



# Stock return and cash flow predictability: The role of volatility risk<sup>☆</sup>



Tim Bollerslev<sup>a,b,\*</sup>, Lai Xu<sup>a,1</sup>, Hao Zhou<sup>c,2</sup>

<sup>a</sup> Department of Economics, Duke University, Durham NC 27708, USA

<sup>b</sup> NBER and CREATES, Denmark

<sup>c</sup> PBC School of Finance, Tsinghua University, Beijing, 100083, PR China

## ARTICLE INFO

### Article history:

Available online 23 March 2015

### JEL classification:

G12  
G13  
C12  
C13

### Keywords:

Return and dividend growth predictability  
Variance risk premium  
Equilibrium pricing  
Stochastic volatility and uncertainty  
“Structural” factor GARCH

## ABSTRACT

We examine the joint predictability of return and cash flow within a present value framework, by imposing the implications from a long-run risk model that allow for both time-varying volatility and volatility uncertainty. We provide new evidence that the expected return variation and the variance risk premium positively forecast both short-horizon returns *and* dividend growth rates. We also confirm that dividend yield positively forecasts long-horizon returns, but that it does not help in forecasting dividend growth rates. Our equilibrium-based “structural” factor GARCH model permits much more accurate inference than univariate regression procedures traditionally employed in the literature. The model also allows for the direct estimation of the underlying economic mechanisms, including a new volatility leverage effect, the persistence of the latent long-run growth component and the two latent volatility factors, as well as the contemporaneous impacts of the underlying “structural” shocks.

© 2015 Elsevier B.V. All rights reserved.

## 1. Introduction

Counter to the “old” efficient market hypothesis dictum that speculative returns are largely unpredictable over time, it is now generally accepted that equity returns are both time-varying and predictable. It is also widely believed that the predictability of the aggregate stock market as a whole is the strongest over longer multi-year horizons.<sup>3</sup> At the same time, to the extent that a consensus has emerged it suggests that expected dividend growth

<sup>☆</sup> We would like to thank Michael McAleer (the Editor), two anonymous referees, Andrew Patton, George Tauchen and seminar participants at the Duke Financial Econometrics Lunch Group and the July 2013 China International Conference in Finance for their helpful comments. Bollerslev’s research was supported by a grant from the National Science Foundation (SES-0957330) to the NBER, and CREATES (DNRF78) funded by the Danish National Research Foundation.

\* Corresponding author at: Department of Economics, Duke University, Durham NC 27708, USA. Tel.: +1 919 660 1846.

E-mail addresses: [boller@econ.duke.edu](mailto:boller@econ.duke.edu) (T. Bollerslev), [lai.xu@duke.edu](mailto:lai.xu@duke.edu) (L. Xu), [zhouh@pbcfs.tsinghua.edu.cn](mailto:zhouh@pbcfs.tsinghua.edu.cn) (H. Zhou).

<sup>1</sup> Tel.: +1 919 257 0059.

<sup>2</sup> Tel.: +86 10 62790655.

<sup>3</sup> Some of the predictor variables used in establishing long-run return predictability include: dividend-price, earning-price, and other valuation ratios (Campbell and Shiller, 1988a,b; Fama and French, 1988; Lamont, 1998; Lewellen, 2004); firms’ net equity payout (Boudoukh et al., 2007) and equity issuance (Baker et al., 2000); interest-rate variables such as *t*-bill and *t*-bond rates, term spreads, and default spreads (Campbell, 1987; Fama and French, 1989; Hodrick,

rates for the aggregate market portfolio, or aggregate cash flows, are much less predictable than the expected returns.<sup>4</sup>

Much of the literature underlying these findings, and the choice of predictor variables in particular, have been guided by the present-value framework pioneered by Campbell and Shiller (1988a,b), and the implication that the dividend-price ratio, or the dividend yield, is identically equal to the expected value of the future returns discounted by the future dividend growth rates. As emphasized by Cochrane (2008, 2011), this intimate link between dividend growth and stock return predictability also implies that the seemingly stronger empirical evidence for long-run return predictability is not surprisingly accompanied by seemingly weaker empirical evidence for long-run dividend growth predictability.

Set against this background, a number of recent studies have argued that the variance risk premium, or the difference between options implied and expected variances, possesses superior forecasting power for stock market returns over shorter within-year horizons; see, e.g., Bollerslev et al. (2009), Drechsler and Yaron

1992, ); and macroeconomic variables like total investment (Cochrane, 1991), the consumption-wealth ratio (Lettau and Ludvigson, 2001), and inflation (Campbell and Vuolteenaho, 2004).

<sup>4</sup> Maio and Santa-Clara (forthcoming) have recently challenged this view, showing that for portfolios comprised of small and value stocks, the dividend-price ratio is primarily related to future changes in cash flows. With a few notable exceptions (e.g., Fama and French, 1988; Lettau and Ludvigson, 2005) cash flow predictability has historically received much less attention in the literature.

(2011), and Kelly (2011). Motivated by these more recent empirical findings, we show how explicitly incorporating priced volatility risk into the present-value framework affords important new insights into the return vis-à-vis dividend growth predictability debate across all horizons.

The reduced form VAR framework, as exemplified by Hodrick (1992) and Campbell (2001), traditionally used for empirically implementing present value relations does not naturally lend itself to the estimation of models involving priced volatility risk. Instead, we follow Sentana and Fiorentini (2001) and Rigobon (2003) in designing a “structural” factor GARCH model, in which the factors exhibit time-varying volatility. The dynamics of the factors is derived endogenously from an extended long-run risk model explicitly incorporating time-varying consumption volatility and volatility-of-volatility, or economic uncertainty. The resulting econometric model separately identifies the long-run risk, volatility, and economic uncertainty components, as well as the corresponding structural shocks and their contemporaneous impact on both returns and dividend growth.

Estimating the “structural” factor GARCH model by standard GMM techniques on data for the S&P 500 market portfolio, we confirm existing empirical evidence that the dividend-price ratio is useful for predicting long-horizon multi-year returns, but that it has no predictive power for dividend growth.<sup>5</sup> More important, we document a number of new results pertaining to the predictability of the volatility factors. In particular, while the variance risk premium shows significant predictability for returns over short within-year horizons, it also helps predict dividend growth. Similarly, the expected return variation appears to be very informative for predicting dividend growth.

These results are consistent with the findings in Kojien and Nieuwerburgh (2011) that the high-frequency component of the dividend-price ratio, which in our setup is driven by two separate volatility factors, contains useful information for predicting expected dividend growth. Our results are also related to Binsbergen et al. (2012) and their findings that the term structure of equity risk premia is particularly steep in the short end, while standard asset pricing models without priced volatility risk typically imply higher equity premia at the long end.

In addition to the new empirical evidence pertaining to the short-run predictability of returns and dividend growth, by explicitly identifying the systematic risk factors at work, our “structural” factor GARCH approach also helps shed new light on the underlying economic mechanisms. Specifically, we find that the long-run expected growth component is highly persistent with a first-order autocorrelation coefficient close to one ( $\rho_x = 0.988$ ) at the monthly level, consistent with the idea in Bansal and Yaron (2004) that it acts as the most important driver of the risk premium dynamics over long horizons.<sup>6</sup> The model also clearly differentiates and is able to accurately estimate the persistence of the consumption volatility component ( $\rho_\sigma = 0.64$ ) and the volatility-of-volatility, or economic uncertainty, component ( $\rho_q = 0.46$ ), advocated by Bollerslev et al. (2009), both of which are intimately linked to the shorter-run predictability patterns in the data. In terms of the underlying “structural” shocks, we find a negative relationship between the long-run growth and consumption

volatility shocks (akin to a “leverage effect”), as well as a negative relationship between the consumption volatility and volatility uncertainty shocks (interpretable as a separate new “leverage effect”). The price-dividend ratio also responds negatively to both consumption volatility and volatility uncertainty shocks.<sup>7</sup>

The basic motivation behind the new “structural” factor GARCH model is in line with a growing recent literature seeking to explicitly incorporate the effect of stochastic volatility in asset pricing models. For example, Bansal et al. (2014) demonstrate that ignoring the variation in volatility leads to counter-intuitive economic interpretation of risk premium dynamics. Similarly, Campbell et al. (2013) examine the cross-sectional return predictability in an ICAPM framework that allows for stochastic volatility.<sup>8</sup> In contrast to these studies, our focus is on the joint predictability of returns and cash flows within the context of a “structural” econometric model explicitly designed to accommodate time-varying volatility in an internally consistent fashion. Recent studies by Binsbergen and Kojien (2010) and Piatti and Trojani (2012) have also relied on a latent variable approach with heteroskedastic shocks for incorporating the effect of time-varying volatility within a present-value framework. Importantly, however, we differ from both of these studies by specifying an empirically more realistic two-factor volatility structure and by explicitly including both the actual and risk-neutral expected variation in the formulation and estimation of the model.<sup>9</sup>

The rest of the paper is organized as follows. Section 2 presents the equilibrium asset pricing model underlying our empirical investigations. Section 3 describes the data and the formulation of the “structural” factor GARCH model and the GMM-based parameter estimation results. Section 4 details the return and cash flow predictability implied by the model, and contrast the results with those obtained by other less structured reduced form estimation procedures. Section 5 concludes.

