludvigson (2001). Consumption, aggregate wealth, and expected stock returns. *Journal of Finance* 56, 815–849.

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

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

Marcellino, M., J. H. Stock, and M.W. Watson (2006). A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. *Journal of Econometrics*, 135(1-2), 499-526.

Martin, I. (2017). What is the expected return on the market? *The Quarterly Journal of Economics*, 132, 367–433.

Medeiros, M.C., and E. Mendes (2016). $l1$-regularization of high-dimensional time-series models with non-Gaussian and heteroskedastic errors. *Journal of Econometrics* 191, 255–271.

Medeiros, M.C., and G.F.R. Vasconcelos (2016). Forecasting macroeconomic variables in data-rich environments. *Economics Letters* 138, 50–52.

Menzly, L., T. Santos, and P. Veronesi (2004). Understanding predictability. *Journal of Political Economy* 112, 1–47.

Møller, S.V., and M. Sander (2017). Dividends, earnings and predictability. *Journal of Banking and Finance* 78, 153–163.

Owyang, M. T., Ramey, V. A., and Zubairy, S. (2013). Are government spending multipliers greater during periods of slack? Evidence from twentieth-century historical data. *American Economic Review*, 103(3), 129-34.

Ramey, V.A. (2016). Macroeconomic Shocks and Their Propagation. In H. Uhlig and J. Taylor (eds), *Handbook of Macroeconomics*. Elsevier.

Ramey, V.A., and S. Zubairy (2018). Government spending multipliers in good times and in bad: evidence from US historical data. *Journal of Political Economy*, 126(2), 850-901.

Rangvid, J., M. Schmeling, and A. Schrimpf (2014). Dividend Predictability Around the World. *Journal of Financial and Quantitative Analysis* 49, 1255–1277.

Rapach, D.E., and G. Zhou (2013). Forecasting stock returns. In eds. Elliott, G. and Timmermann, *Handbook of Economic Forecasting*, volume 2A, pages 329–383. North-Holland.

Rapach, D.E., J.K. Strauss, and G. Zhou (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *Review of Financial Studies* 23, 821–862.

Shiller, R.J. (1981). Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? *American Economic Review*, 71(3), 421–436.

Taylor, J. W. (2008). Using exponentially weighted quantile regression to estimate value at risk and expected shortfall. *Journal of Financial Econometrics*, 6(3), 382-406.

Tibshirani, R. (1996). Regression Shrinkage and Selection via the lasso. *Journal of the Royal Statistical Society. Series B*, 58, 267–288.

Timmermann, A. (2006). Forecast combinations, in G. Elliott, C.W.J. Granger, and A. Timmermann (eds.), *Handbook of Economic Forecasting*. Elsevier, Amsterdam.

Timmermann, A. (2008). Elusive return predictability. *International Journal of Forecasting* 24, 1–18.

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

Zhu, X. (2015). Tug-of-War: Time-Varying Predictability of Stock Returns and Dividend Growth. *Review of Finance* 19, 2317–2358.

Zhu, X., and J. Zhu (2013). Predicting stock returns: A regime-switching combination approach and economic links. *Journal of Banking and Finance* 37, 4120–4133.

Internet Appendix

Discount Rates and Cash Flows:
A Local Projection Approach

September 18, 2019

This Internet Appendix presents various supplemental results to the main results reported in the paper. Sections I and II present detailed estimation (regression) results for the local projections containing one and three state variables which generate the volatility decompositions reported in Sections 3.2 and 3.3 of the paper. Additional results and extensions to the LASSO analyses (Section 3.4) are presented in Section III. Together with the parameter estimates of the LASSO local projections for the monthly data, these include volatility decompositions obtained with the 'post-LASSO step' (i.e. the use of the Ordinary Least Squares in the final estimation for the LASSO-selected state variables), cross-validation-based tuning parameter selection and elastic net as an alternative to the LASSO. Finally, Section IV reports the volatility decompositions using *(i)* subsamples of the data, *(ii)* dividends that are market re-invested and cash re-invested (as opposed to dividends re-invested at the risk-free rate, as in the main paper), and *(iii)* S&P 500 index returns and dividends (as opposed to CRSP value-weighted market returns and dividends, as in the main paper).

I One state variable local projections: Estimation results

This section and Section II describe the detailed estimation (regression) results of the local projections leading to the dividend yield volatility decompositions reported in Sections 3.2 and 3.3 of the paper.

Tables I and II report the estimated local projections when the dividend yield (dp_t) is the single state variable. As introduced in equation (7) of the paper, the dividend yield at time t is thus used to predict $t + k$-period-ahead left hand side (LHS) variable which is the (discounted) cumulative return, cumulative dividend growth or the (forward) dividend yield. Similarly as in the reported volatility decompositions, the horizon k is given in annualized terms from one year up to 15 years.

As expected, the highest predictive power is obtained for short-term horizons for the dividend yield itself. The predictability of cumulative returns (in terms of the adjusted-R^2) generally increases when the horizon increases, consistent with the well-known long-horizon predictability documented in the literature (see the survey by Welch and Goyal, 2008, and the critical appraisal of this literature by Boudoukh, Richardson and Whitelaw, 2008). In line with our reported volatility decompositions (Table 2 in the paper), the predictability of dividend growth rates is very small in the case of a single state variable (dividend yield).

II Multiple state variable local projections: Estimation results

Tables III and IV present corresponding regression results as Tables I and II with the difference that now local projections contain three state variables (see equation (9) of the paper, i.e. the lagged cumulative return, cumulative dividend growth rate and dividend yield are used as predictors). Compared with the dividend yield as the single state variable, the increasing predictable patterns in the cumulative dividend growth rate are evident. This applies also to the dividend yield: lagged cumulative returns and dividend growth rates help to predict them in the long-run (when the horizon k increases).

Table I: Estimation results of the one state variable (dividend yield) local projections. Annual data (sample period 1952–2017).

This table reports the estimated regression coefficients, standard errors and adjusted-R^2s of local projections at different horizons k for the three left hand side (LHS) variables described in equation (7) of the paper. The reported standard errors are the Newey-West heteroskedasticity and autocorrelation (HAC) robust standard errors with the lag length equal to the horizon k.

k (years)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
LHS: Cumulative return															
const.	0.55	1.00	1.29	1.53	1.87	2.18	2.48	2.76	3.05	3.27	3.44	3.51	3.63	3.73	3.92
(s.e.)	(0.16)	(0.29)	(0.33)	(0.30)	(0.18)	(0.16)	(0.18)	(0.22)	(0.33)	(0.35)	(0.41)	(0.48)	(0.66)	(0.67)	(6.02)
dp	0.13	0.23	0.28	0.33	0.40	0.47	0.53	0.59	0.65	0.70	0.73	0.73	0.74	0.76	0.79
(s.e.)	(0.05)	(0.09)	(0.10)	(0.10)	(0.06)	(0.05)	(0.04)	(0.05)	(0.07)	(0.08)	(0.10)	(0.12)	(0.17)	(0.18)	(2.09)
$\text{adj} - R^2$	0.07	0.14	0.20	0.23	0.26	0.33	0.40	0.45	0.48	0.51	0.51	0.49	0.49	0.49	0.51
LHS: Cumulative dividend growth															
const.	0.78	1.47	2.12	2.74	3.35	3.94	4.52	5.10	5.65	6.20	6.71	7.17	7.66	8.12	8.58
(s.e.)	(0.06)	(0.15)	(0.19)	(0.21)	(0.19)	(0.19)	(0.17)	(0.17)	(0.18)	(0.19)	(0.20)	(0.20)	(0.23)	(0.20)	(0.21)
dp	0.01	0.01	0.00	-0.01	-0.02	-0.03	-0.04	-0.04	-0.04	-0.05	-0.06	-0.07	-0.08	-0.09	-0.10
(s.e.)	(0.02)	(0.04)	(0.06)	(0.06)	(0.05)	(0.05)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.05)	(0.06)
$\text{adj} - R^2$	-0.01	-0.01	-0.02	-0.02	-0.01	-0.00	0.00	0.00	0.01	0.01	0.02	0.05	0.06	0.08	0.09
LHS: Dividend yield															
const.	-0.25	-0.52	-0.70	-0.84	-1.07	-1.29	-1.46	-1.60	-1.78	-1.86	-1.94	-2.00	-2.04	-2.07	-2.17
(s.e.)	(0.17)	(0.35)	(0.46)	(0.47)	(0.40)	(0.34)	(0.31)	(0.29)	(0.32)	(0.41)	(0.52)	(0.62)	(0.72)	(0.58)	(0.61
dp	0.90	0.80	0.73	0.66	0.57	0.49	0.42	0.36	0.28	0.24	0.19	0.15	0.12	0.09	0.04
(s.e.)	(0.05)	(0.10)	(0.14)	(0.15)	(0.13)	(0.11)	(0.10)	(0.07)	(0.06)	(0.07)	(0.09)	(0.12)	(0.14)	(0.12)	(0.15)
$\text{adj} - R^2$	0.89	0.78	0.69	0.61	0.48	0.37	0.28	0.21	0.13	0.09	0.06	0.03	0.01	0.00	-0.02

Table II: Estimation results of the one state variable (dividend yield) local projections. Monthly data (sample period 1952–2017).

This table reports the estimated regression coefficients, standard errors and adjusted-R^2s of local projections at different horizons k for the three left hand side (LHS) variables described in equation (7) of the paper. The reported standard errors are the Newey-West heteroskedasticity and autocorrelation (HAC) robust standard errors with the lag length equal to the horizon k.

k (months)	12	24	36	48	60	72	84	96	108	120	132	144	156	168	180
LHS: Cumulative return															
const.	0.51	0.94	1.25	1.54	1.84	2.12	2.41	2.73	2.99	3.17	3.33	3.43	3.50	3.62	3.82
(s.e.)	(0.29)	(0.58)	(0.64)	(0.51)	(0.37)	(0.24)	(0.20)	(0.26)	(0.32)	(0.35)	(0.41)	(0.54)	(0.56)	(0.70)	(1.46)
dp	0.12	0.21	0.28	0.33	0.40	0.45	0.51	0.58	0.64	0.68	0.70	0.71	0.71	0.73	0.77
(s.e.)	(0.08)	(0.18)	(0.20)	(0.16)	(0.12)	(0.08)	(0.05)	(0.06)	(0.08)	(0.09)	(0.11)	(0.15)	(0.15)	(0.19)	(0.45)
$\mathrm{adj}-R^2$	0.08	0.15	0.21	0.25	0.30	0.35	0.42	0.49	0.51	0.52	0.52	0.51	0.49	0.50	0.52
LHS: Cumulative dividend growth															
const.	0.14	0.28	0.39	0.43	0.42	0.46	0.51	0.58	0.65	0.72	0.74	0.72	0.73	0.76	0.78
(s.e.)	(0.24)	(0.43)	(0.53)	(0.55)	(0.38)	(0.32)	(0.28)	(0.26)	(0.33)	(0.46)	(0.53)	(0.53)	(0.43)	(0.42)	(0.36)
dp	0.00	0.02	0.03	0.02	0.00	0.00	0.00	0.00	0.00	0.00	-0.01	-0.03	-0.04	-0.04	-0.05
(s.e.)	(0.07)	(0.13)	(0.16)	(0.17)	(0.11)	(0.09)	(0.07)	(0.06)	(0.09)	(0.12)	(0.14)	(0.14)	(0.11)	(0.11)	(0.09)
$\mathrm{adj}-R^2$	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01
LHS: Dividend yield															
const.	-0.27	-0.51	-0.69	-0.85	-1.09	-1.29	-1.47	-1.62	-1.77	-1.86	-1.94	-2.00	-2.04	-2.07	-2.17
(s.e.)	(0.66)	(1.52)	(1.45)	(1.57)	(1.17)	(0.99)	(0.74)	(0.88)	(0.57)	(0.52)	(0.51)	(1.49)	(1.55)	(0.51)	(0.69)
dp	0.90	0.80	0.73	0.66	0.57	0.49	0.41	0.35	0.28	0.23	0.19	0.15	0.12	0.09	0.04
(s.e.)	(0.19)	(0.46)	(0.43)	(0.48)	(0.40)	(0.35)	(0.24)	(0.33)	(0.18)	(0.13)	(0.11)	(0.21)	(0.23)	(0.14)	(0.20)
$\mathrm{adj}-R^2$	0.89	0.78	0.70	0.61	0.49	0.38	0.29	0.22	0.15	0.11	0.07	0.05	0.03	0.02	0.00

Table III: Estimation results of the three state variable local projections. Annual data (sample period 1952–2017).

This table reports the estimated regression coefficients, standard errors and adjusted-R^2s of local projections at different horizons k for the three left hand side (LHS) variables described in equation (7) of the paper. The reported standard errors are the Newey-West heteroskedasticity and autocorrelation (HAC) robust standard errors with the lag length equal to the horizon k.

k (years)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
LHS: Cumulative return															
const	0.62	0.98	1.29	1.91	2.39	2.59	2.76	3.64	3.70	3.64	3.26	3.17	3.30	2.64	3.01
(s.e.)	(0.21)	(0.43)	(0.41)	(0.38)	(0.37)	(0.45)	(0.51)	(0.56)	(0.92)	(0.71)	(0.81)	(1.04)	(0.98)	(0.83)	(1.31)
$\sum_{j=1}^{k} \rho^{j-1} r_{t+j-k}$	-0.15	-0.29	-0.12	-0.05	-0.05	0.00	0.00	-0.04	-0.08	0.00	-0.07	-0.16	-0.25	-0.43	-0.45
(s.e.)	(0.09)	(0.11)	(0.13)	(0.13)	(0.10)	(0.10)	(0.11)	(0.15)	(0.32)	(0.23)	(0.35)	(0.27)	(0.19)	(0.14)	(0.6)
$\sum_{j=1}^{k} \rho^{j-1} \Delta d_{t+j-k}$	-0.33	0.28	-0.13	-0.60	-0.63	-0.46	-0.28	-0.66	-0.41	-0.14	0.17	0.25	0.21	0.51	0.40
(s.e.)	(0.26)	(0.16)	(0.26)	(0.33)	(0.40)	(0.44)	(0.40)	(0.39)	(0.42)	(0.37)	(0.64)	(0.76)	(0.63)	(0.50)	(1.57)
dp_t	0.14	0.22	0.27	0.38	0.48	0.53	0.57	0.73	0.75	0.77	0.69	0.62	0.60	0.39	0.44
(s.e)	(0.06)	(0.12)	(0.11)	(0.08)	(0.07)	(0.08)	(0.10)	(0.13)	(0.27)	(0.19)	(0.20)	(0.17)	(0.17)	(0.16)	(0.29)
$\text{adj} - R^2$	0.08	0.21	0.15	0.26	0.29	0.31	0.35	0.45	0.45	0.48	0.49	0.53	0.59	0.73	0.81
LHS: Cumulative dividend growth															
const	-0.02	0.27	0.53	0.74	0.80	0.82	0.84	1.03	1.40	1.43	1.46	1.80	2.00	1.69	2.01
(s.e.)	(0.09)	(0.22)	(0.24)	(0.27)	(0.37)	(0.56)	(0.84)	(0.66)	(0.59)	(1.78)	(0.43)	(0.42)	(0.41)	(0.45)	(0.61)
$\sum_{j=1}^{k} \rho^{j-1} r_{t+j-k}$	0.27	0.06	-0.16	-0.25	-0.14	-0.14	-0.14	-0.15	-0.18	-0.27	-0.30	-0.27	-0.19	-0.16	-0.08
(s.e.)	(0.06)	(0.10)	(0.08)	(0.06)	(0.08)	(0.12)	(0.15)	(0.10)	(0.16)	(0.27)	(0.07)	(0.08)	(0.07)	(0.08)	(0.10)
$\sum_{j=1}^{k} \rho^{j-1} \Delta d_{t+j-k}$	0.20	-0.23	-0.35	-0.45	-0.50	-0.41	-0.34	-0.43	-0.63	-0.62	-0.62	-0.78	-0.83	-0.66	-0.79
(s.e.)	(0.09)	(0.12)	(0.15)	(0.20)	(0.16)	(0.24)	(0.45)	(0.36)	(0.32)	(0.88)	(0.21)	(0.12)	(0.09)	(0.21)	(0.29)
dp_t	-0.02	0.03	0.05	0.07	0.06	0.04	0.03	0.04	0.08	0.04	0.00	0.04	0.08	0.01	0.08
(s.e.)	(0.03)	(0.07)	(0.07)	(0.07)	(0.10)	(0.14)	(0.17)	(0.13)	(0.14)	(0.42)	(0.10)	(0.11)	(0.10)	(0.09)	(0.09)
$\text{adj} - R^2$	0.32	0.02	0.20	0.47	0.27	0.15	0.09	0.15	0.27	0.41	0.45	0.49	0.41	0.28	0.24
LHS: Dividend yield															
const	-0.25	-0.53	-0.58	-0.85	-1.25	-1.44	-1.51	-2.06	-2.24	-2.16	-1.80	-1.38	-0.87	-0.30	-0.47
(s.e.)	(0.13)	(0.38)	(0.45)	(0.51)	(0.55)	(0.76)	(1.10)	(1.12)	(227.53)	(2.38)	(0.89)	(0.70)	(0.96)	(2.29)	(0.74)
$\sum_{j=1}^{k} \rho^{j-1} r_{t+j-k}$	-0.21	0.03	-0.21	-0.32	-0.28	-0.28	-0.28	-0.27	-0.29	-0.47	-0.47	-0.41	-0.25	-0.05	-0.01
(s.e.)	(0.08)	(0.12)	(0.10)	(0.14)	(0.16)	(0.13)	(0.18)	(0.31)	(77.96)	(0.45)	(0.26)	(0.22)	(0.24)	(0.19)	(0.12)
$\sum_{j=1}^{k} \rho^{j-1} \Delta d_{t+j-k}$	0.43	0.02	0.11	0.38	0.48	0.37	0.22	0.45	0.32	0.04	-0.30	-0.57	-0.79	-1.01	-0.98
(s.e.)	(0.20)	(0.15)	(0.18)	(0.27)	(0.5)	(0.66)	(0.85)	(0.71)	(135.28)	(1.12)	(0.38)	(0.44)	(0.78)	(1.43)	(0.49)
dp_t	0.91	0.80	0.75	0.66	0.53	0.45	0.38	0.24	0.14	0.04	0.04	0.08	0.20	0.35	0.29
(s.e.)	(0.04)	(0.11)	(0.13)	(0.14)	(0.13)	(0.14)	(0.20)	(0.21)	(58.90)	(0.60)	(0.25)	(0.19)	(0.16)	(0.26)	(0.12)
$\text{adj} - R^2$	0.91	0.76	0.70	0.65	0.54	0.44	0.36	0.31	0.23	0.27	0.25	0.21	0.14	0.12	0.09

Table IV: Estimation results of the three state variable local projections. Monthly data (sample period 1952–2017).

This table reports the estimated regression coefficients, standard errors and adjusted-R^2s of local projections at different horizons k for the three left hand side (LHS) variables described in equation (7) of the paper. The reported standard errors are the Newey-West heteroskedasticity and autocorrelation (HAC) robust standard errors with the lag length equal to the horizon k.

k (months)	12	24	36	48	60	72	84	96	108	120	132	144	156	168	180
LHS: Cumulative return															
const	0.59	0.98	1.45	1.99	2.32	2.34	2.71	3.35	3.40	3.27	3.23	3.23	2.86	2.44	2.73
(s.e.)	(0.28)	(0.74)	(0.63)	(0.38)	(0.45)	(0.56)	(0.62)	(0.68)	(1.13)	(0.65)	(0.88)	(1.24)	(1.05)	(0.58)	(0.67)
$\sum_{j=1}^{k} \rho^{j-1} r_{t+j-k}$	-0.14	-0.22	-0.11	-0.08	-0.07	0.00	0.01	0.00	-0.02	-0.01	-0.09	-0.17	-0.31	-0.44	-0.48
(s.e.)	(0.15)	(0.12)	(0.14)	(0.16)	(0.15)	(0.11)	(0.11)	(0.16)	(0.78)	(0.22)	(0.27)	(0.21)	(0.17)	(0.12)	(0.15)
$\sum_{j=1}^{k} \rho^{j-1} \Delta d_{t+j-k}$	-0.04	-0.06	-0.39	-0.57	-0.49	-0.21	-0.25	-0.40	-0.20	0.04	0.13	0.18	0.36	0.58	0.49
(s.e.)	(0.16)	(0.35)	(0.28)	(0.34)	(0.55)	(0.54)	(0.47)	(0.43)	(0.78)	(0.32)	(0.58)	(0.8)	(0.58)	(0.34)	(0.21)
dp_t	0.13	0.21	0.29	0.40	0.46	0.48	0.56	0.69	0.71	0.70	0.67	0.63	0.50	0.36	0.38
(s.e.)	(0.08)	(0.21)	(0.17)	(0.09)	(0.09)	(0.09)	(0.11)	(0.14)	(0.37)	(0.17)	(0.18)	(0.21)	(0.20)	(0.12)	(0.2)
adj $- R^2$	0.11	0.18	0.24	0.32	0.31	0.33	0.39	0.48	0.49	0.52	0.53	0.58	0.64	0.77	0.84
LHS: Cumulative dividend growth															
const	-0.03	0.27	0.59	0.82	0.87	0.81	0.87	1.07	1.32	1.14	1.28	1.55	1.83	1.71	2.25
(s.e.)	(0.13)	(0.38)	(0.41)	(0.27)	(0.44)	(0.69)	(0.82)	(0.59)	(0.45)	(2.27)	(0.57)	(0.50)	(0.37)	(0.42)	(0.55)
$\sum_{j=1}^{k} \rho^{j-1} r_{t+j-k}$	0.23	0.17	-0.10	-0.28	-0.22	-0.18	-0.17	-0.19	-0.26	-0.32	-0.33	-0.30	-0.23	-0.20	-0.12
(s.e.)	(0.15)	(0.23)	(0.18)	(0.09)	(0.12)	(0.14)	(0.14)	(0.22)	(0.13)	(0.35)	(0.08)	(0.09)	(0.08)	(0.08)	(0.07)
$\sum_{j=1}^{k} \rho^{j-1} \Delta d_{t+j-k}$	0.49	0.02	-0.26	-0.31	-0.34	-0.23	-0.22	-0.32	-0.45	-0.33	-0.42	-0.55	-0.67	-0.60	-0.82
(s.e.)	(0.14)	(0.3)	(0.34)	(0.22)	(0.21)	(0.33)	(0.39)	(0.33)	(0.24)	(1.21)	(0.22)	(0.09)	(0.13)	(0.19)	(0.23)
dp_t	-0.02	0.02	0.06	0.07	0.06	0.04	0.03	0.04	0.05	-0.03	-0.04	-0.01	0.04	0.00	0.10
(s.e.)	(0.03)	(0.13)	(0.12)	(0.07)	(0.11)	(0.16)	(0.17)	(0.10)	(0.09)	(0.51)	(0.14)	(0.14)	(0.10)	(0.10)	(0.10)
adj $- R^2$	0.28	0.08	0.11	0.35	0.29	0.16	0.16	0.20	0.34	0.41	0.46	0.46	0.42	0.34	0.34
LHS: Dividend yield															
const	-0.25	-0.50	-0.61	-0.88	-1.24	-1.47	-1.64	-2.12	-2.05	-2.10	-1.76	-1.26	-0.82	-0.20	-0.01
(s.e.)	(0.30)	(0.83)	(0.86)	(0.86)	(1.02)	(1.67)	(1.65)	(1.54)	(6.39)	(4.7)	(1.05)	(1.02)	(1.12)	(1.93)	(0.68)
$\sum_{j=1}^{k} \rho^{j-1} r_{t+j-k}$	-0.26	0.01	-0.23	-0.32	-0.30	-0.28	-0.29	-0.32	-0.40	-0.54	-0.51	-0.43	-0.26	-0.09	0.00
(s.e.)	(0.18)	(0.18)	(0.15)	(0.17)	(0.19)	(0.21)	(0.29)	(0.34)	(2.64)	(0.81)	(0.33)	(0.22)	(0.21)	(0.14)	(0.11)
$\sum_{j=1}^{k} \rho^{j-1} \Delta d_{t+j-k}$	0.14	-0.01	0.17	0.35	0.41	0.37	0.31	0.46	0.18	0.01	-0.28	-0.59	-0.79	-1.08	-1.19
(s.e.)	(0.18)	(0.31)	(0.24)	(0.45)	(1.22)	(1.33)	(1.18)	(0.90)	(0.85)	(2.48)	(0.45)	(0.79)	(0.86)	(1.06)	(0.29)
dp_t	0.90	0.81	0.75	0.65	0.53	0.44	0.36	0.22	0.14	0.03	0.04	0.10	0.20	0.33	0.35
(s.e.)	(0.08)	(0.24)	(0.26)	(0.24)	(0.17)	(0.27)	(0.26)	(0.27)	(1.02)	(1.08)	(0.29)	(0.18)	(0.18)	(0.26)	(0.16)
adj $- R^2$	0.90	0.77	0.71	0.67	0.56	0.46	0.39	0.37	0.32	0.36	0.34	0.31	0.24	0.23	0.22

III LASSO extensions

III.I LASSO state variable selection: Monthly data

Table V reports the LASSO estimation and model selection results for the monthly data. When comparing the corresponding results for the annual data (see Table 5), we obtain clearly less exclusions (i.e. state variables completely left out from the resulting LASSO estimates and local projections) than with the annual data. In this monthly case, the most exclusions happen largely for variables that are also excluded in the annual case, including short-term interest rates (levels), which have been found useful to predict stock returns in Ang and Bekaert (2007). However, it is important to note that the relative risk-free rate ($RREL$) seems systematically an important state variable for all the LHS variables, so the information transmission from the (short-term) interest rates is coming strongest through the relative short-rate instead of the levels.

III.II LASSO: Shrinkage also in dividend yield

In Table VI, we consider the change to the previous LASSO setting so that also the dividend yield as a state variable is subject to possible shrinkage (penalization) in the local projections. In practice, the dividend yield turns out to be consistently involved in the estimated local projections for all the LHS variables in the monthly data and especially for expected returns in the annual data frequency. These findings, in addition to strong theoretical backing as discussed in Section 2 of the paper, are the reasons why the main analyses have been carried out by always including the lagged dividend yield. Given these findings, the volatility decomposition reported in Table VI is, as expected, very similar to the decomposition reported in Section 3.4 (Table 4).[1]

[1] At the one-year horizon ($k = 1$, annual data) the discount rate variation, $\widehat{\sigma}(r, 1)$, is zero since all state variables are excluded by the LASSO model selection. The cash flow–discount rate ratio (equation (13)) is thus not defined as all volatility is coming from the cash flow variation.