## 2. Asset pricing model

Our equilibrium-based approach combines the long-run risk model pioneered by Bansal and Yaron (2004), with the model in Bollerslev et al. (2009) explicitly allowing for stochastic volatility-of-volatility, or time-varying economic uncertainty. This general setup naturally accommodates the magnitude of both the equity and variance risk premia, as well as the long- and short-horizon predictability patterns in the returns and cash flows within a unified framework.

### 2.1. Model setup and assumptions

Following the long-run risk literature, we assume an endowment economy with a representative agent equipped with Epstein and Zin (1991) recursive preferences. The logarithm of the intertemporal marginal substitution for this agent may consequently be expressed as,

$$m_{t+1} = \theta \log \delta - \frac{\theta}{\psi} \Delta c_{t+1} + (\theta - 1)r_{c,t+1}, \quad (1)$$

<sup>5</sup> Compared to earlier empirical findings based on univariate regressions (Rozeff, 1984; Fama and French, 1988; Campbell and Shiller, 1988b) and traditional present-value homoskedastic VAR's (Hodrick, 1992; Campbell, 2001; Cochrane, 2008), our “structural” factor GARCH model results in much sharper inference, with the actual point estimates systematically falling within the standard error bands obtained from the more conventional procedures.

<sup>6</sup> Nakamura et al. (2012) have recently shown how the long-run growth factor may also be identified from cross-country aggregate consumption data under additional simplifying assumptions.

<sup>7</sup> The importance of economic uncertainty for explaining asset prices has also recently been emphasized from different perspectives by Bekaert et al. (2009), Nieto and Rubio (2011), and Corradi et al. (2013), among others.

<sup>8</sup> Our “structural” factor GARCH estimate for the persistence in consumption volatility  $\rho_\sigma$ , and in turn the effect of allowing for time-varying volatility, are much larger than the estimates reported in Campbell et al. (2013) based on simple VAR procedures and imprecise variance measures.

<sup>9</sup> Other recent studies seeking to incorporate more realistic two-factor volatility structures in the standard long-run risk model include Zhou and Zhu (2013), Branger and Völkert (2012), and Branger et al. (2011), among others.

where  $r_{c,t+1} \equiv \log(R_{c,t+1})$  refers to the logarithmic return on the consumption asset,  $\Delta C_{t+1} \equiv \log(C_{t+1}/C_t)$  denotes the growth rate of consumption,  $0 < \delta < 1$  is the time discount factor,  $\gamma > 0$  denotes the risk aversion parameter, and  $\theta \equiv \frac{1-\gamma}{1-\psi}$  where  $\psi > 0$  refers to the intertemporal elasticity of the substitution. As is standard in the long-run risk literature, we will assume that  $\gamma > 1$ , implying that the representative agent is more risk averse than log utility, and that  $\psi > 1$ , and therefore  $\theta < 0$ , implying a preference for early resolution of uncertainty.

Let  $x_t$  denote the long-run mean of consumption growth as in [Bansal and Yaron \(2004\)](#), and  $\sigma_t^2$  and  $q_t$  refer to two separate volatility factors along the lines of [Bollerslev et al. \(2009\)](#). For notational convenience, collect the consumption growth  $\Delta c_t$ , the log dividend growth  $\Delta d_t$ , and the latent state variables in the vector  $Y_t = (\Delta c_t, x_t, \sigma_t^2, q_t, \Delta d_t)'$ . The importance of allowing for multiple volatility factors in accurately describing both short- and long-horizon time-varying return and volatility dynamics has also recently been highlighted by [Bollerslev et al. \(2009\)](#), [Drechsler and Yaron \(2011\)](#), [Bollerslev et al. \(2012\)](#), [Zhou and Zhu \(2013\)](#), [Branger and Völkert \(2012\)](#), among others.

We will assume that the state vector  $Y_t$  has affine conditional mean and variance dynamics,

$$Y_{t+1} = \mu + FY_t + HG_t Z_{t+1}, \tag{2}$$

where  $Z_{t+1} \equiv [z_{c,t+1}, z_{x,t+1}, z_{\sigma,t+1}, z_{q,t+1}, z_{d,t+1}]'$  denotes a vector of independent standard normally distributed shocks. We rank all of the “structural” consumption shocks, including the two volatility shocks  $z_{\sigma,t}$  and  $z_{q,t}$ , before shocks to dividends  $z_{d,t}$ . Based on the intuition that level shocks are more “fundamental” than shocks to volatility, we also put the  $z_{c,t}$  and  $z_{x,t}$  shocks before the two volatility shocks. The conditional mean of  $Y_t$  is in turn determined by the constant vector  $\mu$  and the loading matrix  $F$ . We assume that this loading matrix takes the sparse form,

$$F = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & \rho_x & 0 & 0 & 0 \\ 0 & 0 & \rho_\sigma & 0 & 0 \\ 0 & 0 & 0 & \rho_q & 0 \\ 0 & \phi_{dx} & 0 & 0 & \rho_d \end{pmatrix}, \tag{3}$$

in which the diagonal elements characterize the own lagged dependencies and the off-diagonal elements describe the dynamic first-order cross dependencies. In particular,  $\phi_{dx}$  allows the dividend growth rate  $\Delta d_{t+1}$  to directly load on the lagged long-run consumption growth component  $x_t$ . Allowing  $\Delta d_{t+1}$  to also depend on its own lag permits a non-redundant pricing effect of dividend growth risk on the equity premium. Restricting this coefficient  $\rho_d$  to be zero reduces the model's growth dynamics to that of a “standard” long-run risk model. However, our estimates of the model discussed below strongly rejects such a specification.

The conditional second-order dynamics of the state vector is determined by the time-varying diagonal volatility matrix  $G_t$  and the constant loading matrix  $H$ ,

$$G_t = \begin{pmatrix} \sigma_t & 0 & 0 & 0 & 0 \\ 0 & \sqrt{q_t} & 0 & 0 & 0 \\ 0 & 0 & \sqrt{q_t} & 0 & 0 \\ 0 & 0 & 0 & \sqrt{q_t} & 0 \\ 0 & 0 & 0 & 0 & \sigma_t \end{pmatrix} \tag{4}$$

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \varphi_x & 0 & 0 & 0 \\ 0 & \varphi_x s_{\sigma,x} & 1 & 0 & 0 \\ 0 & \varphi_x s_{q,x} & s_{q,\sigma} & \varphi_q & 0 \\ 0 & \varphi_x s_{d,x} & s_{d,\sigma} & \varphi_q s_{d,q} & \varphi_d \end{pmatrix}.$$

Our choice of  $G_t$  differs from the models in [Drechsler and Yaron \(2011\)](#) and [Branger and Völkert \(2012\)](#) by allowing both  $x_{t+1}$  and

$\sigma_{t+1}^2$  to have time-varying volatility  $\sqrt{q_t}$ . Our choice of  $G_t$  also nests the model in [Bollerslev et al. \(2009\)](#) by zeroing out the long-run growth component, equating the dividend and consumption growth, and fixing  $s_{i,j} = 0$  for  $i \neq j$ , thereby rendering  $H$  diagonal.<sup>10</sup> Identification of the lower triangular volatility loading matrix  $H$  is effectively accomplished through heteroskedasticity, and cross-dependencies between the different state variables implied by the form of the time-varying volatility.

Further, denoting the columns of  $H \equiv [h_1, h_2, h_3, h_4, h_5]$ , the “square” of  $HG_t$  may be conveniently expressed in affine form as,

$$HG_t G_t' H' = \sum_{j=1,5} h_j h_j' \sigma_t^2 + \sum_{j=2,3,4} h_j h_j' q_t. \tag{5}$$

This two-factor volatility structure is distinctly different from the one-factor setup recently employed in [Campbell et al. \(2013\)](#). As discussed in more detail below, it affords an empirically much more realistic description of the return and cash flow dynamics, and in turn the predictability patterns obtained by imposing the equilibrium-based restrictions.

### 2.2. Model implications

In order to deduce the “structural” model restrictions that guide our empirical analysis, we begin by solving the consumption-based asset pricing model using similar techniques to the ones in [Bansal and Yaron \(2004\)](#), [Bansal et al. \(2007b\)](#), and [Drechsler and Yaron \(2011\)](#). In the spirit of [Campbell \(1993, 1996\)](#), we then substitute out the hard-to-measure consumption and its volatility dynamics with directly observable market return and its variance measures.

Standard solution methods applied in the long-run risk literature readily imply that the stochastic discount factor  $m_{t+1}$ , the return on consumption  $r_{c,t+1}$ , and the market return on dividends  $r_{t,t+1}$ , must satisfy

$$\begin{aligned} m_{t+1} - E_t(m_{t+1}) &= -\Lambda' HG_t Z_{t+1}, \\ r_{c,t+1} - E_t(r_{c,t+1}) &= \Lambda_c' HG_t Z_{t+1}, \\ r_{t,t+1} - E_t(r_{t,t+1}) &= \Lambda_d' HG_t Z_{t+1}, \end{aligned} \tag{6}$$

where  $\Lambda = \gamma e_1 + \kappa_1(1 - \theta)A$ , for  $A = (0, A_x, A_\sigma, A_q, 0)$ , denotes the price of risk for the factor shocks,  $\Lambda_c = e_1 + \kappa_1 A$ ,  $\Lambda_d = e_5 + \kappa_{d,1} A_d$ ,  $\kappa_1$  and  $\kappa_{d,1}$  refer to the [Campbell and Shiller \(1988b\)](#) log-linearization constants based on the “usual” approximations for consumption return  $r_{c,t+1} \approx \kappa_0 + \kappa_1 v_{t+1} - v_t + \Delta c_{t+1}$  and the aggregate market return  $r_{t,t+1} \approx \kappa_{d,0} + \kappa_{d,1} w_{t+1} - w_t + \Delta d_{t+1}$ , respectively, and the two selection vectors are defined by  $e_1 \equiv [1, 0, 0, 0, 0]'$  and  $e_5 \equiv [0, 0, 0, 0, 1]'$ .<sup>11</sup> Given these expressions, it is possible to solve for the market return variance  $Var_t(r_{t,t+1})$ , the variance risk premium  $VRP_t$ , and the log dividend-price ratio  $dp_t$ , as

$$Var_t(r_{t,t+1}) = (1 + \kappa_{d,1} A_{d,d})^2 \varphi_d^2 \sigma_t^2 + \sum_{j=2,3,4} \Lambda_d' h_j h_j' \Lambda_d q_t, \tag{7}$$

<sup>10</sup> We also experimented with two alternative setups, one closer to [Drechsler and Yaron \(2011\)](#) with  $G_t = \text{diag}[\sigma_t, \sqrt{q_t}, \sigma_t, \sqrt{q_t}, \sigma_t]$ , and the other one closer to [Branger and Völkert \(2012\)](#) with  $G_t = \text{diag}[\sigma_t, \sigma_t, \sigma_t, \sqrt{q_t}, \sigma_t]$ , resulting in qualitatively similar predictability results to the ones reported below. However, both of these alternative specifications were rejected at conventional significance levels by the corresponding GMM-based  $J$ -tests for over-identifying restrictions. Further details concerning these alternative models and empirical results are reported in the supplementary [Appendix A](#).

<sup>11</sup> As further detailed in the supplementary [Appendix A](#), the market prices of risks also depend implicitly on the coefficients in the wealth-consumption ratio  $v_t = A_0 + [0, A_x, A_\sigma, A_q, 0]' Y_t$  and the price-dividend ratio  $w_t \equiv -dp_t = A_{d,0} + [0, A_{d,x}, A_{d,\sigma}, A_{d,q}, A_{d,d}]' Y_t$ .

$$VRP_t = \left( \sum_{j=1,5} \Lambda'_d h_j h'_j \Lambda_d s_{q,1} + \sum_{j=2,3,4} \Lambda'_d h_j h'_j \Lambda_d s_{q,2} \right) q_t, \tag{8}$$

$$dp_t = -A_{0,d} - A_{d,x}x_t - A_{d,\sigma}\sigma_t^2 - A_{d,q}q_t - A_{d,d}\Delta d_t, \tag{9}$$

where  $s_{q,1} = -(\varphi_x s_{\sigma,x} h'_2 + h'_3)\Lambda$  and  $s_{q,2} = -(\varphi_x s_{q,x} h'_2 + s_{q,\sigma} h'_3 + \varphi_q h'_4)\Lambda$ . We will impose these “structural” restrictions on the empirical model estimated below.

Even though our empirical strategy of substituting out consumption means that some of the parameters in the autoregressive loading matrix  $F$  and the volatility loading matrix  $H$  are not identified, the specific structures for the two loading matrices still provide useful guidance on how to restrict the dynamics. In particular, denote the sub-vector of  $Y_t$  that excludes consumption growth by  $f_t \equiv [\sigma_t^2, q_t, \Delta d_t, x_t]'$ , it follows that

$$f_{t+1} = \mu + \rho f_t + S \epsilon_{t+1}, \tag{10}$$

where

$$\rho = \begin{pmatrix} \rho_\sigma & 0 & 0 & 0 \\ 0 & \rho_q & 0 & 0 \\ 0 & 0 & \rho_d & \phi_{dx} \\ 0 & 0 & 0 & \rho_x \end{pmatrix} \quad S = \begin{pmatrix} 1 & 0 & 0 & s_{\sigma,x} \\ s_{q,\sigma} & 1 & 0 & s_{q,x} \\ s_{d,\sigma} & s_{d,q} & 1 & s_{d,x} \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{11}$$

and the vector of innovations  $\epsilon_{t+1} \equiv [\sqrt{q_t}z_{\sigma,t+1}, \varphi_q \sqrt{q_t}z_{q,t+1}, \varphi_d \sigma_t z_{d,t+1}, \varphi_x \sqrt{q_t}z_{x,t+1}]'$  is conditionally heteroskedastic.<sup>12</sup>

### 3. “Structural” estimation results

The consumption-based asset pricing model with volatility uncertainty, outlined in the previous section, imposes a number of restrictions pertaining to the dynamic dependencies and possible feedback effects between the expected variance, the variance risk premium, the dividend growth rate, and the dividend-price ratio. Our new “structural” factor GARCH model is designed to honor these restrictions within a tractable econometric framework.

#### 3.1. Data description

Our empirical investigations are based on end-of-month S&P 500 index returns, as a proxy for the aggregate market portfolio, and the S&P 500 dividend payments, as a proxy for the corresponding aggregate cash flows. All of our S&P 500 data are obtained from DataStream, and cover the period from January 1990 to November 2011, for a total of 262 monthly observations.<sup>13</sup>

Following standard practice in the literature, we use the trailing 12-month dividend-price ratio to account for the strong seasonality inherent in the dividend payouts; see, e.g., the discussion in [Bollerslev and Hodrick \(1995\)](#). Accordingly, the month  $t$  log dividend-price ratio  $dp_t$ , is defined by,

$$dp_t = \log \left( \frac{Div_{t-11} + \dots + Div_t}{12P_t} \right), \tag{12}$$

where  $Div_t$  denotes the dividend payments from the end-of-month  $t - 1$  to the end-of-month  $t$ , and  $P_t$  denotes the end-of-month  $t$  price. Our measures for the month  $t + 1$  log dividend growth rate

$\Delta d_{t+1}$  and the log returns including dividends  $r_{t,t+1}$ , are similarly defined from this ratio as,

$$\Delta d_{t+1} = \log \left( \frac{Div_{t-10} + \dots + Div_{t+1}}{Div_{t-11} + \dots + Div_t} \right), \tag{13}$$

$$r_{t,t+1} = \log \left( \frac{P_{t+1} + \frac{Div_{t-10} + \dots + Div_{t+1}}{12}}{P_t} \right), \tag{14}$$

with longer-run dividend growth rates and multi-period returns obtained by summation.

We consider three distinct empirical variation measures: the options implied variation  $IV_t$ , the expected return variation  $ERV_t$ , and the variance risk premium  $VRP_t$ . Our measure for the options implied variation is based on the square of the Chicago Board of Options Exchange (CBOE) VIX volatility index. This model-free measure is (approximately) equal to the market risk-neutral, or  $\mathbb{Q}$ , expectation of the one-month-ahead return variation under very general assumptions. Our construction of the corresponding actual, or  $\mathbb{P}$ , expectation, is based on the linear projection of the monthly realized variance  $RV_{t,t+1}$  on its lagged daily  $RV_{t-\frac{1}{22},t}$ , weekly  $RV_{t-\frac{5}{22},t}$ , and monthly  $RV_{t-1,t}$  values, along with the implied variation  $IV_t$ ; i.e.,<sup>14</sup>

$$ERV_{t,t+1} = \alpha_0 + \alpha_1 RV_{t-\frac{1}{22},t} + \alpha_2 RV_{t-\frac{5}{22},t} + \alpha_3 RV_{t-1,t} + \alpha_4 IV_t. \tag{15}$$

This mimics the popular HAR-RV model proposed by [Corsi \(2009\)](#). Importantly, the addition of  $IV_t$  as an additional right-hand-side variable imbues the formulation in (15) with an additional persistent long-run predictor variable, which in the traditional HAR-RV model would be captured by longer-run realized variation measures.<sup>15</sup> Finally, our measure for the variance risk premium is simply given by the difference between our risk-neutral and statistical expectations of the one-month-ahead return variation; i.e.,  $VRP_t = IV_t - ERV_t$ .

To illustrate the basic features of the different variables, [Fig. 1](#) plots the monthly time series of stock returns, dividend growth rates, dividend-price ratios, and variance risk premia. The large losses in market values and the increased volatility during the recent economic downturn are immediately evident in the plots of the returns and cash flows. The plot for the dividend-yields shows a sharp drop throughout the 1990s, but an increase after the burst of the tech bubble in 2001, reaching a new peak in the fourth quarter of 2008 around the advent of the global financial crisis and the stock market crash.<sup>16</sup> The variance risk premium shown in the last panel is on average positive with occasional negative spikes, the largest of which occurs in the fall of 2008 at the onset of the financial crises. Summary statistics for the same four variables, along with the options implied and expected variation measures underlying the variance risk premium, are reported in [Table 1](#).