Table V: LASSO model selection and estimation results for the monthly data (sample period 1952–2017). Empty cells indicate the exclusion of those state variables. The results are reported for the three left hand side (LHS) variables. See also the notes to Table 5 in the main paper.

k (months)	const	r	Δd	dp	EP	DE	$SVAR$	BM	$NTIS$	TBL	LTY	$RFREE$	TMS	LTR	DFY	DFR	$CORPR$	$INFL$	CAY	$RREL$
LHS: Cumulative return																				
12	0.33	-0.16	0.06	0.01	0.09	-0.01	-1.26	0.20	0.21	-2.33		-0.00			7.00	0.14	0.35	-4.02	3.13	1.48
24	1.05	-0.24	0.25	-0.12	0.41	0.26			1.22	-2.14		-0.00	2.53		7.33		0.12	-1.83	4.55	4.29
36	2.03	-0.11	-0.01	-0.50	1.01	0.76	-3.06	-0.23	1.11	-2.86		-0.01	2.93		11.44	0.26	0.27	3.08	5.41	4.66
48	1.91	-0.07	-0.37	-0.29	0.74	0.59			0.12	-2.01		-0.01	2.46		12.42	0.27		3.34	6.13	3.94
60	1.55	-0.11	-0.30	-0.34	0.69	0.64		0.21		-1.50		-0.00	2.67		16.24		0.13	3.99	7.53	3.85
72	1.32	-0.18		-0.20	0.51	0.51	3.54	0.22	1.09				4.06	0.12	10.54	-0.09		7.63	8.43	2.84
84	1.28	-0.20	-0.25	-0.06	0.32	0.34	2.78	0.39			1.62		3.36	0.41	5.51			7.86	8.25	2.10
96	1.11	-0.31	-0.71	0.28	-0.16	0.04	2.77	0.71	-1.42		4.64		2.14	0.34	0.49		0.02	5.38	8.88	2.00
100	2.45	-0.31	-0.61	0.43	0.03		5.19	0.26	-2.49		3.94		4.53		3.36		0.62	6.60	7.08	2.97
120	3.92	-0.25	-0.42	0.11	0.76	0.41	7.59	-0.32	-3.28		1.83		7.42	0.35	12.56		0.23	10.87	4.94	5.17
132	5.09	-0.06	-0.61	0.07	1.14	0.54	7.31	-0.68	-4.59		1.42		10.88	0.51	23.53	-0.29		11.96	0.98	7.37
144	3.15	-0.12	-0.16	-0.11	0.81	0.61	6.65		-4.47		0.39		11.34	0.31	17.53	-0.66		12.04	2.59	8.53
156	2.02	-0.24	0.29	-0.32	0.71	0.70	5.69	0.19	-3.49		0.55		12.15	0.70	10.27	-0.24		11.42	2.58	8.35
168	1.14	-0.44	0.46	-0.19	0.26	0.39	-1.64	0.57	-1.72		0.56		8.63	0.45		-0.16		2.34	4.37	5.15
180	1.36	-0.29	0.04	-0.12	0.20	0.39	-1.35	0.84	-3.17		2.31		5.70	0.54	-4.91		0.19	1.11	3.39	2.35
LHS: Cumulative dividend growth																				
12	-0.09	0.08	0.33	-0.06	0.05	-0.11		0.10	0.75	-0.53		-0.00	1.87		-2.93					3.40
24	-1.19	0.16	0.08	-0.06	-0.22	-0.31	-1.01	0.77	0.54	-2.77		-0.02	2.41		-1.58	0.38	0.26	0.74	1.73	5.30
36	-1.18	-0.01	-0.22	-0.12	-0.19	-0.30	-1.40	0.92		-2.79		-0.01	2.20		-3.83	0.24	0.15	0.37	1.95	4.65
48	-0.67	-0.15	-0.24	-0.24		-0.05	-1.58	0.71		-1.60		-0.01	2.08		-5.83	0.09	0.12		1.08	2.90
60	-1.21	0.01	-0.39	-0.39			-0.89	0.96	-0.43	-1.92		-0.00	2.40		-3.95		0.07	0.57	0.26	3.10
72	-1.51	0.12	-0.38	-0.30	-0.17	-0.14	-0.51	1.10	-0.25	-2.50		-0.00	1.75		-1.10	0.11	0.15	1.68	-0.78	3.03
84	-0.94	0.07	-0.36	-0.24	-0.11	-0.12		0.84	0.39	-1.63		-0.00	0.12	-0.08	0.86	0.05		-0.17	-1.64	1.73
96	0.10		-0.38	-0.13		-0.10		0.41	0.89	-0.66	-0.08	-0.10					0.06		-2.72	0.85
108	0.59	-0.03	-0.20	0.17	-0.17	-0.38	1.79	0.21	2.40	-1.54		-0.00	0.48	0.10					-3.04	1.67
120	1.18	-0.09	-0.06	0.10		-0.18			2.08	-2.10		-0.32			-2.68				-2.83	1.42
132	2.11	-0.12	-0.27	0.02	0.24			-0.13	0.79	-3.10		-0.46	0.94	0.02	-4.58		0.19		-3.46	2.91
144	3.02	-0.10	-0.38	-0.13	0.58	0.26	-1.76	-0.41	0.32	-2.99		-10.54	0.83		-0.75		0.14	-0.58	-4.54	2.73
156	2.39	-0.11	-0.21	-0.12	0.38	0.23	-1.87	-0.23	1.17	-4.18	-0.27	-0.49		0.06	1.06			-0.54	-3.88	2.68
168	1.87	-0.07	-0.16	-0.20	0.34	0.27	-1.93	-0.09	0.83	-4.12		-0.07		0.13		-0.02		1.59	-3.96	2.25
180	1.74		-0.29	-0.11	0.21	0.20	-0.84		0.98	-2.49		-0.18	0.41	0.06	-3.32	-0.17		4.13	-4.06	1.14
LHS: Dividend yield																				
12	-0.15	-0.08	0.00	0.36	0.58	0.49	1.96		-0.67		0.03		1.07		-4.63		-0.23	2.43	-1.28	2.17
24	-1.18	0.08	-0.22	0.65			1.71	0.34	-0.81						-2.77		-0.04	2.21	-2.49	1.99
36	-1.72	-0.04	-0.08	0.50	-0.02			0.55	-1.28	-0.38					-3.86				-3.10	0.36
48	-2.47	-0.07		0.39	-0.15			0.79	-2.44						-8.65			-1.52	-4.08	
60	-2.50			0.27	-0.05		0.54	0.68	-1.44	-0.43					-6.81			-1.16	-5.48	
72	-3.42	0.19	-0.23	0.36	-0.40	-0.32		1.16	-3.25	-1.38	-0.35	-0.28			-7.26	0.31		-5.01	-7.28	0.74
84	-2.48	0.15	-0.06	0.18				0.58	-1.27	-0.24	-1.45	-0.02				0.21		-5.97	-8.67	
96	-2.55	0.10	0.28	0.16		-0.04		0.38			-2.29					0.14		-5.00	-8.77	-0.45
108	-2.81	0.19	0.37	-0.05	0.16			0.42	1.79		-3.83		-0.51	-0.22	2.69			-4.79	-8.69	-0.52
120	-2.74		0.25	0.00		0.08	-2.87	0.38	2.07		-2.57		-1.67	-0.09	-1.92	0.13		-6.73	-7.59	
132	-3.13	-0.13	0.27	-0.21		0.26	-4.97	0.70	2.60		-4.01		-3.86	-0.34	-11.07	0.14		-9.72	-4.42	-2.43
144	-1.31	-0.23	-0.17	0.11		0.11	-3.41	0.19	2.59		-2.93		-5.97	-0.23	-15.79	0.54		-9.77	-4.41	-5.28
156	-0.73	-0.02	-0.41	0.23	0.05		-5.08	0.16	3.03		-4.22		-5.97	-0.45	-7.48	0.29		-8.44	-4.49	-4.75
168	-0.10	0.16	-0.60	0.23	0.25	0.09			2.28		-4.80		-4.23	-0.35					-5.01	-3.64
180	0.25	0.12	-0.30	-0.00	0.60	0.30	0.26	-0.47	2.68		-5.62		-1.41	-0.39	5.33		-0.18	4.96	-5.70	-1.49

Table VI: LASSO-based volatility decomposition when the dividend yield is also subject to shrinkage.

This table reports the annual (Panel A) and monthly (Panel B) volatility decomposition using the LASSO estimator for different horizons k where the dividend yield as a state variable is also subject to shrinkage (cf. Table 4 in the paper).

A: LASSO - Annual					B: LASSO - Monthly				
k (years)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	k (months)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
1	0.00	0.11	0.90	–	12	0.23	0.24	0.94	1.07
2	0.43	0.19	0.81	0.44	24	0.39	0.28	0.85	0.72
3	0.43	0.24	0.79	0.55	36	0.52	0.31	0.82	0.59
4	0.61	0.22	0.74	0.35	48	0.62	0.30	0.81	0.49
5	0.75	0.30	0.61	0.40	60	0.73	0.30	0.76	0.42
6	0.70	0.30	0.64	0.43	72	0.80	0.32	0.79	0.40
7	0.73	0.30	0.62	0.40	84	0.83	0.31	0.75	0.37
8	0.84	0.33	0.68	0.39	96	0.94	0.34	0.74	0.37
9	0.93	0.30	0.74	0.32	108	0.92	0.39	0.73	0.43
10	0.92	0.28	0.69	0.30	120	0.91	0.40	0.68	0.44
11	0.93	0.32	0.63	0.34	132	0.93	0.47	0.68	0.51
12	0.95	0.50	0.61	0.53	144	0.93	0.53	0.65	0.58
13	0.85	0.39	0.57	0.45	156	0.94	0.53	0.61	0.56
14	0.93	0.37	0.55	0.40	168	0.89	0.49	0.54	0.55
15	0.91	0.39	0.59	0.43	180	0.95	0.40	0.55	0.43

III.III Cross-validation-based tuning parameters

For cross-sectional datasets, the tuning parameter λ in the LASSO estimator (equation (15)) is often determined by using cross-validation. As argued in Section 2.3, our main results are based on the use of the BIC to determine λ, which follows the conventional use of information criteria in financial econometrics.

Tables VII–VIII report the estimated coefficients in the annual and monthly local projections and the resulting volatility decompositions are presented in Table IX. The results are based on the 10-fold cross-validation instead of the BIC as employed in the results of Section 3.4. It turns out that the local projections (inclusions and exclusions of the state variables) are in large extent the same as reported in Table 5 and Table V above (slightly more inclusions due to a typically somewhat smaller shrinkage than in the BIC case). The resulting volatility decompositions in Table IX are hence the same in their main conclusions as in Table 4 (Section 3.4).

Table VII: Cross-validation-based LASSO estimation and model selection results (annual data).

In this table, empty cells indicate the exclusion of those state variables in the LASSO estimation results (10-fold cross-validation with the value of the parameter λ at which the minimum mean square error is achieved). The results are reported for the three left hand side (LHS) variables at the time and for all horizons k (years).