<sup>12</sup> The value of  $\mu$  is immaterial to all of our predictability results. Also, the reordering of the elements in  $f_t$  relative to  $Y_t$  merely serves to facilitate comparisons with other benchmark models below, and does not affect any of the results.

<sup>13</sup> While the S&P 500 data are obviously available over a much longer sample period, some of the key variation measures employed in our analysis are only available starting in 1990.

<sup>14</sup> Our regression-based estimates of the  $\alpha$ 's rely on overlapping daily observations for all of the variation measures, thus implicitly assuming that the same relationship holds every day of the month. This greatly enhances the accuracy of the estimates compared to the estimates obtained by the use of non-overlapping monthly observations only.

<sup>15</sup> We also experimented with decomposing the realized variation measures into their continuous and discontinuous parts. Although this often helps for shorter-run forecasting, consistent with the results in [Andersen et al. \(2007\)](#), we found that the monthly forecasts and  $R^2$ 's from these more elaborate models were virtually the same as the ones from the simple-to-implement HAR-RV type formulation in (15).

<sup>16</sup> The sharp decline observed in the 1990s has been attributed to firms' substitution of dividend payments by share repurchases; see, e.g., [Kojien and Nieuwerburgh \(2011\)](#), along with the earlier related discussion in [Bagwell and Shoven \(1989\)](#).

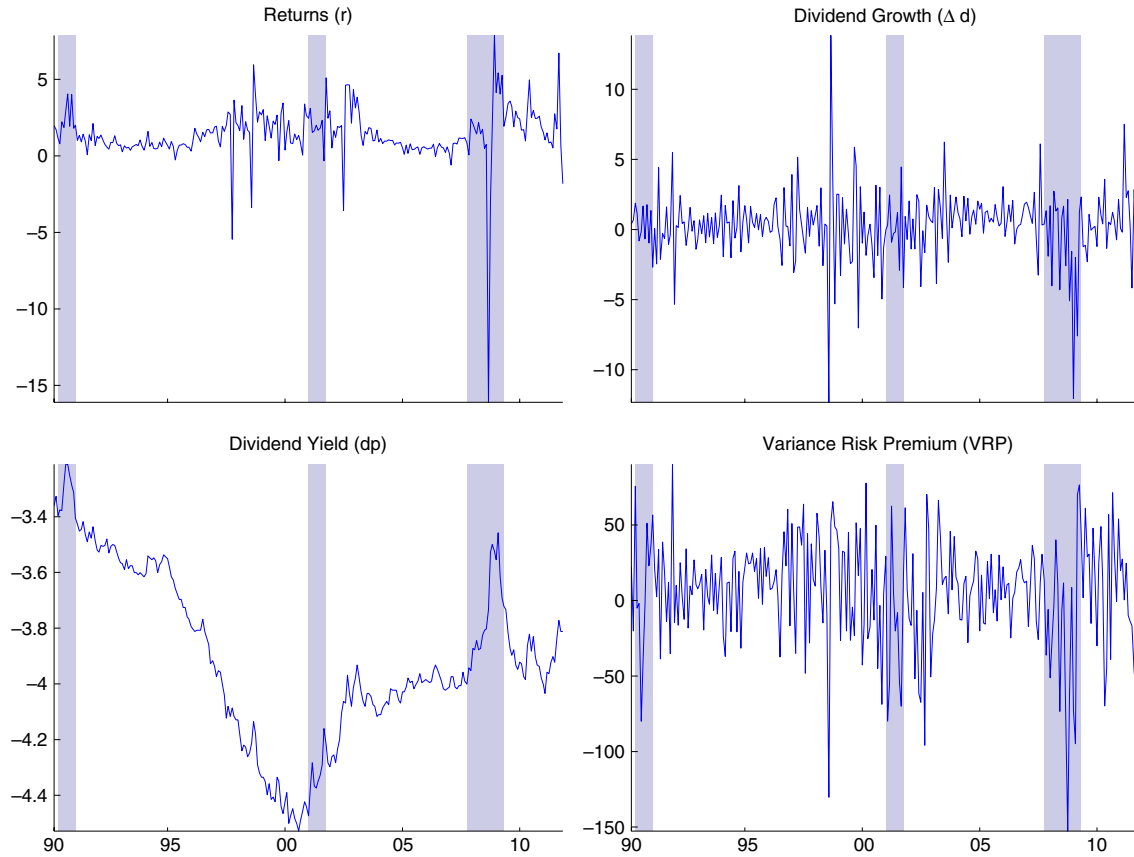


Fig. 1. Returns and dividends.

We turn next to our new present value framework and “structural” model designed to describe these general features and inherent dynamic dependencies.

3.2. “Structural” factor GARCH

The dynamics of the asset pricing model in Section 2 is succinctly summarized by the state vector  $f_t$  and Eqs. (10) and (11). The state vector  $f_t$  is, of course, not directly observable. To circumvent this, we define the “observable” state vector  $X_t \equiv [ERV_t, VRP_t, \Delta d_t, dp_t]'$ . From the solution of the model, the  $X_t$  vector is directly related to the latent  $f_t$  vector by the linear equations,<sup>17</sup>

$$X_t = \mu_X + Qf_t \quad Q = \begin{pmatrix} Q_{1,1} & Q_{1,2} & 0 & 0 \\ 0 & Q_{2,2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -A_{d,\sigma} & -A_{d,q} & -A_{d,d} & -A_{d,x} \end{pmatrix}, \quad (16)$$

where  $Q_{1,1} = (1 + \kappa_{d,1}A_{d,d})^2\varphi_d^2\rho_\sigma$ ,  $Q_{1,2} = \sum_{j=2,3,4} \Lambda_d h_j h_j' \Lambda_d \rho_q$ , and  $Q_{2,2} = (1 + \kappa_{d,1}A_{d,d})^2 s_{q,1} + \sum_{j=2,3,4} \Lambda_d' h_j h_j' \Lambda_d s_{q,2}$ . Given the standard set of assumptions about the structural parameter values typically employed in the long-run risk literature, all of the  $Q$  parameters would be positive. Conversely,  $A_{d,\sigma}$ ,  $A_{d,q}$ , and  $A_{d,d}$  would all be negative, while  $A_{d,x}$  is naturally expected to be positive.

Now combining the model for  $f_t$  in Eqs. (10) and (11) with the expression for  $X_t$  in Eq. (16), it follows that

$$BX_{t+1} = \tilde{\mu} + \tilde{\rho}BX_t + \tilde{S}\tilde{\epsilon}_{t+1}, \quad \tilde{\epsilon}_{t+1} = \tilde{G}_t\tilde{z}_{t+1}, \quad (17)$$

where  $\tilde{G}_t = \text{diag}[Q_{1,1}\sqrt{q_t}, Q_{2,2}\varphi_q\sqrt{q_t}, \varphi_d\sigma_t, -A_{d,x}\varphi_x\sqrt{q_t}]$ , and<sup>18</sup>

$$B = \begin{pmatrix} 1 & -\frac{Q_{1,2}}{Q_{2,2}} & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ A_{d,\sigma} & \frac{Q_{1,1}A_{d,q} - A_{d,\sigma}Q_{1,2}}{Q_{1,1}Q_{2,2}} & \frac{\rho_d}{1 - \kappa_{d,1}\rho_d} & 1 \end{pmatrix} \quad (18)$$

$$\tilde{\rho} = \begin{pmatrix} \rho_\sigma & 0 & 0 & 0 \\ 0 & \rho_q & 0 & 0 \\ 0 & 0 & \rho_d & \frac{\phi_{dx}}{-A_{d,x}} \\ 0 & 0 & 0 & \rho_x \end{pmatrix}$$

$$\tilde{S} = \begin{pmatrix} 1 & 0 & 0 & \frac{Q_{1,1}}{-A_{d,x}}s_{\sigma,x} \\ \frac{Q_{2,2}}{Q_{1,1}}s_{q,\sigma} & 1 & 0 & \frac{Q_{2,2}}{-A_{d,x}}s_{q,x} \\ \frac{1}{Q_{1,1}}s_{d,\sigma} & \frac{1}{Q_{2,2}}s_{d,q} & 1 & \frac{1}{-A_{d,x}}s_{d,x} \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (19)$$

Multiplying the “structural” VAR in Eq. (17) by  $B^{-1}$ , the corresponding reduced form VAR(1) representation for  $X_{t+1}$  becomes,

$$X_{t+1} = B^{-1}\tilde{\mu} + \Phi X_t + u_{t+1}, \quad (20)$$

where  $\Phi = B^{-1}\tilde{\rho}B$ ,  $u_{t+1} = \Phi_0^{-1}\tilde{\epsilon}_{t+1}$ , and  $\Phi_0^{-1} = B^{-1}\tilde{S}$ . As this representation makes clear, ignoring the heteroskedasticity in the

<sup>17</sup> Additional details concerning the solution of the model is available in the supplementary Appendix A.

<sup>18</sup> As explained in more detail in the supplementary Appendix A, the matrix  $B$  matrix is obtained from the matrix  $Q$  by normalizing its diagonal elements to unity.

**Table 1**

Summary statistics. The table reports standard summary statistics and correlations for the S&P 500 return  $r_{t,t+1}$ , dividend growth rate  $\Delta d_t$ , dividend-price ratio  $dp_t$ , options implied variance  $IV_t$ , expected variance  $ERV_t$ , and variance risk premium  $VRP_t$ . The returns, dividend growth, and dividend-price ratio are all in annualized percentage form. All of the variance variables are in monthly percentage form. The sample period extends from February 1990 to November 2011, for a total of 262 monthly observations.

	Mean	Std	Skew	Kurt	AC1
$r_{t,t+1}$	8.19	15.33	-0.76	4.48	0.07
$\Delta d_t$	3.92	8.79	-0.46	10.02	-0.26
$dp_t$	-3.91	0.31	0.08	2.32	0.98
$IV_t$	40.30	36.47	3.23	18.07	0.81
$ERV_t$	28.54	36.64	4.62	30.08	0.69
$VRP_t$	11.75	14.93	-3.37	38.42	0.27

Correlations						
	$r_{t,t+1}$	$\Delta d_t$	$dp_t$	$IV_t$	$ERV_t$	$VRP_t$
$r_{t,t+1}$	1.00	0.34	-0.03	-0.42	-0.48	0.15
$\Delta d_t$		1.00	-0.02	-0.25	-0.25	-0.01
$dp_t$			1.00	-0.05	-0.02	-0.07
$IV_t$				1.00	0.92	0.19
$ERV_t$					1.00	-0.21
$VRP_t$						1.00

reduced form shocks  $u_{t+1}$ , and interpreting the model for  $X_{t+1}$  in (17) as a standard homoskedastic VAR(1), the  $B$  and  $\tilde{S}$  matrices would not be jointly identified. In empirical macroeconomics, this lack of identification is usually “solved” by imposing that  $\Phi_0$  is lower triangular. However, as argued by Sentana and Fiorentini (2001), Rigobon (2003) and Rigobon and Sack (2003), among others, under the maintained assumption that the underlying “structural” shocks are independent, it is possible to identify the  $\Phi_0$  matrix, and in turn both  $B$  and  $\tilde{S}$ , through the heteroskedasticity in  $\tilde{\epsilon}_{t+1}$ .

Meanwhile, rather than specifying the time-varying covariance matrix for the “structural” shocks to be an explicit function of the latent  $q_t$  and  $\sigma_t^2$  risk factors, in the implementation reported on below we adopt a more flexible and empirically realistic GARCH approach for characterizing the dynamic dependencies in  $\tilde{\epsilon}_{t+1}$ . Specifically, let  $\Sigma_{t+1}$  denote the conditional covariance matrix of  $\tilde{\epsilon}_{t+1}$ . We will then assume that  $\Sigma_{t+1}$  may be described by the following relatively simple yet flexible diagonal GARCH(1,1) model,

$$\text{diag}(\Sigma_{t+1}) = (I - \Gamma - \Upsilon)\theta_0^{-1}\varpi_u + \Gamma\text{diag}(\Sigma_t) + \Upsilon\tilde{\epsilon}_t^2, \quad (21)$$

where  $\theta_0 = \Phi_0^{-1} \odot \Phi_0^{-1}$ , and  $\varpi_u$  denotes the unconditional covariance matrix of the reduced form shocks  $u_{t+1} = \Phi_0^{-1}\tilde{\epsilon}_{t+1}$ . Consequently, the second order dynamics of  $u_{t+1}$  will follow the more complicated non-diagonal GARCH(1,1) structure,<sup>19</sup>

$$\text{vec}(\Omega_{t+1}) = \Theta_1(I - \Gamma - \Upsilon)\theta_0^{-1}\varpi_u + \Theta_1\Gamma\theta_0^{-1}\text{diag}(\Omega_t) + \Theta_1\Upsilon\theta_2\text{vec}(u_t u_t'). \quad (22)$$

By explicitly parameterizing this implied conditional heteroskedasticity in  $u_{t+1}$ , it is possible to identify and separately estimate all of the “structural” parameters in (17)–(19).

The diagonal GARCH(1,1) model in (21) freely parametrizes the persistence in the “structural” shocks. Consistent with our initial estimates of the model, and the implication from the underlying consumption-based asset pricing model, we impose the restriction that the autoregressive dependencies in the GARCH expected

variance and the dividend-price ratio are the same, i.e.,  $\Gamma_{1,1} + \Upsilon_{1,1} = \Gamma_{4,4} + \Upsilon_{4,4} = \rho_q$ . Guided by our initial diagnostic tests, we also restrict the dividend growth shock to have only ARCH and no GARCH effect, i.e.,  $\Gamma_{3,3} = 0$ . All-in-all, this leaves us with a total of nine conditional variance parameters to be estimated.

Let  $\xi$  denote the vector of stacked parameters comprised of the conditional mean parameters in  $B$ ,  $\tilde{S}$ ,  $\tilde{\mu}$ , and  $\tilde{\rho}$ , along with the conditional variance parameters in  $\Gamma$ ,  $\Upsilon$ , and  $\varpi_h$ . Assuming that the reduced form shocks  $u_{t+1}$  are jointly normally distributed, the logarithm of the density for  $X_{t+1}$  conditional on  $X_t$  and  $\Omega_{t+1}$ , or equivalently the contribution to the log-likelihood function coming from  $X_{t+1}$ , may be expressed as,

$$\begin{aligned} L_t(X_{t+1}, \xi) &= -2 \log 2\pi - \frac{1}{2} \log |\Omega_t| - \frac{1}{2} (X_{t+1} - B^{-1}\tilde{\mu} \\ &\quad - \Phi X_t)' \Omega_t^{-1} (X_{t+1} - B^{-1}\tilde{\mu} - \Phi X_t) \\ &= -2 \log 2\pi - \frac{1}{2} \log |\Sigma_t| + \log |\tilde{S}^{-1}B| \\ &\quad - \frac{1}{2} (X_{t+1} - B^{-1}\tilde{\mu} - B^{-1}\tilde{\rho}BX_t)' \tilde{S}^{-1} \\ &\quad \times B\Sigma_t^{-1}B'\tilde{S}^{-1}' (X_{t+1} - B^{-1}\tilde{\mu} - B^{-1}\tilde{\rho}BX_t). \end{aligned} \quad (23)$$

Even if the assumption of conditional normality is violated empirically, the estimate for  $\xi$  obtained by maximizing the resulting log-likelihood function, defined by summing (23) over the full sample, remains consistent and asymptotically normally distributed under quite general conditions; see, e.g., Bollerslev and Wooldridge (1992).

The long-run implications from multivariate GARCH models can be very sensitive to estimation errors and small perturbations in a few parameters. To help guard against this, we augment the Gaussian-based score for the “structural” VAR-GARCH model with an additional set of moment conditions designed to ensure that the unconditional variances of the reduced form errors implied by the model match their standard VAR-based analogs.<sup>20</sup> Expressing this additional set of moments in parallel to Eq. (23) and the contribution to the likelihood function coming from  $X_{t+1}$ , we have

$$\begin{aligned} W_t(X_{t+1}, \xi) &= \varpi_u - \text{diag} \left( (X_{t+1} - \mu^{\text{OLS}} - \Phi^{\text{OLS}}X_t) \right. \\ &\quad \left. \times (X_{t+1} - \mu^{\text{OLS}} - \Phi^{\text{OLS}}X_t)' \right), \end{aligned} \quad (24)$$

where the “OLS” superscript indicates the parameters obtained from equation-by-equation least squares estimation of the reduced form VAR. The estimates for  $\xi$  reported below are obtained by applying standard iterated GMM to the conditional set of moments defined by the score for the conditional density in (23), say  $\partial_\xi L_t(X_{t+1}, \xi)$ , augmented with the moment conditions in (24),<sup>21</sup>

$$g(X_{t+1}, \xi) = \begin{pmatrix} \partial_\xi L_t(X_{t+1}, \xi) \\ W_t(X_{t+1}, \xi) \end{pmatrix}. \quad (25)$$

We turn next to a discussion of the resulting  $\hat{\xi}$ , and the implications of the estimates in regard to the dynamics of the systematic risk factors and the dependencies among the “structural” shocks.

<sup>19</sup> More formally,  $\Theta_1 = (\Phi_0^{-1} \otimes \Phi_0^{-1})I$ ,  $\Theta_2 = [\text{vec}(\Phi_0^{-1'}\Phi_0^{-1}), \text{vec}(\Phi_0^{-1'}\Phi_0^{-1}), \text{vec}(\Phi_0^{-1'}\Phi_0^{-1}), \text{vec}(\Phi_0^{-1'}\Phi_0^{-1})]'$ , where  $\Phi_0^{-1(i)}$  denotes the  $i$ th row of the square matrix  $\Phi_0^{-1}$ , and the  $16 \times 4$  matrix  $I$  helps to transform the vector  $\text{vec}(\Omega_t)$  into diagonal matrix form.