k (years)	const	r	Δd	dp	EP	DE	$SVAR$	BM	$NTIS$	TBL	LTY	$RFREE$	TMS	LTR	DFY	DFR	$CORPR$	$INFL$	CAY	$EQIS$	IK
LHS: Cumulative return																					
1	0.77			0.18			0.45						0.09						0.23	-0.33	
2	1.56	-0.25	0.42	0.21	0.20		0.94		2.45	-2.78		-0.00	0.44		6.52				4.64	-0.34	
3	3.24	-0.26	0.10	0.04	0.90	0.43	0.58	-0.83	1.81	-6.66	-0.56	-0.02			7.12	0.17	0.31	4.24	7.09	0.58	13.22
4	2.46	-0.15	-0.32	0.34	0.25		1.39		0.06	-1.98		-0.00	1.12						6.04		
5	2.60	-0.11	-0.32	0.57			1.88						2.26						4.53	-0.16	-0.96
6	2.79	-0.12	-0.20	0.53			1.51						3.54			-0.83		1.92	5.94		-11.21
7	2.76	-0.08		0.51		-0.02	0.47						5.28			-0.80		2.90	5.66		-14.62
8	3.54	-0.18	-0.50	0.56			1.29		-0.22		2.42		3.11			-0.25		2.67	6.03		-23.11
9	4.23	-0.22	-0.68	0.47	0.11		2.06		-2.93		4.39		0.86	0.06		-0.10		1.83	4.86	-0.04	-36.29
10	5.00	-0.17	-0.47	0.37	0.43	0.00	3.42	-0.31	-4.01		3.08		4.89	0.44	-14.84	-0.16		4.51	3.49	0.25	-39.82
11	5.20	-0.11	-0.12	0.03	0.97	0.44	3.17	-0.85	-2.72		3.83		7.31	0.42	-6.29	-0.69		5.21		-0.11	-29.95
12	3.83	-0.24	-0.01	-0.11	0.58	0.41	4.13		-2.31		7.76		4.51	0.38	-22.08	-1.45		1.97		-0.56	-40.75
13	4.32	-0.36	-0.22	-0.16	0.63	0.62	3.93		-4.93		7.45		4.08	0.23	-24.57	-0.55		1.36		-0.38	-36.85
14	3.27	-0.44	0.15	0.38		0.09	1.20		-2.51		3.33		7.59	0.16	-14.13	-1.50		2.83	1.17		-17.31
15	3.41	-0.35		0.47					-3.23		2.34		5.71	0.04		-0.41		1.64	0.39		-9.12
LHS: Cumulative dividend growth																					
1	-0.14	0.35	0.12	-0.45	0.36	0.26	0.78	0.17	0.75	0.50	0.21	0.00			-4.97	0.07	0.05	-0.81	-0.82	-0.15	-7.18
2	-0.39	0.04	-0.13	-0.13		-0.15	0.17	0.48	2.40				0.13		-0.01	0.31	0.15	-1.80	-0.05	-0.49	-4.31
3	-0.12	-0.08	-0.22	-0.06		-0.09		0.33	1.79	-1.03		-0.00	1.03		-2.24	0.37	0.20			-0.19	
4	-0.02	-0.20	-0.39	-0.10		-0.00		0.28	0.80	-0.00		-0.00	0.04		-2.84	0.38	0.12				
5	-0.75	-0.00	-0.45	-0.30		0.00		0.60	0.51	-0.81		-0.00	0.27			0.43			-0.00	0.15	-2.79
6	-1.42	0.10	-0.40	-0.52		0.08		0.98		-0.60	-0.72	-0.23			-0.44	0.83	0.17		-0.87	0.05	-7.30
7	-0.93	0.05	-0.39	-0.39	-0.01		0.00	0.73	0.44		-1.59				2.34	0.56	0.01		-0.95	0.31	-5.85
8	-0.21		-0.50	-0.23		-0.05	0.15	0.35	0.67		-0.76		-0.99		1.69	0.32	0.10	0.32	-2.03	0.56	-2.45
9	0.59		-0.51	-0.05		-0.03		0.10	0.86							0.04	0.02	0.14	-3.36	0.28	
10	0.83	-0.05	-0.28	-0.01					1.70										-3.03		
11	1.47	-0.15	-0.47	0.05					0.23	-0.79	-1.47	-0.02			-2.29	0.15	0.07		-2.15	0.40	
12	3.62	-0.15	-0.93	0.01	0.56	0.34	1.33	-0.63			-0.76	-1.12			0.36	0.26	0.13	-1.05	-4.71	0.38	6.51
13	1.47		-0.43	0.06		0.03	-0.13		1.09	-1.30	-0.00								-3.93		
14	0.68			-0.08			-0.14		1.95	-1.43	-0.41	-0.00							-3.89		
15	0.91		-0.09	-0.08		0.07	-0.31		1.33	-0.48	-1.47	-0.02							-3.95		
LHS: Dividend yield																					
1	-0.43	-0.13	0.27	0.87													-0.10	0.14	-0.43	0.25	
2	-0.77	0.00		0.75						0.69									-2.81	0.06	
3	-1.40			0.57				0.26											-3.45		
4	-2.09			0.38		0.05		0.43	-0.64										-4.45		2.03
5	-2.31		0.16	0.33				0.26											-5.47	0.22	4.42
6	-1.91			0.40															-6.99	0.33	7.10
7	-2.26			0.30				0.02											-7.44	0.46	7.99
8	-2.55			0.22									-0.67						-7.78	0.48	11.45
9	-3.15	0.23	0.37	-0.25	0.64	0.31	-0.39		1.47	-0.97	-9.10				21.99	1.42	0.14	0.56	-6.13	1.00	36.87
10	-3.89	0.01	0.09	-0.11	0.15	0.21	-0.67	0.44	2.94		-3.32		-2.38	-0.12	12.47	0.69		-3.09	-5.98	0.23	35.86
11	-3.09	-0.03		0.14		0.17	-0.87	0.19	2.44		-4.42			-0.01		0.49			-5.13	0.34	33.88
12	-2.25	-0.09	-0.10	0.29		0.09	-0.76		2.97		-6.28		-0.12			0.69			-3.54	0.51	35.00
13	-3.18	0.16	-0.04	-0.48	0.49	0.32	-4.64	0.72	3.98		-8.98		-4.10	-0.17	22.61	1.73		-2.80	-1.42	0.08	18.79
14	-1.42	0.08	-0.27	0.37		-0.03	-0.99		3.24		-7.18		-1.42	-0.26	12.21	0.32			-3.23		21.78
15	-1.56			0.11					1.68		-5.22		-2.13	-0.35			-0.04		-3.38		

Table VIII: Cross-validation-based LASSO estimation and model selection results (monthly data).

In this table, empty cells indicate the exclusion of those state variables in the LASSO estimation results (10-fold cross-validation with the value of the parameter λ at which the minimum mean square error is achieved). The results are reported for the three left hand side (LHS) variables at the time and for all horizons k (months).

k (monthly)	const	r	Δd	dp	EP	DE	$SVAR$	BM	$NTIS$	TBL	LTY	$RFREE$	TMS	LTR	DFY	DFR	$CORPR$	$INFL$	CAY	$RREL$
LHS: Cumulative return																				
12	0.35	-0.16	0.06	0.02	0.09	-0.01	-1.19	0.20	0.20	-2.30		-0.00			6.90	0.13	0.35	-3.97	3.09	1.43
24	1.06	-0.23	0.26	-0.19	0.48	0.33	-0.08		1.26	-2.32		-0.00	2.55		7.78	0.07	0.14	-1.86	4.66	4.48
36	2.02	-0.11	-0.01	-0.49	1.01	0.76	-3.04	-0.23	1.11	-2.85		-0.01	2.93		11.41	0.26	0.27	3.07	5.41	4.66
48	2.01	-0.05	-0.37	-0.44	0.93	0.75	-1.47	-0.03	0.12	-2.51		-0.02	2.45		14.18	0.38	0.02	4.14	6.25	4.46
60	1.58	-0.07	-0.33	-0.56	0.93	0.86	-1.72	0.24	-0.42	-2.28		-0.01	2.72		17.57	0.04	0.21	5.01	7.69	4.78
72	1.42	-0.14	-0.04	-0.42	0.75	0.72	2.16	0.23	0.95	-0.89		-0.00	3.78	0.16	13.12	-0.10		9.05	8.66	3.64
84	1.37	-0.19	-0.30	-0.21	0.48	0.49	1.86	0.38	-0.30		1.33		3.70	0.37	6.76		0.08	8.37	8.23	2.69
96	1.11	-0.31	-0.71	0.28	-0.15	0.04	2.76	0.71	-1.41		4.64		2.14	0.34	0.52		0.02	5.39	8.87	2.00
108	2.49	-0.30	-0.67	0.31	0.16	0.12	4.46	0.28	-2.92		3.72		4.76		4.47	0.11	0.64	6.98	7.16	3.44
120	3.86	-0.24	-0.42	0.12	0.74	0.39	7.61	-0.30	-3.24		1.84		7.39	0.34	12.28		0.23	10.82	4.99	5.12
132	5.07	-0.06	-0.61	0.07	1.14	0.53	7.32	-0.68	-4.58		1.42		10.88	0.51	23.46	-0.28		11.95	0.99	7.36
144	3.39	-0.13	-0.18	-0.15	0.90	0.69	6.59	-0.09	-4.67		0.36		11.37	0.30	18.51	-0.73		12.32	2.41	8.62
156	2.02	-0.24	0.29	-0.32	0.71	0.70	5.69	0.19	-3.49		0.55		12.15	0.70	10.27	-0.24		11.42	2.58	8.35
168	1.14	-0.44	0.46	-0.18	0.26	0.39	-1.62	0.57	-1.71		0.57		8.63	0.45		-0.16		2.33	4.37	5.14
180	1.37	-0.29	0.04	-0.11	0.20	0.39	-1.30	0.84	-3.16		2.32		5.69	0.54	-4.91		0.19	1.10	3.39	2.34
LHS: Cumulative dividend growth																				
12	-0.10	0.09	0.34	-0.08	0.07	-0.09	-0.08	0.11	0.74	-0.57		-0.00	1.92		-3.00					3.46
24	-1.19	0.16	0.08	-0.06	-0.22	-0.31	-1.01	0.77	0.54	-2.77		-0.02	2.41		-1.58	0.38	0.26	0.74	1.73	5.30
36	-1.22	-0.01	-0.22	-0.11	-0.21	-0.32	-1.33	0.94	-0.02	-2.80		-0.01	2.20		-3.99	0.24	0.16	0.40	2.00	4.67
48	-0.97	-0.15	-0.25	-0.18	-0.14	-0.17	-1.12	0.85	-0.13	-1.70		-0.01	2.04		-6.79	0.12	0.15	0.27	1.47	3.10
60	-1.24	0.01	-0.40	-0.39	-0.00		-0.95	0.99	-0.49	-1.96		-0.00	2.42		-4.10		0.08	0.64	0.30	3.18
72	-1.49	0.12	-0.38	-0.31	-0.16	-0.13	-0.54	1.09	-0.23	-2.49		-0.00	1.74		-1.01	0.11	0.14	1.67	-0.79	3.00
84	-0.93	0.07	-0.36	-0.24	-0.11	-0.11		0.83	0.39	-1.62		-0.00	0.12	-0.08	0.85	0.05		-0.16	-1.65	1.72
96	-0.17	0.01	-0.38	-0.05	-0.14	-0.23	1.02	0.53	0.94	-0.90	-0.19	-0.15		0.00	0.74		0.11	0.64	-2.47	1.33
108	0.51	-0.03	-0.20	0.20	-0.22	-0.42	1.99	0.24	2.41	-1.52		-0.00	0.52	0.11	-0.07	-0.02			-2.96	1.77
120	1.10	-0.07	-0.10	0.10		-0.23	0.19	0.09	2.03	-2.40		-0.32	0.72		-4.33	0.05	0.20	0.71	-2.96	2.00
132	2.38	-0.13	-0.29	0.00	0.32	0.06	0.13	-0.23	0.70	-3.07		-0.46	1.12	0.06	-4.02		0.18	-0.12	-3.81	2.97
144	3.02	-0.10	-0.38	-0.13	0.58	0.26	-1.76	-0.41	0.32	-2.99		-10.54	0.83		-0.75		0.14	-0.58	-4.54	2.73
156	2.40	-0.11	-0.21	-0.12	0.39	0.23	-1.88	-0.24	1.16	-4.18	-0.27	-0.48		0.06	1.14	-0.00		-0.53	-3.90	2.68
168	1.92	-0.08	-0.17	-0.21	0.36	0.28	-1.98	-0.11	0.82	-4.13		-0.07		0.13	0.21	-0.03		1.68	-4.01	2.25
180	1.76	0.01	-0.30	-0.14	0.26	0.24	-0.99		0.87	-2.55		-0.57	0.55	0.07	-2.74	-0.20		4.33	-4.13	1.31
LHS: Dividend yield																				
12	-0.15	-0.09		0.39	0.55	0.46	2.01		-0.61		0.04		1.00		-4.42	-0.01	-0.23	2.40	-1.25	2.14
24	-1.30	0.13	-0.26	0.51	0.10	0.11	2.22	0.43	-1.56	-0.17			0.41	-0.08	-5.52	0.09		2.69	-2.52	2.37
36	-2.41	-0.03	-0.12	0.59	-0.28	-0.19	0.56	0.89	-2.18	-0.43			0.51	-0.02	-8.53	0.08		-1.67	-2.67	1.22
48	-3.18	-0.09	-0.02	0.60	-0.54	-0.32	2.17	1.12	-2.94		0.19		0.14	0.05	-12.66	-0.12		-3.04	-3.60	0.46
60	-3.32	0.04	-0.06	0.32	-0.31	-0.17	2.22	1.11	-2.61	-0.54			0.80		-12.61	-0.15	-0.06	-3.25	-5.23	0.67
72	-3.44	0.19	-0.23	0.37	-0.41	-0.32	0.04	1.17	-3.27	-1.39	-0.34	-0.29			-7.34	0.31		-5.03	-7.27	0.76
84	-3.00	0.23	-0.19	0.21	-0.15	-0.15	-0.88	0.89	-2.25	-1.13	-1.65	-0.44		-0.11	3.48	0.25		-6.96	-8.57	1.95
96	-2.72	0.16	0.33	0.11	0.03	-0.07	0.87	0.49		-0.44	-3.18	-0.03		-0.21	3.41	0.32		-5.38	-8.75	-0.18
108	-2.61	0.25	0.37	-0.21	0.40	0.16	-1.02	0.37	1.87		-4.72		-0.44	-0.30	6.38		-0.04	-4.93	-8.90	-0.68
120	-2.74		0.25	0.00		0.08	-2.87	0.38	2.07		-2.57		-1.67	-0.09	-1.92	0.13		-6.73	-7.59	
132	-2.74	-0.11	0.17	-0.12		0.23	-4.02	0.57	2.31		-3.78		-3.43	-0.27	-9.30			-8.84	-5.04	-1.85
144	-1.13	-0.21	-0.17	0.17		0.07	-2.78	0.08	2.64		-2.81		-5.38	-0.14	-13.79	0.36		-8.15	-4.98	-4.75
156	-0.61	-0.03	-0.40	0.29	0.02		-4.46	0.06	2.87		-3.92		-5.64	-0.36	-6.38	0.14		-6.44	-4.86	-4.12
168	-0.07	0.19	-0.61	0.16	0.33	0.16		0.02	2.16		-5.04		-4.21	-0.41	1.20			-0.35	-5.02	-3.65
180	0.24	0.12	-0.30	0.00	0.60	0.30	0.26	-0.46	2.68		-5.62		-1.41	-0.39	5.30		-0.18	4.95	-5.69	-1.49

Table IX: Volatility decompositions build upon the cross-validation-based LASSO estimator and resulting local projections (both annual and monthly data).