<sup>20</sup> This mirrors the variance targeting approach originally advocated by Engle and Mezrich (1996). However, in contrast to that two-step approach, the GMM-based procedure applied here jointly estimates all of the parameters in  $\xi$  in a single step.

<sup>21</sup> This idea of augmenting the likelihood function with additional information mirrors the use of quasi-Bayesian priors, applied in a different context by, e.g., Hamilton (1991), and may also be seen as a form of shrinkage type estimation.

3.3. Estimation results

The dynamic dependencies in the observable state vector  $X_t = [ERV_t, VRP_t, \Delta d_t, dp_t]'$  underlying our GMM estimation is directly related to the latent state vector  $f_t = [\sigma_t^2, q_t, \Delta d_t, x_t]'$  of interest by the affine equation  $X_t = \mu_x + Qf_t$ . This allows us to infer both the contemporaneous interaction matrix  $Q$  and the autoregressive matrix  $\rho$  describing the mean dynamics in  $f_{t+1} = \mu + \rho f_t + S\epsilon_{t+1}$  from the estimates for  $B$  and  $\tilde{\rho}$  based on  $BX_{t+1} = \tilde{\mu} + \tilde{\rho}BX_t + \tilde{S}\tilde{\epsilon}_{t+1}$ , and the relations in Eq. (18) above. Similarly, the estimated volatility loading matrix  $\tilde{S}$  for the observable state vector  $X_t$  allow us to infer the volatility loading matrix  $S$  for the latent state vector  $f_t$  from Eq. (19), while the estimated volatility dynamics of the  $\tilde{\epsilon}_{t+1}$  shocks effectively determines the implied volatility dynamics of the “structural”  $\epsilon_{t+1}$  shocks.

We begin with a discussion of the estimates for  $B$  and  $\tilde{\rho}$ ,

$$\hat{B} = \begin{pmatrix} 1 & -0.02 & 0 & 0 \\ & (0.11) & & \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -0.60 & -1.44 & -0.19 & 1 \\ & (0.03) & (0.10) & \end{pmatrix} \tag{26}$$

$$\hat{\rho} = \begin{pmatrix} 0.64 & 0 & 0 & 0 \\ & (0.05) & & \\ 0 & 0.46 & 0 & 0 \\ & (0.07) & & \\ 0 & 0 & -0.23 & -0.002 \\ & & (0.03) & (0.004) \\ 0 & 0 & 0 & 0.988 \\ & & & (0.009) \end{pmatrix}$$

where the numbers in parentheses represent asymptotic standard errors. With the exception of  $B_{1,2}$  and  $\tilde{\rho}_{3,4}$ , all of the individual parameter estimates are highly statistically significant. All of the estimates also have the “correct” signs vis-a-vis the implications from the equilibrium-based model and the “structural” VAR.

In particular, the negative estimates for the loadings for the dividend price ratio reported in the last row of the  $B$  matrix are consistent with the idea that the two volatility components  $\sigma_t^2$  and  $q_t$ , and cash flow growth  $\Delta d_t$ , are all genuine risk factors with negative market prices of risks.<sup>22</sup> Within the context of the standard [Bansal and Yaron \(2004\)](#) long-run risk model, these negative contemporaneous relationships between the dividend-price ratio and the other state variables, or risk factors, are critically dependent on the risk aversion parameter  $\gamma > 1$  and the intertemporal elasticity of substitution  $\psi > 1$ . As such, our “structural” estimation results indirectly support this commonly invoked set of assumptions.

Our estimate for  $\tilde{\rho}_{4,4} = \rho_x = 0.988$  also points to a highly persistent and very accurately estimated long-run risk factor. This contrasts with the typical practice of simply fixing the long-run persistence coefficient at some “large” value, as in, e.g., [Bansal et al. \(2007a\)](#), and clearly highlights the advantages of the more structured GMM estimation approach and richer data sources applied here. Meanwhile, even though our estimate for  $\phi_{dx} = \tilde{\rho}_{3,4} = \frac{\phi_{dx}}{-A_{d,x}} = -0.002$  is “correctly” signed, the parameter is not significantly different from zero, and as such offers only limited support to the idea that the long-run risk factor  $x_t$  contemporaneously impacts cash flows  $\Delta d_t$ .

Interestingly, our use of more accurate volatility measures results in a much more persistent consumption variance estimate  $\tilde{\rho}_{1,1} = \rho_\sigma = 0.64$  compared to the estimates recently reported in [Campbell et al. \(2013\)](#). Moreover, our estimates for  $\tilde{\rho}_{1,1} = \rho_\sigma =$

$0.64 > \tilde{\rho}_{2,2} = \rho_q = 0.46$  imply that the consumption variance  $\sigma_t^2$  is more persistent than the variance-of-variance  $q_t$ , or economic uncertainty, which is directly in line with the implicit assumptions invoked in the calibrations reported in [Bollerslev et al. \(2009\)](#).

Turning to our estimates for the volatility dependence matrix  $\tilde{S}$ ,

$$\hat{\tilde{S}} = \begin{pmatrix} 1 & 0 & 0 & 0.08 \\ & & & (0.04) \\ -0.29 & 1 & 0 & -0.09 \\ & (0.06) & & (0.02) \\ -0.36 & -0.09 & 1 & 0.15 \\ & (0.05) & (0.08) & (0.03) \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{27}$$

all of the individual parameters, except  $\tilde{S}_{3,2}$ , are again highly statistically significant. This clearly underscores the idea that multiple volatility factors are indeed needed to accurately describe the dynamic dependencies observed in the data, and that the standard long-run risk model with a single stochastic volatility factor is misspecified. To more fully appreciate this and the other implications of the estimates recall again the relationship between  $\tilde{S}$  and the “structural”  $S$  matrix for the latent state vector in Eq. (19).

It follows from this relation that shocks to cash flow growth  $\Delta d_t$  are adversely affected by shocks to the long-run risk component  $x_t$ , as  $s_{d,x} \propto -\tilde{S}_{3,4} = -0.15$ .<sup>23</sup> This is consistent with the idea that companies tend to distribute more in dividends when long-run growth opportunities are poor. The “structural” long-run risk shock affects the two variance processes  $\sigma_t^2$  and  $q_t$  in opposite directions. Good news about long-run consumption growth reduces the consumption variance, as  $s_{\sigma,x} \propto -\tilde{S}_{1,4} = -0.08 < 0$ , but increases economic uncertainty, as  $s_{q,x} \propto -\tilde{S}_{2,4} = 0.09 > 0$ . The first effect represents the well known “leverage effect”, whereby a negative growth shock is associated with higher volatility, and vice versa. The second effect, however, is more subtle. Since  $q_t$  directly affects the time-varying volatility of the long-run risk component, a positive  $s_{q,x}$  implies that when a positive  $x_{t}$  shock occurs, the volatility of next period’s  $\epsilon_{x,t+1}$  will also be higher, and vice versa. Intuitively, this could happen when good news in consumption growth is accompanied by better investment opportunities, in turn resulting in higher economic uncertainty, possibly due to over-investment. Interestingly, our estimates for  $\tilde{S}$  also suggest that  $s_{q,\sigma} \propto \tilde{S}_{2,1} = -0.29 < 0$ , implying that a positive “structural” shock to consumption volatility  $\sigma_t^2$  reduces the uncertainty of volatility  $q_t$ . This effect is naturally interpreted as a new “leverage effect” between volatility and volatility-of-volatility.<sup>24</sup>

Our identification and estimation of the “structural” model parameters rely crucially on the presence of time-varying conditional heteroscedasticity in the  $\epsilon_{t+1}$  shocks. The GMM parameter estimates for the “structural” factor GARCH model describing this heteroscedasticity are reported in [Table 2](#). As the table shows, all of the shocks do indeed exhibit highly significant (G)ARCH effects.<sup>25</sup> The overall good fit of the model is also supported by the conventional  $J$ -test statistic for general model misspecification and the minimized value of the GMM objective function equal to 12.76,

<sup>23</sup> We use the symbol  $\propto$  to denote proportional to.

<sup>24</sup> This new equilibrium-based “leverage effect” is also consistent with the asymmetries in daily and high-frequency intraday VIX and S&P 500 returns documented in [Aboura and Niklas \(2012\)](#) and [Bollerslev et al. \(2013\)](#), respectively.

<sup>25</sup> The significance of the (G)ARCH effects is also indirectly supported by Ljung-Box tests for residual serial correlation in the raw and standardized absolute residuals from the model; further details concerning these results are available upon request. This, of course, is directly in line with the burgeoning literature on the estimation of reduced form GARCH and stochastic volatility models for a wide array of other financial and macroeconomic time series.

<sup>22</sup> Note that the market price of dividend risk  $B_{4,3} = -0.19$  is imputed to by the constraint  $A_{d,d} = \frac{\rho_d}{1-\kappa_{d,1}\rho_d}$  imposed in Eq. (18).

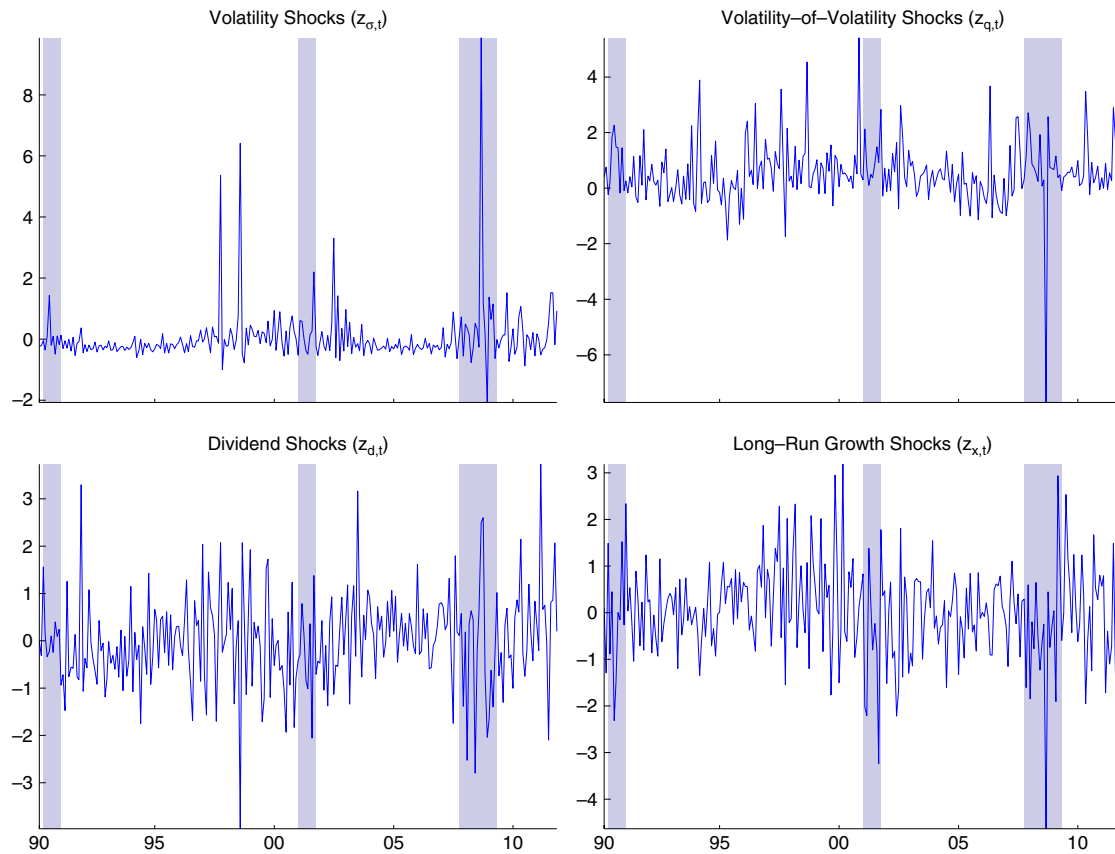


Fig. 2. Model implied structural shocks.

Table 2

“Structural” factor GARCH estimation. The table reports the GMM estimation result for the conditional variance parameters for the “structural” factor GARCH model discussed in the main text. The column labeled  $\varpi_u$  gives the unconditional variance of the reduced form shocks  $u_t$ .  $\Gamma$  and  $\Upsilon$  denote the ARCH and GARCH parameters, respectively, for the “structural” shocks  $\tilde{\epsilon}_t$ . The estimates are based on monthly data from February 1990 to November 2011, for a total of 262 observations.

$\tilde{\epsilon}_t$	$\varpi_u$		$\Gamma$		$\Upsilon$	
$ERV_t$	0.0011	(0.0002)	0.189		0.273	(0.075)
$VRP_t$	0.0003	(0.0000)	0.758	(0.080)	0.239	(0.077)
$\Delta d_t$	0.0006	(0.0001)	0		0.524	(0.100)
$d_{t+1}/p_t$	0.0016	(0.0002)	0.299		0.163	(0.082)

which corresponds to a p-value of 0.12 in the relevant asymptotic chi-square distribution.<sup>26</sup>

In order to further gauge the quality of the fit afforded by the model, Fig. 2 plots the time-series of “structural” shocks associated with each of the four equations. The top two panels show the volatility shocks  $z_{\sigma,t}$  and  $z_{q,t}$ . Both of these shocks experienced unprecedented large, albeit opposite signed, realizations during the 2007–2009 “Great Recession”. Interestingly, neither one of the earlier 1990–1991 and 2001–2002 NBER-dated recessions were accompanied by especially large “structural” volatility shocks. The general time-series pattern of the equilibrium-based cash flow shocks  $z_{\Delta d,t}$  appear quite similar to that of the normalized cash flow news in Campbell et al. (2013). Although not quite as dramatic as for the two volatility shocks, the permanent growth shocks  $z_{x,t}$

also experienced their most extreme realizations during the “Great Recession”. This basic dynamic pattern in the equilibrium-based growth shocks is again quite similar to that of the normalized discount rate news shocks reported in Campbell et al. (2013).<sup>27</sup>

In lieu of these findings and generally supportive diagnostic tests for the “structural” factor GARCH model, we turn next to our main empirical investigations, showing how incorporating the additional variance-related state variables in the equilibrium-based model help shed new light on the return and dividend growth predictability patterns inherent in the data.

#### 4. Model implied return and cash flow predictability

Our predictability analysis is based on recasting the “structural” factor GARCH model in the form of an expanded VAR system, along with the use of the standard Campbell–Shiller approximation for expressing the return as a function of the observable state variables.

##### 4.1. VAR and predictability

The first order VAR for the state vector  $X_t = [ERV_t, VRP_t, \Delta d_t, dp_t]$  implied by the “structural” factor GARCH model in Eq. (20) does not directly involve the return. However, by the standard Campbell–Shiller approximation, the return may be conveniently expressed as  $r_{t,t+1} = \kappa_{d,0} - \kappa_{d,1}dp_{t+1} + dp_t + \Delta d_{t+1}$ .<sup>28</sup> Combining

<sup>26</sup> By contrast, the two alternative specifications discussed in the supplementary Appendix A, one closer to Drechsler and Yaron (2011) with  $G_t = \text{diag}[\sigma_t, \sqrt{q_t}, \sigma_t, \sqrt{q_t}, \sigma_t]$ , and one closer to Branger and Völkert (2012) with  $G_t = \text{diag}[\sigma_t, \sigma_t, \sigma_t, \sqrt{q_t}, \sigma_t]$ , result in GMM-based  $J$ -statistics equal to 26.31 and 37.02, respectively, with corresponding p-values essentially zero.

<sup>27</sup> This is also consistent with the findings in Lettau and Ludvigson (2013), who suggest that large negative permanent growth shocks might have adversely affected housing wealth.

<sup>28</sup> The accuracy of the Campbell–Shiller approximation has recently been corroborated by Engsted et al. (2012). By definition  $\kappa_{d,1} = \exp(-E(dp_t)) [1 +$



this equation for  $r_{t,t+1}$  with the VAR for  $X_{t+1}$ , it follows that

$$r_{t,t+1} = \mu_r + (l_1\Phi + e_4)X_t + l_1\Phi_0^{-1}\tilde{\epsilon}_{t+1}, \tag{28}$$

where  $\mu_r$  collects all of the relevant constant terms,  $l_1 \equiv (0, 0, 1, -k_{1,d})$ , and the selection vector  $e_4 \equiv (0, 0, 0, 1)$ . Iterating the VAR for  $X_t$  forward, it is therefore possible to derive closed-form expressions for the model-implied multi-period return  $r_{t,t+h} = r_{t,t+1} + \dots + r_{t+h-1,t+h}$  regressions based on any explanatory variable spanned by the  $X_t$  state vector.