This table reports the annual and monthly dividend yield volatility decompositions for different horizons k based on the local projections (equation (7)) where in the LASSO estimator (equation (15)) the tuning parameter λ is determined by the 10-fold cross-validation, with the value of the parameter λ at which the minimum mean square error is achieved.

A: Annual					B: Monthly				
k (years)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	k (months)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
1	0.14	0.16	0.92	1.14	12	0.23	0.24	0.94	1.08
2	0.42	0.21	0.83	0.50	24	0.39	0.28	0.86	0.72
3	0.54	0.24	0.81	0.45	36	0.52	0.31	0.84	0.60
4	0.56	0.27	0.77	0.49	48	0.63	0.32	0.81	0.51
5	0.61	0.27	0.72	0.44	60	0.73	0.31	0.77	0.43
6	0.71	0.30	0.69	0.43	72	0.81	0.32	0.79	0.40
7	0.73	0.29	0.67	0.40	84	0.84	0.32	0.80	0.38
8	0.85	0.31	0.65	0.36	96	0.95	0.34	0.77	0.36
9	0.90	0.30	0.75	0.34	108	0.93	0.40	0.77	0.42
10	0.96	0.28	0.70	0.29	120	0.91	0.42	0.70	0.46
11	0.95	0.40	0.63	0.42	132	0.92	0.49	0.67	0.53
12	0.96	0.51	0.60	0.53	144	0.93	0.54	0.64	0.57
13	0.96	0.37	0.65	0.38	156	0.94	0.52	0.59	0.56
14	0.93	0.37	0.53	0.40	168	0.92	0.49	0.54	0.54
15	0.91	0.38	0.50	0.42	180	0.95	0.40	0.55	0.43

III.IV Post-LASSO step

Instead of the direct use of the LASSO estimator (equation (15)), we can also consider a 'post-LASSO' step where the state variables for information-rich local projections are first selected by the LASSO but the final estimation of the selected model is carried out by OLS. Belloni and Chernozhukov (2013), among others, have shown that the post-LASSO step has the (near) oracle property, that is the estimator is consistent both in parameter estimation and variable selection. Put differently, the inference is as efficient as if 'an oracle' had revealed the true model and estimation had been carried out using only the relevant state variables. If the LASSO selects the 'true' model in the first-stage, the OLS estimator in the second stage becomes the oracle estimator.

Table X reports the volatility decomposition of the post-LASSO step local projections. The selected state variables for each LHS variable are the same as presented in Table 5 and Table V above (i.e. the components with non-empty cells), but the final estimation is carried out by OLS. The essential message is still the same as in the previous analyses: The main contributor for asset volatility is the discount rate variation but the cash flows also matter. The expected dividend growth contributions at different horizons k and both data frequencies (annual and monthly) are generally diverging only marginally from the ones in Table 4 obtained with the LASSO. In other words, replacing the shrinkage-based estimation by the OLS step does not change the main findings and, hence, our results are robust also to this change of methodology.

Table X: Volatility decompositions based on the post-LASSO step.

This table reports the annual (Panel A) and monthly (Panel B) volatility decomposition of the dividend yield based on the local projections where the predictors for each LHS and horizon k are selected by the LASSO but the final estimation of the selected model is carried out with OLS.

A: Annual					B: Monthly				
k (years)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	k (months)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
1	0.13	0.15	0.90	1.17	12	0.23	0.25	0.94	1.10
2	0.38	0.23	0.83	0.60	24	0.40	0.29	0.86	0.71
3	0.44	0.28	0.84	0.62	36	0.52	0.31	0.83	0.60
4	0.64	0.31	0.82	0.48	48	0.64	0.32	0.81	0.50
5	0.75	0.31	0.77	0.41	60	0.74	0.32	0.77	0.43
6	0.81	0.33	0.75	0.41	72	0.82	0.32	0.79	0.39
7	0.80	0.31	0.72	0.39	84	0.84	0.32	0.79	0.38
8	0.90	0.32	0.71	0.36	96	0.95	0.35	0.76	0.36
9	0.96	0.34	0.72	0.36	108	0.93	0.40	0.77	0.43
10	0.97	0.35	0.69	0.35	120	0.91	0.43	0.70	0.47
11	0.96	0.40	0.68	0.42	132	0.93	0.49	0.68	0.53
12	0.97	0.51	0.64	0.52	144	0.94	0.54	0.65	0.58
13	0.99	0.41	0.62	0.42	156	0.94	0.53	0.61	0.57
14	0.95	0.44	0.57	0.46	168	0.92	0.50	0.54	0.55
15	0.94	0.44	0.58	0.47	180	0.95	0.41	0.55	0.43

III.V Elastic net

In statistical (machine) learning, the elastic net is a regularized regression method that linearly combines the LASSO and ridge regressions (see Zou and Hastie, 2005, and also, e.g., Hastie, Tibshirani and Friedman, 2009, pp. 662–663). Ridge regression shrinks parameter estimates using the $l2$ penalty, which precludes shrinkage to zero, while the LASSO also performs variable selection by employing the $l1$ penalty. In the past statistical learning studies, a reported potential drawback of the LASSO is that it may tend to arbitrarily select a single predictor from a group of correlated predictors, possibly making it less informative for datasets with several rather strongly correlated regressors. This is to some extent the case in our study. The elastic net is proposed to avoid this problem by including both $l1$ (LASSO) and $l2$ (ridge) penalty terms. See, e.g., Bai and Ng (2008) for a survey of variable selection and parameter estimation with a large number of potential predictors and applications of the LASSO and the elastic net to macroeconomic forecasting.

Formally, like the LASSO estimator introduced in equation (15) of the paper, the elastic net estimator is based on a penalized sum of squared errors where the penalty (shrinkage) term is different than in the LASSO:

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

where $\boldsymbol{\varphi}^{(a,k)} = (\alpha^{(a,k)}, \boldsymbol{\beta}^{(a,k)})$, $a \in \{r, d, dp\}$, and the parameter ν measures the compromise between the LASSO and ridge penalties (all the other notations as in equation (15)). In finance, Rapach, Strauss and Zhou (2013) were the first ones to apply (a modification of) the elastic net estimator in their predictive regressions for international stock returns. We experimented with a range of ν values $(0, 0.25, 0.5, 0.75)$ ($\nu = 1$ corresponds the LASSO), with essentially the same main conclusions on the relative importance of expected discount rates and cash flows channels. Following the common selection in the applications of the elastic net in various fields, in Table XIII we report the results of $\nu = 0.5$, which gives equal weight to ridge regression and LASSO. In this case, we benefit both from the feature selection capability of the LASSO and the ability of ridge regression to handle, for example, the potential multicollinearity issues in data-rich local projections.

Tables XI–XII present the elastic net-based estimated coefficients in the corresponding local projections. In Table XIII, the volatility shares and the ratios between the cash flow and discount rate components are again very close to the ones reported for the LASSO-based local projections and all the main conclusions are intact. So, when modifying the learning part of the local projections, which can be interpreted as agents' (investors') learning on the useful state variables, the main findings are the same as obtained with the LASSO in the main text.

Table XI: Elastic net estimation and model selection results for the annual data.

In this table, empty cells indicate the exclusion of those state variables in the elastic net-based local projections ($\nu = 0.5$). The results are reported for the three left hand side (LHS) variables at the time and for all horizons k.

k (years)	const	r	Δd	dp	EP	DE	$SVAR$	BM	$NTIS$	TBL	LTY	$RFREE$	TMS	LTR	DFY	DFR	$CORPR$	$INFL$	CAY	$EQIS$	IK
LHS: Cumulative return																					
1	0.55			0.13																	
2	0.98			0.22																	
3	1.25			0.28									0.79						1.70		
4	2.36	-0.14	-0.30	0.35	0.20		1.24			-0.80		-0.75	1.50						5.45		-0.52
5	2.70	-0.11	-0.33	0.58			1.90						2.41			-0.12			4.39	-0.24	-2.05
6	2.78	-0.10	-0.19	0.54			1.52						3.49			-0.82		1.72	5.45		-10.69
7	2.79	-0.07		0.53		-0.04	0.69						5.05			-0.81		2.59	5.36		-14.36
8	3.55	-0.18	-0.52	0.56			1.31		-0.36		2.61		3.07			-0.30		2.60	5.78		-23.49
9	4.14	-0.32	-0.84	0.36	0.07		2.96	0.21	-3.90		7.78			0.14	-13.80	-0.50		1.05	4.55	-0.14	-47.36
10	5.33	-0.19	-0.51	0.24	0.63	0.18	3.46	-0.46	-4.42		3.06		4.88	0.51	-15.84			4.87	3.26	0.36	-40.44
11	5.23	-0.13	-0.13	0.06	0.93	0.42	3.21	-0.86	-2.89		3.98		7.30	0.40	-6.97	-0.77		5.20		-0.10	-30.24
12	3.91	-0.25	-0.05	0.10	0.39	0.25	4.38		-2.60		7.59		4.87	0.29	-21.91	-1.74		1.95		-0.47	-39.80
13	4.20	-0.33	-0.15	0.56			4.30		-4.23		4.47		5.63		-15.25	-1.00		1.64	1.07		-27.27
14	3.03	-0.49	0.23	0.28		0.17	1.25		-2.52		4.64		7.79	0.28	-21.96	-1.55		3.59	0.91	-0.13	-19.46
15	3.05	-0.34	-0.14	0.22		0.08		0.40	-5.02		3.99		5.10	0.15	-12.60	-0.41		2.05	0.84		-23.88
LHS: Cumulative dividend growth																					
1	0.00	0.15		0.00		-0.07			0.57				0.10		-0.91				-0.25		
2	-0.11	0.01	-0.12	-0.05		-0.14		0.29	2.31				0.52			0.21	0.15	-1.19	-0.30	-0.36	-3.16
3	-0.09	-0.09	-0.22	-0.06		-0.09		0.32	1.85	-0.53		-0.50	1.07		-2.38	0.37	0.21			-0.19	
4	0.26	-0.19	-0.34	-0.03		-0.02		0.13	1.05							0.29	0.05				
5	-0.26	-0.03	-0.39	-0.17				0.31	0.85				0.96			0.24			-0.70	0.16	-1.61
6	-1.21	0.09	-0.40	-0.46		0.08		0.85		-0.41	-0.53	-0.41			-0.25	0.78	0.16		-1.08	0.12	-6.46
7	-0.71	0.04	-0.37	-0.34			0.06	0.60	0.72		-1.34				2.66	0.51	0.02		-1.25	0.32	-4.99
8	0.11		-0.30	-0.12		-0.01		0.11	1.11									0.35	-2.58	0.40	
9	0.50	-0.02	-0.58	-0.07		-0.05		0.15	0.92				-0.59		1.58	0.25	0.09		-3.11	0.31	
10	0.79	-0.06	-0.24	-0.01					1.79										-2.72		
11	1.03	-0.09	-0.21	0.01					1.69	-0.29	-0.65	-0.28							-2.31		
12	3.44	-0.16	-0.90	0.08	0.43	0.24	1.29	-0.55		-0.61	-0.75	-0.58				0.16	0.08	-1.07	-4.43	0.39	5.88
13	1.59	-0.05	-0.44	0.05		0.07	-0.39		1.11	-0.68	-0.63	-0.61							-3.57		
14	0.94	-0.10	-0.08	-0.14		0.11	-1.40		1.74	-0.64	-1.02	-0.64						0.43	-3.69	-0.18	-2.88
15	0.98		-0.15	-0.11		0.12	-1.10		0.92	-0.73	-1.28	-0.66		0.04				1.19	-3.97		-2.27
LHS: Dividend yield																					
1	-0.25			0.90																	
2	-0.58			0.79															-1.83		
3	-1.53			0.54				0.32											-3.46		
4	-2.04	-0.00		0.40		0.06		0.39	-0.74										-4.47		3.38
5	-2.39		0.18	0.32				0.27											-5.35	0.27	4.73
6	-2.07			0.37				0.05											-6.64	0.39	7.55
7	-2.40			0.26				0.09											-6.90	0.46	7.33
8	-2.45			0.23									-0.77						-7.04	0.44	9.99
9	-2.90			0.17							-1.05		-0.30						-7.46	0.73	18.65
10	-2.94			0.18					1.90		-1.58		-0.04						-7.04	0.43	24.45
11	-3.65	-0.06		-0.02		0.23	-1.54	0.52	2.47		-4.85		-1.39	-0.11	3.92	0.79		-1.57	-3.89	0.28	34.62
12	-2.26	-0.10	-0.11	0.26		0.11	-0.93		2.91		-5.90		-0.52			0.71			-3.55	0.48	33.52
13	-1.98		-0.25	0.26			-2.77	0.19	3.00		-7.96		-1.95	-0.30	14.16	0.58		-1.25	-2.28	0.41	26.52
14	-1.50	0.16	-0.26	0.28	0.15		-1.28		3.67		-8.04		-1.27	-0.22	18.14	0.70			-3.24		23.87
15	-1.17		-0.14	0.17			-0.16		1.64		-5.69		-2.41	-0.35			-0.12		-3.24		

Table XII: Elastic net estimation and model selection results for the monthly data.

In this table, empty cells indicate the exclusion of those state variables in the elastic net estimation results ($\nu = 0.5$). The elastic net estimation results are reported for the three left hand side (LHS) variables at the time and for all horizons k.

k (months)	const	r	Δd	dp	EP	DE	$SVAR$	BM	$NTIS$	TBL	LTY	$RFREE$	TMS	LTR	DFY	DFR	$CORPR$	$INFL$	CAY	$RREL$
									LHS: Cumulative return											
12	0.34	-0.16	0.06	0.02	0.09	-0.02	-1.24	0.20	0.21	-1.26		-12.63			6.98	0.14	0.35	-4.01	3.11	1.47
24	1.05	-0.24	0.26	-0.17	0.45	0.31			1.29	-1.20		-13.05	2.57		7.81	0.06	0.14	-1.92	4.64	4.49
36	2.03	-0.11	-0.01	-0.49	1.01	0.76	-3.06	-0.23	1.12	-1.49		-16.37	2.93		11.47	0.26	0.27	3.11	5.41	4.67
48	1.94	-0.05	-0.37	-0.40	0.86	0.70	-1.11		0.12	-1.33		-13.28	2.43		13.83	0.36	0.02	4.04	6.30	4.40
60	1.50	-0.08	-0.33	-0.47	0.81	0.75	-0.89	0.26	-0.31	-1.10		-11.07	2.68		16.97		0.19	4.78	7.72	4.51
72	1.41	-0.14	-0.04	-0.42	0.75	0.72	2.16	0.23	0.95	-0.61		-3.63	3.77	0.16	13.17	-0.11		9.10	8.68	3.67
84	1.27	-0.20	-0.25	-0.04	0.29	0.31	2.99	0.39			1.69		3.36	0.42	5.40			7.89	8.19	2.12
96	1.11	-0.31	-0.71	0.28	-0.16	0.04	2.77	0.71	-1.43		4.64		2.15	0.34	0.53		0.02	5.41	8.86	2.02
108	2.52	-0.31	-0.60	0.44	0.05		5.27	0.22	-2.38		3.86		4.55		3.77		0.61	6.63	6.89	2.86
120	3.92	-0.25	-0.42	0.11	0.76	0.41	7.61	-0.33	-3.29		1.84		7.42	0.35	12.56		0.23	10.88	4.93	5.18
132	5.09	-0.06	-0.61	0.07	1.14	0.54	7.34	-0.68	-4.60		1.44		10.87	0.51	23.47	-0.29		11.98	0.99	7.38
144	3.15	-0.12	-0.16	-0.09	0.78	0.60	6.78		-4.47		0.47		11.28	0.32	17.26	-0.66		12.08	2.61	8.51
156	1.97	-0.24	0.30	-0.33	0.72	0.71	5.64	0.21	-3.51		0.49		12.19	0.70	10.38	-0.26		11.45	2.64	8.43
168	1.15	-0.44	0.45	-0.19	0.26	0.39	-1.64	0.56	-1.72		0.59		8.64	0.45		-0.16		2.38	4.34	5.14
180	1.39	-0.29	0.04	-0.12	0.21	0.40	-1.37	0.83	-3.17		2.34		5.71	0.54	-4.91		0.20	1.17	3.35	2.33
									LHS: Cumulative dividend growth											
12	-0.05	0.07	0.32	-0.04	0.05	-0.12		0.08	0.77	-0.26		-2.91	1.72		-2.70					3.29
24	-1.19	0.16	0.08	-0.06	-0.22	-0.31	-1.01	0.77	0.54	-1.42		-16.29	2.41		-1.58	0.38	0.26	0.75	1.73	5.31
36	-1.17	-0.02	-0.22	-0.12	-0.19	-0.30	-1.42	0.91		-1.42		-16.39	2.21		-3.83	0.24	0.16	0.41	1.93	4.64
48	-0.62	-0.15	-0.24	-0.23		-0.06	-1.56	0.69		-0.84		-8.60	2.11		-5.81	0.11	0.13	0.15	1.01	2.88
60	-1.03		-0.38	-0.34		-0.01	-0.61	0.87	-0.18	-0.89		-9.90	2.41		-3.38		0.06	0.68		2.87
72	-1.51	0.12	-0.38	-0.30	-0.17	-0.14	-0.51	1.10	-0.25	-1.27		-14.69	1.76		-1.10	0.11	0.15	1.69	-0.79	3.03
84	-0.90	0.07	-0.36	-0.24	-0.10	-0.11		0.82	0.41	-0.83		-9.53	0.13	-0.08	0.99	0.05		-0.14	-1.69	1.72
96	-0.14	0.01	-0.38	-0.05	-0.13	-0.22	0.99	0.52	0.95	-0.47	-0.20	-5.43			0.90		0.11	0.70	-2.50	1.34
108	0.62	-0.03	-0.20	0.16	-0.15	-0.37	1.73	0.20	2.39	-0.77		-9.36	0.49	0.10					-3.06	1.66
120	1.13	-0.08	-0.09	0.10		-0.21		0.06	2.03	-1.19		-13.65	0.54		-3.94		0.15	0.63	-2.91	1.86
132	2.07	-0.13	-0.27	0.02	0.23			-0.12	0.78	-1.55		-18.69	0.92	0.02	-4.79		0.19		-3.39	2.92
144	3.00	-0.10	-0.38	-0.13	0.57	0.25	-1.75	-0.40	0.33	-1.93		-22.96	0.86		-0.95		0.15	-0.62	-4.52	2.75
156	2.37	-0.11	-0.21	-0.11	0.37	0.22	-1.85	-0.23	1.19	-2.12	-0.32	-24.78		0.06	1.00			-0.58	-3.83	2.69
168	1.91	-0.08	-0.16	-0.20	0.35	0.27	-1.97	-0.10	0.84	-2.11		-24.21		0.14	0.15	-0.03		1.70	-3.99	2.25
180	1.76	0.02	-0.30	-0.16	0.27	0.25	-1.05		0.83	-1.36		-15.22	0.63	0.08	-2.48	-0.22		4.45	-4.17	1.40
									LHS: Dividend yield											
12	-0.15	-0.08		0.36	0.57	0.48	2.00		-0.67		0.05		1.07		-4.71	-0.01	-0.23	2.43	-1.28	2.18
24	-1.23	0.09	-0.24	0.63			2.29	0.37	-1.09				0.11	-0.07	-3.93			2.39	-2.51	2.08
36	-1.65	-0.04	-0.08	0.51	-0.02			0.51	-1.23	-0.21		-1.42			-3.64				-3.16	0.36
48	-2.38	-0.07		0.40	-0.14			0.74	-2.34						-8.07			-1.35	-4.16	
60	-2.38			0.29	-0.04		0.72	0.63	-1.34	-0.21		-2.26			-6.25			-0.97	-5.59	
72	-3.40	0.19	-0.23	0.36	-0.39	-0.31		1.15	-3.25	-0.77	-0.35	-7.74			-7.20	0.32		-5.00	-7.30	0.75
84	-3.03	0.23	-0.19	0.22	-0.17	-0.17	-0.86	0.91	-2.27	-0.65	-1.66	-6.44		-0.12	3.58	0.25		-7.03	-8.54	2.02
96	-2.50	0.09	0.28	0.17		-0.05		0.36			-2.21					0.13		-4.75	-8.66	-0.45
108	-2.80	0.18	0.37	-0.04	0.16			0.41	1.83		-3.75		-0.61	-0.23	2.62			-4.77	-8.63	-0.57
120	-2.71	-0.01	0.25	0.01		0.09	-2.88	0.36	2.08		-2.50		-1.77	-0.09	-1.99	0.16		-6.73	-7.52	
132	-3.10	-0.13	0.27	-0.21		0.26	-4.96	0.69	2.59		-3.98		-3.88	-0.34	-11.08	0.14		-9.70	-4.44	-2.45
144	-1.29	-0.23	-0.17	0.11		0.11	-3.39	0.18	2.59		-2.93		-5.87	-0.23	-15.44	0.52		-9.58	-4.47	-5.17
156	-0.74	-0.02	-0.40	0.23	0.05		-5.10	0.16	3.05		-4.21		-6.00	-0.46	-7.64	0.30		-8.51	-4.44	-4.76
168	-0.19	0.15	-0.58	0.27	0.19	0.05		0.02	2.38		-4.81		-4.26	-0.35					-4.87	-3.52
180	0.24	0.12	-0.30	0.00	0.60	0.30	0.28	-0.46	2.68		-5.61		-1.42	-0.39	5.29		-0.18	4.97	-5.69	-1.50

Table XIII: Volatility decompositions based on the elastic net.

This table reports the annual (Panel A) and monthly (Panel B) volatility decomposition based on the elastic net estimator where $\nu = 0.5$.

A: Annual					B: Monthly				
k (years)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	k (months)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
1	0.13	0.12	0.90	0.90	12	0.23	0.24	0.94	1.04
2	0.22	0.19	0.81	0.87	24	0.39	0.28	0.86	0.72
3	0.29	0.25	0.81	0.85	36	0.52	0.30	0.82	0.59
4	0.54	0.24	0.77	0.44	48	0.63	0.31	0.80	0.49
5	0.61	0.24	0.72	0.39	60	0.73	0.30	0.75	0.41
6	0.69	0.30	0.68	0.43	72	0.81	0.32	0.79	0.40
7	0.72	0.29	0.66	0.40	84	0.83	0.32	0.80	0.38
8	0.85	0.25	0.61	0.30	96	0.95	0.34	0.73	0.36
9	0.94	0.32	0.64	0.34	108	0.92	0.39	0.75	0.43
10	0.97	0.26	0.61	0.27	120	0.91	0.41	0.69	0.45
11	0.95	0.30	0.66	0.32	132	0.92	0.48	0.68	0.51
12	0.96	0.50	0.60	0.52	144	0.93	0.54	0.64	0.58
13	0.96	0.40	0.60	0.42	156	0.94	0.52	0.61	0.56
14	0.94	0.45	0.54	0.48	168	0.92	0.49	0.54	0.54
15	0.92	0.42	0.53	0.46	180	0.95	0.40	0.55	0.43

IV Alternative datasets: Dividends and sample periods

IV.I Subsample analysis

In Section 3.5 of the paper, we provide a volatility decomposition based on local projections with time-varying parameters. The empirical evidence favours our main conclusions: Allowing for time-varying parameters in (flexible) local projections, the relative importance of expected cash flows is substantially more important than previously thought in the corresponding volatility decompositions.

Supplementing the time-varying parameter analysis, we provide subsample analyses in this section: we estimate the volatility decomposition on two (partly overlapping) subsamples: *(i)* the period 1952–2007, which excludes the Great Recession of 2008–2009 and it's aftermath; and *(ii)* the period 1985–2017, during which the dividend yield trended downwards (caused by a prolonged upward trend in stock prices) and reached a lower average level than before. We consider again one and three state variables, LASSO and model averaging. Due to the reduced number of observations in these subsamples, we consider exclusively monthly data (facilitated by the local projection approach). We also investigate the earlier sample of 1928–1951. For this sample, we only consider the LPs with one or three state variables. We omit the LASSO and model averaging approaches, since not all variables (listed in Table 1) are available during this period. Moreover, the shorter data sample allows to estimate the volatility decomposition only up to horizons of 10 years (120 months).

Results are reported in Tables XIV, XV and XVI. Both in the 1928–1951 and 1952–2007 samples, the contribution of long-run dividend expectations is substantially higher than in the later 1985–2017 sample. For example, the ratio (equation (13)) of expected dividend volatility to expected return volatility reaches 0.90 in the 1952–2007 subsample (using LASSO model selection), while the ratio even exceeds one in the early 1928–1951 sample (using three state variables). The declining dividend contribution over time is consistent with patterns documented by Chen (2009) and Golez and Koudijs (2018). Interestingly, in the later 1985–2017 subsample, the estimated contribution of dividends is still higher than in the benchmark full-sample results in Table 2. This discrepancy between the full-sample and subsample results is likely caused by the trending behaviour of the dividend yield, which has been documented to generate spurious predictability of long-horizon returns (see, e.g., Boudoukh, Richardson and Whitelaw, 2008).

Table XIV: Volatility decomposition: Subsample 1952:3–2007:12 (monthly data).

This table reproduces the monthly volatility decompositions as reported in Tables 2–4 of the paper, with the difference that the local projections are now estimated on the monthly subsample period 1952:3–2007:12.

	A: One state variable				B: Three state variables			
k (months)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
12	0.11	0.00	0.91	0.03	0.13	0.12	0.92	0.92
24	0.19	0.01	0.82	0.05	0.21	0.05	0.83	0.24
36	0.26	0.02	0.75	0.07	0.24	0.13	0.80	0.54
48	0.31	0.02	0.68	0.06	0.35	0.21	0.73	0.62
60	0.39	0.01	0.59	0.01	0.43	0.24	0.66	0.55
72	0.48	0.00	0.52	0.00	0.51	0.24	0.60	0.47
84	0.56	0.01	0.47	0.02	0.67	0.28	0.57	0.41
96	0.63	0.03	0.43	0.05	0.90	0.32	0.62	0.35
108	0.71	0.06	0.34	0.08	0.93	0.35	0.58	0.38
120	0.79	0.07	0.24	0.09	0.97	0.39	0.56	0.40
132	0.89	0.11	0.13	0.12	1.04	0.53	0.60	0.51
144	0.96	0.17	0.00	0.18	1.03	0.64	0.64	0.62
156	1.03	0.22	0.11	0.21	0.96	0.66	0.61	0.69
168	1.13	0.26	0.23	0.23	0.98	0.70	0.64	0.72
180	1.26	0.32	0.41	0.25	1.09	0.66	0.61	0.61
	C: LASSO				D: Model averaging			
k (months)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
12	0.22	0.18	0.95	0.82	0.13	0.12	0.93	0.94
24	0.38	0.22	0.89	0.58	0.21	0.05	0.84	0.25
36	0.49	0.25	0.86	0.51	0.24	0.13	0.80	0.54
48	0.56	0.29	0.83	0.51	0.35	0.22	0.73	0.61
60	0.64	0.32	0.77	0.49	0.44	0.23	0.67	0.53
72	0.70	0.30	0.77	0.43	0.52	0.24	0.60	0.46
84	0.80	0.33	0.77	0.41	0.69	0.27	0.58	0.39
96	0.96	0.33	0.73	0.34	0.93	0.32	0.64	0.34
108	0.87	0.43	0.72	0.49	0.91	0.35	0.60	0.39
120	0.81	0.49	0.67	0.61	0.92	0.39	0.56	0.42
132	0.73	0.52	0.64	0.71	0.96	0.52	0.58	0.54
144	0.69	0.58	0.65	0.83	0.95	0.63	0.60	0.66
156	0.68	0.54	0.66	0.79	0.92	0.64	0.57	0.70
168	0.64	0.53	0.65	0.83	0.96	0.68	0.61	0.71
180	0.63	0.57	0.59	0.90	1.06	0.65	0.59	0.61

Table XV: Volatility decomposition: Subsample 1985:1–2017:12 (monthly data).

This table reproduces the monthly volatility decompositions as reported in Tables 2–4 of the paper, with the difference that the local projections are now estimated on the monthly subsample period 1985:1–2017:12.

	A: One state variable				B: Three state variables			
k (months)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
12	0.17	0.05	0.85	0.30	0.17	0.43	0.87	2.60
24	0.30	0.04	0.71	0.15	0.40	0.23	0.71	0.57
36	0.38	0.06	0.59	0.16	0.82	0.49	0.58	0.59
48	0.48	0.11	0.45	0.22	1.08	0.80	0.56	0.74
60	0.56	0.17	0.28	0.31	0.99	0.73	0.57	0.73
72	0.62	0.22	0.16	0.35	0.87	0.36	0.41	0.42
84	0.72	0.24	0.02	0.34	0.97	0.29	0.48	0.30
96	0.84	0.22	0.10	0.26	1.03	0.30	0.40	0.29
108	0.90	0.18	0.20	0.20	1.47	0.26	0.26	0.18
120	0.95	0.16	0.26	0.17	1.61	0.29	0.16	0.18
132	0.96	0.17	0.31	0.18	1.99	0.83	0.10	0.42
144	0.93	0.19	0.33	0.21	1.93	0.77	0.19	0.40
156	0.87	0.21	0.31	0.24	1.29	0.66	0.16	0.51
168	0.82	0.21	0.26	0.25	1.36	0.28	0.51	0.21
180	0.80	0.20	0.24	0.25	1.49	0.16	0.15	0.11

	C: LASSO				D: Model averaging			
k (months)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
12	0.48	0.61	0.84	1.27	0.18	0.45	0.87	2.55
24	0.77	0.80	0.69	1.04	0.42	0.27	0.71	0.63
36	0.94	0.83	0.63	0.88	0.82	0.48	0.58	0.58
48	1.17	0.69	0.66	0.59	1.09	0.79	0.56	0.72
60	1.12	0.74	0.49	0.65	1.02	0.71	0.55	0.70
72	1.41	0.39	0.34	0.28	0.92	0.36	0.38	0.39
84	1.53	0.39	0.31	0.25	0.97	0.29	0.42	0.30
96	1.13	0.44	0.83	0.39	1.04	0.29	0.38	0.28
108	1.27	0.78	0.35	0.61	1.46	0.28	0.26	0.19
120	1.07	0.66	0.65	0.62	1.59	0.36	0.15	0.23
132	1.52	0.68	0.57	0.45	1.96	0.81	0.14	0.41
144	1.88	0.52	0.37	0.28	1.90	0.73	0.18	0.38
156	1.62	0.46	0.51	0.28	1.26	0.63	0.11	0.50
168	0.96	0.25	0.54	0.26	1.28	0.25	0.46	0.20
180	1.38	0.18	0.19	0.13	1.58	0.17	0.14	0.11

Table XVI: Volatility decomposition: 1928:6–1951:12 (monthly data).

This table reproduces the monthly volatility decompositions as reported in Tables 2–3 of the paper, with the difference that the local projections are now estimated on the monthly subsample period 1928:6–1951:12. Due to the reduces sample size, the maximum horizon k considered is 120 months.

	A: One state variable				B: Three state variables			
k (months)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
12	0.53	0.78	0.50	1.49	0.58	0.51	0.51	0.89
24	1.05	0.83	0.02	0.79	0.96	0.76	0.21	0.79
36	1.49	0.74	0.32	0.50	0.75	0.98	0.19	1.30
48	1.59	0.50	0.36	0.32	0.94	0.90	0.34	0.95
60	1.48	0.26	0.09	0.18	1.03	0.64	0.36	0.62
72	1.45	0.14	0.06	0.10	0.90	0.48	0.43	0.54
84	1.37	0.20	0.13	0.14	1.09	0.87	0.37	0.80
96	1.47	0.20	0.14	0.13	1.18	0.57	0.37	0.49
108	1.42	0.09	0.27	0.06	1.06	0.52	0.47	0.49
120	1.30	0.30	0.42	0.23	1.14	1.34	0.29	1.18

IV.II S&P 500 index returns

In accordance with, e.g., van Binsbergen and Koijen (2010) and their latent variable approach, we also repeat our empirical analysis for the S&P 500 index returns and dividends, as opposed to CRSP value weighted returns and dividends. In the return predictability literature, both the use of CRSP and S&P 500 data are widespread (see, e.g., Ang and Bekaert, 2007; Welch and Goyal, 2008). We use the S&P 500 returns and S&P 500 12-month cumulative dividends reported in the (updated) dataset by Welch and Goyal (2008). Other than the dividend series analyzed in the main paper, the cumulative dividends in the Welch-Goyal dataset are simply aggregated over 12 months, without a re-investment scheme for dividends received during the 12-month windows. We will more closely investigate the effects of dividend re-investment in the next subsection.

The results, reported in Tables XVII and XVIII are very similar to the CRSP-based results in the paper. With a single state variable, the impact of long-run dividend expectations is almost neglible, while this impact (in particular relative to the impact of long-run return predictability) increases substantially when larger information sets (three state variables, LASSO, model averaging) are used for predicting dividend growth rates.

Table XVII: Volatility decomposition: S&P 500 (annual data).

This table reproduces the annual volatility decompositions as reported in Tables 2–4 of the paper, with the difference that the local projections are estimated with S&P 500 returns and dividends instead of CRSP value-weighted average returns and dividends.

	A: One state variable				B: Three state variables			
k (years)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
1	0.10	0.02	0.89	0.16	0.12	0.09	0.97	0.77
2	0.18	0.03	0.80	0.16	0.24	0.04	0.81	0.15
3	0.24	0.03	0.72	0.11	0.23	0.13	0.78	0.57
4	0.32	0.01	0.64	0.05	0.35	0.27	0.74	0.76
5	0.37	0.00	0.55	0.01	0.38	0.28	0.66	0.75
6	0.44	0.01	0.49	0.02	0.48	0.20	0.64	0.43
7	0.51	0.00	0.42	0.01	0.60	0.16	0.66	0.27
8	0.58	0.02	0.36	0.03	0.65	0.17	0.62	0.27
9	0.64	0.03	0.30	0.05	0.70	0.17	0.56	0.24
10	0.71	0.03	0.25	0.05	0.75	0.23	0.57	0.31
11	0.73	0.03	0.20	0.03	0.79	0.32	0.55	0.41
12	0.74	0.01	0.16	0.01	0.87	0.35	0.59	0.40
13	0.75	0.01	0.12	0.01	0.96	0.35	0.60	0.37
14	0.77	0.02	0.08	0.03	1.06	0.35	0.64	0.33
15	0.79	0.04	0.04	0.05	1.05	0.35	0.53	0.33
	C: LASSO				D: Model averaging			
k (years)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
1	0.10	0.13	0.97	1.32	0.12	0.09	0.97	0.78
2	0.17	0.14	0.83	0.78	0.24	0.04	0.82	0.17
3	0.22	0.17	0.80	0.74	0.24	0.13	0.78	0.56
4	0.52	0.26	0.78	0.49	0.36	0.27	0.74	0.74
5	0.57	0.29	0.72	0.50	0.39	0.28	0.67	0.73
6	0.72	0.29	0.74	0.40	0.48	0.21	0.65	0.43
7	0.71	0.24	0.72	0.34	0.60	0.17	0.67	0.28
8	0.83	0.27	0.59	0.33	0.65	0.18	0.63	0.28
9	0.90	0.24	0.62	0.26	0.71	0.17	0.56	0.25
10	0.97	0.23	0.65	0.24	0.75	0.23	0.56	0.31
11	0.98	0.23	0.62	0.23	0.78	0.33	0.54	0.42
12	0.96	0.29	0.60	0.30	0.86	0.36	0.57	0.42
13	1.00	0.31	0.63	0.31	0.96	0.37	0.58	0.38
14	0.98	0.37	0.51	0.38	1.06	0.37	0.63	0.35
15	0.96	0.31	0.49	0.32	1.04	0.36	0.52	0.35

Table XVIII: Volatility decomposition : S&P 500 (monthly data).

This table reproduces the monthly volatility decompositions as reported in Tables 2–4 of the paper, with the difference that the local projections are estimated with S&P 500 returns and dividends instead of CRSP value-weighted average returns and dividends.

	A: One state variable				B: Three state variables			
k (months)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
12	0.10	0.02	0.89	0.19	0.12	0.16	0.96	1.38
24	0.18	0.03	0.80	0.16	0.20	0.08	0.82	0.40
36	0.25	0.03	0.71	0.11	0.27	0.11	0.79	0.42
48	0.32	0.02	0.63	0.06	0.34	0.26	0.73	0.76
60	0.38	0.01	0.54	0.02	0.40	0.23	0.66	0.57
72	0.44	0.00	0.48	0.01	0.52	0.16	0.64	0.31
84	0.51	0.01	0.41	0.03	0.59	0.16	0.63	0.27
96	0.58	0.03	0.35	0.04	0.64	0.16	0.59	0.25
108	0.65	0.03	0.29	0.05	0.70	0.20	0.57	0.28
120	0.69	0.03	0.25	0.05	0.73	0.28	0.59	0.38
132	0.71	0.03	0.21	0.04	0.77	0.31	0.58	0.41
144	0.72	0.02	0.17	0.03	0.84	0.34	0.59	0.41
156	0.74	0.01	0.13	0.01	0.91	0.35	0.59	0.38
168	0.75	0.00	0.09	0.00	0.97	0.35	0.54	0.36
180	0.80	0.00	0.05	0.00	0.98	0.32	0.38	0.32
	C: LASSO				D: Model averaging			
k (months)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
12	0.17	0.22	0.97	1.32	0.12	0.17	0.96	1.41
24	0.34	0.24	0.88	0.71	0.20	0.08	0.83	0.42
36	0.50	0.27	0.82	0.54	0.27	0.11	0.79	0.42
48	0.62	0.30	0.79	0.49	0.34	0.26	0.73	0.76
60	0.68	0.28	0.76	0.41	0.40	0.23	0.67	0.57
72	0.75	0.24	0.75	0.33	0.53	0.17	0.66	0.32
84	0.80	0.22	0.75	0.28	0.59	0.17	0.65	0.28
96	0.96	0.23	0.72	0.24	0.64	0.17	0.60	0.26
108	0.94	0.31	0.71	0.33	0.70	0.19	0.57	0.27
120	0.98	0.35	0.70	0.36	0.74	0.27	0.59	0.36
132	0.97	0.41	0.65	0.42	0.77	0.31	0.58	0.41
144	0.95	0.43	0.62	0.45	0.83	0.35	0.58	0.42
156	0.98	0.47	0.58	0.48	0.91	0.37	0.58	0.40
168	0.96	0.46	0.56	0.48	0.97	0.36	0.54	0.38
180	0.98	0.37	0.48	0.37	0.97	0.32	0.38	0.33

IV.III Dividend re-investment

Finally, we consider the robustness of our results to different re-investment schemes. There is not a single convention prevailing in the literature. For example, Welch and Goyal (2008) accumulate dividends without adjusting for returns, Cochrane (2008) assumes that dividends received during the year are re-invested in the market portfolios, while in van Binsbergen and Koijen (2010) dividends received during the year are re-invested at the risk-free rate. Chen (2009) and van Binsbergen and Koijen (2010) caution that the assumption of market-reinvestment may possibly complicate disentangling dividend predictability from return predictability because market-reinvested dividends contain a substantial return component.

In the main paper, we follow the assumption by van Binsbergen and Koijen (2010) of risk-free rate re-investment. In this section, we reproduce our volatility decompositions reconstructed using market re-invested dividends and non-reinvested dividends, for both monthly and annual data. The results, reported in Tables XIX–XXII, reveal that our results are not very sensitive to the choice of re-investment strategy. Our main result of a significant volatility contribution of long-run dividend expectations when the information set is extended beyond the single state variable holds for all re-investment assumptions, at both monthly and annual frequency.

Table XIX: Volatility decomposition: No dividend re-investment (annual data).

This table reproduces the annual volatility decompositions as reported in Tables 2–4 of the paper, with the difference that the 12-month dividend growth rates are constructed under the assumption that dividends received during the 12 months are not re-invested.

	A: One state variable				B: Three state variables			
k (years)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
1	0.13	0.00	0.90	0.04	0.15	0.11	0.92	0.72
2	0.23	0.01	0.80	0.06	0.30	0.07	0.80	0.24
3	0.28	0.02	0.73	0.08	0.28	0.16	0.77	0.59
4	0.33	0.01	0.66	0.04	0.40	0.25	0.73	0.63
5	0.40	0.01	0.57	0.01	0.47	0.22	0.66	0.47
6	0.47	0.02	0.49	0.04	0.48	0.18	0.58	0.37
7	0.53	0.02	0.42	0.04	0.53	0.15	0.52	0.28
8	0.59	0.01	0.36	0.02	0.63	0.17	0.46	0.27
9	0.65	0.00	0.28	0.01	0.68	0.22	0.40	0.33
10	0.70	0.00	0.24	0.00	0.74	0.30	0.48	0.41
11	0.73	0.02	0.19	0.02	0.76	0.33	0.48	0.44
12	0.73	0.04	0.15	0.05	0.77	0.34	0.45	0.45
13	0.74	0.06	0.12	0.07	0.79	0.32	0.38	0.40
14	0.76	0.08	0.09	0.10	0.85	0.27	0.36	0.31
15	0.79	0.09	0.04	0.11	0.89	0.26	0.31	0.29
	C: LASSO				D: Model averaging			
k (years)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
1	0.13	0.08	0.90	0.66	0.15	0.11	0.92	0.73
2	0.24	0.19	0.81	0.80	0.30	0.08	0.81	0.25
3	0.34	0.25	0.82	0.72	0.28	0.16	0.78	0.59
4	0.56	0.25	0.75	0.44	0.40	0.25	0.73	0.63
5	0.62	0.26	0.73	0.42	0.48	0.22	0.66	0.47
6	0.72	0.29	0.70	0.39	0.50	0.18	0.59	0.36
7	0.75	0.28	0.67	0.38	0.54	0.15	0.52	0.28
8	0.86	0.24	0.66	0.28	0.64	0.17	0.47	0.27
9	0.95	0.27	0.68	0.28	0.70	0.22	0.42	0.32
10	0.94	0.25	0.62	0.27	0.75	0.30	0.47	0.40
11	0.92	0.30	0.56	0.32	0.76	0.35	0.48	0.45
12	0.96	0.46	0.60	0.48	0.77	0.36	0.45	0.46
13	0.95	0.33	0.55	0.35	0.80	0.33	0.38	0.41
14	0.94	0.31	0.54	0.33	0.86	0.28	0.37	0.32
15	0.92	0.40	0.54	0.43	0.89	0.26	0.30	0.29

Table XX: Volatility decomposition: No dividend re-investment (monthly data).

This table reproduces the monthly volatility decompositions as reported in Tables 2–4 of the paper, with the difference that the 12-month dividend growth rates are constructed under the assumption that dividends received during the 12 months are not re-invested.

	A: One state variable				B: Three state variables			
k (months)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
12	0.12	0.01	0.88	0.12	0.13	0.14	0.88	1.10
24	0.22	0.00	0.78	0.01	0.23	0.09	0.80	0.38
36	0.29	0.02	0.72	0.07	0.31	0.12	0.76	0.38
48	0.34	0.02	0.66	0.07	0.41	0.22	0.72	0.54
60	0.41	0.01	0.57	0.02	0.45	0.20	0.64	0.44
72	0.46	0.00	0.50	0.00	0.46	0.14	0.57	0.31
84	0.52	0.00	0.43	0.00	0.52	0.13	0.50	0.25
96	0.58	0.00	0.38	0.01	0.59	0.16	0.45	0.26
108	0.63	0.00	0.31	0.01	0.65	0.23	0.42	0.36
120	0.66	0.00	0.28	0.01	0.69	0.29	0.47	0.43
132	0.70	0.00	0.24	0.00	0.71	0.33	0.47	0.46
144	0.71	0.02	0.21	0.02	0.72	0.31	0.42	0.44
156	0.74	0.03	0.18	0.04	0.74	0.29	0.38	0.39
168	0.77	0.03	0.14	0.04	0.79	0.25	0.36	0.32
180	0.82	0.04	0.09	0.05	0.84	0.24	0.31	0.28

	C: LASSO				D: Model averaging			
k (months)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
12	0.19	0.21	0.91	1.10	0.13	0.15	0.88	1.13
24	0.35	0.25	0.85	0.72	0.23	0.09	0.81	0.39
36	0.48	0.27	0.81	0.57	0.31	0.12	0.76	0.39
48	0.58	0.28	0.79	0.49	0.41	0.22	0.72	0.55
60	0.68	0.26	0.75	0.39	0.45	0.20	0.64	0.44
72	0.75	0.28	0.78	0.37	0.47	0.15	0.57	0.31
84	0.79	0.28	0.76	0.35	0.52	0.13	0.51	0.26
96	0.89	0.30	0.73	0.33	0.60	0.16	0.45	0.27
108	0.88	0.35	0.72	0.39	0.66	0.23	0.42	0.34
120	0.86	0.35	0.69	0.40	0.70	0.29	0.48	0.42
132	0.88	0.42	0.67	0.48	0.71	0.34	0.48	0.48
144	0.88	0.47	0.62	0.53	0.72	0.33	0.43	0.46
156	0.89	0.45	0.59	0.51	0.74	0.30	0.38	0.41
168	0.86	0.41	0.52	0.47	0.79	0.27	0.36	0.34
180	0.89	0.33	0.52	0.38	0.84	0.24	0.30	0.29

Table XXI: Volatility decomposition: Market re-investment of dividends (annual data).

This table reproduces the annual volatility decompositions as reported in Tables 2–4 of the paper, with the difference that the 12-month dividend growth rates are constructed under the assumption that dividends received during the 12 months are re-invested at the market rate of return.

	A: One state variable				B: Three state variables			
k (years)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
1	0.13	0.02	0.90	0.18	0.15	0.12	0.91	0.80
2	0.23	0.03	0.80	0.13	0.29	0.13	0.80	0.43
3	0.28	0.02	0.73	0.08	0.28	0.21	0.77	0.76
4	0.33	0.01	0.66	0.02	0.40	0.30	0.73	0.76
5	0.40	0.01	0.57	0.02	0.50	0.29	0.65	0.58
6	0.47	0.02	0.49	0.04	0.49	0.22	0.58	0.45
7	0.53	0.02	0.42	0.03	0.54	0.24	0.51	0.44
8	0.59	0.01	0.36	0.02	0.63	0.25	0.46	0.40
9	0.65	0.01	0.28	0.02	0.72	0.27	0.44	0.38
10	0.70	0.00	0.24	0.01	0.74	0.33	0.49	0.44
11	0.73	0.02	0.19	0.02	0.76	0.37	0.48	0.49
12	0.73	0.05	0.15	0.07	0.77	0.40	0.44	0.52
13	0.74	0.06	0.12	0.08	0.79	0.36	0.41	0.45
14	0.76	0.08	0.09	0.10	0.85	0.32	0.44	0.38
15	0.79	0.09	0.04	0.11	0.92	0.30	0.52	0.33

	C: LASSO				D: Model averaging			
k (years)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
1	0.13	0.05	0.90	0.38	0.15	0.12	0.92	0.80
2	0.24	0.06	0.81	0.24	0.29	0.13	0.81	0.43
3	0.34	0.12	0.81	0.36	0.28	0.21	0.78	0.75
4	0.62	0.27	0.80	0.43	0.40	0.30	0.73	0.75
5	0.60	0.34	0.72	0.56	0.51	0.29	0.66	0.58
6	0.72	0.29	0.70	0.41	0.51	0.22	0.58	0.44
7	0.75	0.31	0.74	0.42	0.55	0.24	0.52	0.43
8	0.86	0.33	0.66	0.39	0.64	0.25	0.47	0.40
9	0.91	0.36	0.69	0.39	0.74	0.27	0.45	0.37
10	0.95	0.31	0.62	0.32	0.75	0.32	0.48	0.42
11	0.95	0.30	0.56	0.32	0.77	0.38	0.48	0.49
12	0.95	0.48	0.60	0.50	0.77	0.41	0.44	0.53
13	0.93	0.48	0.58	0.51	0.80	0.37	0.39	0.46
14	0.95	0.46	0.56	0.48	0.86	0.33	0.42	0.38
15	0.94	0.51	0.54	0.55	0.92	0.31	0.49	0.33

Table XXII: Volatility decomposition: Market re-investment of dividends (monthly data).

This table reproduces the monthly volatility decompositions as reported in Tables 2–4 of the paper, with the difference that the 12-month dividend growth rates are constructed under the assumption that dividends received during the 12 months are re-invested at the market rate of return.

	A: One state variable				B: Three state variables			
k (months)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
12	0.12	0.05	0.88	0.42	0.13	0.11	0.88	0.85
24	0.22	0.05	0.78	0.21	0.23	0.12	0.81	0.52
36	0.29	0.05	0.72	0.18	0.30	0.19	0.77	0.63
48	0.34	0.05	0.66	0.14	0.38	0.27	0.72	0.70
60	0.41	0.04	0.57	0.11	0.45	0.25	0.64	0.56
72	0.46	0.02	0.50	0.05	0.47	0.20	0.57	0.42
84	0.52	0.02	0.43	0.05	0.52	0.19	0.50	0.36
96	0.58	0.02	0.38	0.04	0.59	0.20	0.45	0.34
108	0.63	0.03	0.31	0.05	0.65	0.26	0.42	0.40
120	0.66	0.02	0.28	0.03	0.69	0.38	0.47	0.55
132	0.70	0.02	0.24	0.03	0.71	0.38	0.47	0.54
144	0.71	0.00	0.21	0.01	0.72	0.40	0.43	0.56
156	0.74	0.01	0.18	0.02	0.74	0.35	0.38	0.47
168	0.77	0.02	0.14	0.02	0.79	0.33	0.39	0.42
180	0.82	0.01	0.09	0.01	0.83	0.29	0.34	0.35

	C: LASSO				D: Model averaging			
k (months)	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$	$\widehat{\sigma}(r,k)$	$\widehat{\sigma}(d,k)$	$\widehat{\sigma}(dp,k)$	$\frac{\widehat{\sigma}(d,k)}{\widehat{\sigma}(r,k)}$
12	0.20	0.18	0.91	0.90	0.13	0.11	0.88	0.86
24	0.35	0.26	0.84	0.74	0.23	0.12	0.81	0.53
36	0.48	0.31	0.82	0.65	0.30	0.19	0.77	0.63
48	0.56	0.32	0.79	0.57	0.38	0.27	0.73	0.69
60	0.66	0.32	0.75	0.49	0.45	0.25	0.64	0.55
72	0.75	0.29	0.77	0.39	0.47	0.20	0.57	0.42
84	0.78	0.32	0.77	0.41	0.52	0.19	0.51	0.36
96	0.91	0.31	0.73	0.35	0.60	0.20	0.45	0.34
108	0.87	0.38	0.72	0.44	0.66	0.26	0.42	0.39
120	0.87	0.42	0.70	0.49	0.70	0.39	0.48	0.56
132	0.87	0.48	0.67	0.55	0.71	0.41	0.48	0.57
144	0.89	0.58	0.63	0.65	0.72	0.44	0.43	0.61
156	0.87	0.48	0.58	0.55	0.74	0.38	0.37	0.50
168	0.88	0.49	0.53	0.56	0.79	0.35	0.37	0.44
180	0.89	0.40	0.51	0.45	0.83	0.29	0.32	0.35

References

Ang, A., and G. Bekaert (2007). Return predictability: Is it there? *Review of Financial Studies* 20, 651–707.

Bai, J., and S. Ng (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics* 146, 304–317.

Belloni, A., and V. Chernozhukov (2013). Least squares after model selection in high-dimensional sparse models. *Bernoulli* 19, 521–547.

van Binsbergen, J.H., and R.S.J. Koijen (2010). Predictive regressions: A present-value approach. *The Journal of Finance* 65, 1493–1471.

Boudoukh, J., M. Richardson, and R.F. Whitelaw (2008). The myth of long-horizon predictability. *Review of Financial Studies* 21, 1577–1605.

Chen, L. (2009). On the reversal of return and dividend growth predictability: A tale of two periods. *Journal of Financial Economics* 92, 128–151.

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

Golez, B., and P. Koudijs (2018). Four centuries of return predictability. *Journal of Financial Economics* 127, 248–263.

Hastie, T., R. Tibshirani, and J. Friedman (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Second Edition, Springer-Verlag, New York.

Rapach, D.E., J.K. Strauss, and G. Zhou (2013). International stock return predictability: What is the role of the United States? *The Journal of Finance* 68, 1633–1662.

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

Zou, H., and T. Hastie (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society, Series B* 67, 301–320.