In the analysis reported on below we will focus on the three key predictor variables: the log dividend-price ratio  $dp_t$ , the variance risk premium  $VRP_t$ , and the expected variation  $ERV_t$ . In particular, consider the regression of the  $h$ -period returns on the dividend-price ratio,

$$\frac{1}{h} \sum_{i=1}^h r_{t,t+i} = \alpha_{r,dp} + \beta_{r,dp}(h) \cdot dp_t + \varsigma_{t,t+h}. \tag{29}$$

By similar arguments to the ones in Hodrick (1992) and Campbell (2001), it is possible to show that

$$\beta_{r,dp}(h) = \frac{(l_1\Phi + e_4)(I - \Phi)^{-1}(I - \Phi^h)C(0)e'_4}{e_4C(0)e'_4} \tag{30}$$

where  $C(0) = \sum_{j=0}^{\infty} \Phi^j \Phi_0^{-1} \text{diag}(\vartheta_0^{-1} \varpi_{ii}) \Phi_0^{-1'} \Phi^j$  denotes the model-implied unconditional covariance matrix for  $X_t$ , and  $e_4 \equiv (0, 0, 0, 1)$ .<sup>29</sup> Similarly, the implied coefficients for the return predictability regressions based on  $VRP_t$  and  $ERV_t$  may be expressed in close form as,

$$\beta_{r,VRP}(h) = \frac{(l_1\Phi + e_4)(I - \Phi)^{-1}(I - \Phi^h)C(0)e'_2}{e_2C(0)e'_2} \tag{31}$$

$$\beta_{r,ERV}(h) = \frac{(l_1\Phi + e_4)(I - \Phi)^{-1}(I - \Phi^h)C(0)e'_1}{e_1C(0)e'_1} \tag{32}$$

where the  $e_1$  and  $e_2$  selection vectors are defined in an obvious manner.<sup>30</sup>

In parallel to Eq. (28) for the returns, the growth rate dynamics implied by the “structural” factor GARCH may be expressed in linear form as,

$$\Delta d_{t+1} = \mu_d + e_3\Phi X_t + e_3\Phi_0^{-1}\tilde{\epsilon}_{t+1}, \tag{33}$$

where  $\mu_d$  collects all the relevant constant terms. Thus, replacing  $l_1\Phi + e_4$  with  $e_3\Phi$  in the formulas for the regression coefficients above, comparable expressions for the cash flow predictability regression coefficients  $\beta_{\Delta d,dp}(h)$ ,  $\beta_{\Delta d,VRP}(h)$ , and  $\beta_{\Delta d,ERV}(h)$  are readily available. When interpreting these coefficients, it is important to keep in mind the relationship  $E_t(\Delta d_{t+1}) = \phi_{dx}X_t + \rho_d\Delta d_t$  implied by Eqs. (10) and (11), and the fact that within the “structural” model the expected value of next periods dividend growth rate is linearly related to the lagged dividend growth rate and the long-run risk component.

$\exp(-E(dp_t))^{-1}$ . In the estimation results reported on below we rely on the sample average of the monthly dividend-price ratio from January 1965 to November 2011 when calculating  $E(dp_t)$ , implying a value of  $\kappa_{d,1} = 0.9976$ .

<sup>29</sup> In the empirical results reported on below, we truncate the infinite sum in the expression for  $C(0)$  at 120, or ten years; see Bollerslev and Hodrick (1995) for further discussion along these lines.

<sup>30</sup> Analytical expressions for the  $R^2$ s from the regressions may be derived in a similar manner. Specifically, for the dividend-price ratio regression  $R^2_{r,dp}(h) = h^2\beta^2_{r,dp}(h)e_4C(0)e'_4/\text{Var}(\sum_{j=1}^h r_{t,t+j})$ , where  $\text{Var}(\sum_{j=1}^h r_{t,t+j}) = h(l_1\Phi + e_3)C(0)(l_1\Phi + e_3)' + hl_1(I - \Phi)C(0)l'_1 + \sum_{i=1}^{h-1} 2(h-i)((l_1\Phi + e_3)\Phi^i C(0)(l_1\Phi + e_3)' + (l_1\Phi + e_3)\Phi^{i-1}(I - \Phi)C(0)e'_3)$ .

#### 4.2. Model-implied reduced form VAR estimates

The reduced form VAR parameter matrix  $\Phi$  and the unconditional covariance matrix  $C(0)$  for  $X_t$  entering the expressions for the predictive regression coefficients in Eqs. (30)–(32) could, of course, be estimated directly by OLS equation-by-equation. However, that obviously would ignore any of the equilibrium-based “structural” restrictions. It also would not permit the separate identification of the contemporaneous  $\Phi_0$  matrix entering the expressions for the return and dividend growth rate in Eqs. (28) and (33), respectively.

Instead, the  $\Phi_0$  and  $\Phi$  parameter matrices may both be deduced from the “structural” factor GARCH model parameters and the relations  $\Phi = B^{-1}\tilde{\rho}B$  and  $\Phi_0^{-1} = B^{-1}\tilde{S}$  derived above. Substituting the previously discussed estimates for  $B$ ,  $\tilde{\rho}$  and  $\tilde{S}$  into these expressions, yields,

$$\hat{\Phi} = \begin{pmatrix} 0.64 & -0.003 & 0 & 0 \\ (0.05) & (0.020) & & \\ 0 & 0.46 & 0 & 0 \\ & (0.07) & & \\ 0.001 & 0.002 & -0.23 & -0.002 \\ (0.002) & (0.005) & (0.03) & (0.004) \\ -0.21 & -0.76 & -0.23 & 0.988 \\ (0.03) & (0.13) & (0.03) & (0.087) \end{pmatrix} \tag{34}$$

$$\hat{\Phi}_0 = \begin{pmatrix} 0.995 & 0.02 & 0 & 0.08 \\ (0.033) & (0.11) & & (0.04) \\ -0.29 & 1 & 0 & -0.09 \\ (0.06) & & & (0.02) \\ -0.34 & -0.09 & 1 & 0.15 \\ (0.06) & (0.08) & & (0.03) \\ 0.11 & 1.44 & 0.19 & 0.94 \\ (0.09) & (0.13) & (0.02) & (0.03) \end{pmatrix}$$

where the numbers in parentheses represent standard errors derived by the delta-method.

Based on these estimates for  $\Phi$  and  $\Phi_0$ , the return equation in (28) may be expressed numerically as,

$$r_{t,t+1} = 0.05 + 0.20 ERV_t + 0.76 VRP_t - 0.0013 \Delta d_t + 0.013 dp_t \\ (0.03) \quad (0.03) \quad (0.12) \quad (0.002) \quad (0.011) \\ - 0.47 \tilde{\epsilon}_{\sigma,t+1} - 1.52 \tilde{\epsilon}_{q,t+1} + 0.81 \tilde{\epsilon}_{\Delta d,t+1} - 0.79 \tilde{\epsilon}_{x,t+1}. \tag{35}$$

Of course, this “estimated” return equation does not actually rely on the return data, but instead is deduced from our estimates for the equilibrium-based model and the observable state vector involving the dividend growth rate and the log dividend-price ratio. Again, this mirrors the approach of Cochrane (2008). However, in contrast to the return equation therein, which only involves the dividend-price ratio, we purposely include the two variance variables, both of which enters with highly significant coefficients.

Further underscoring the importance of incorporating the variation measures into the analysis, the model-implied loadings for all of the “structural” shocks are also highly significant. Among the four shocks, the ones for the long-run risk component and the consumption variance uncertainty have the largest impacts, accounting for 43% ( $z_{x,t}$ ) and 26% ( $z_{q,t}$ ) of the unexpected unconditional return variation, respectively. The “estimated” return equation in (35) also implies that the total one-month explainable return variation equals 9%, far exceeding that afforded by traditional univariate return predictability regressions that does not include  $ERV_t$  and  $VRP_t$ .

Explicitly writing out the second equation for the variance risk premium in the model-implied VAR,

$$VRP_{t+1} = 0.001 + 0.46 VRP_t - 0.29 \tilde{\epsilon}_{\sigma,t+1} \\ (0.001) \quad (0.07) \quad (0.06) \\ + \tilde{\epsilon}_{q,t+1} - 0.09 \tilde{\epsilon}_{x,t+1} \tag{36}$$

**Table 3**

Predictive Regressions based on the Dividend-Price Ratio. The table reports the slope coefficients in the return and cash flow predictability regressions,

$$\frac{1}{h} \sum_{i=1}^h r_{t,t+i} = \alpha_{r,dp} + \beta_{r,dp}(h) \cdot dp_t + \zeta_{t,t+h}$$

$$\frac{1}{h} \sum_{i=1}^h \Delta d_{t,t+i} = \alpha_{\Delta d,dp} + \beta_{\Delta d,dp}(h) \cdot dp_t + \zeta_{t,t+h}$$

implied by the parameter estimates for the “structural” factor GARCH model discussed in the main text, with asymptotic standard errors in parentheses. The table also reports the slope coefficients implied by a two-variable reduced form homoskedastic VAR for the dividend growth rate and the dividend-price ratio, as in Cochrane (2008), along with the results from simple univariate predictive regressions. The time horizon  $h$  runs from one to ten years in the first two panels, and from one to twelve months in the bottom three panels. All of the results are based on monthly data from February 1990 to November 2011.

Years	1	2	3	4	5	6	7	8	9	10
Structural model implied										
$\beta_{r,dp}(h)$	0.0126 (0.0099)	0.0114 (0.0085)	0.0106 (0.0074)	0.0098 (0.0065)	0.0092 (0.0057)	0.0086 (0.0050)	0.0081 (0.0045)	0.0077 (0.0039)	0.0072 (0.0035)	0.0068 (0.0031)
$\beta_{\Delta d,dp}(h)$	-0.0013 (0.0028)	-0.0012 (0.0027)	-0.0011 (0.0026)	-0.0011 (0.0024)	-0.0010 (0.0023)	-0.0009 (0.0022)	-0.0009 (0.0021)	-0.0008 (0.0020)	-0.0008 (0.0019)	-0.0007 (0.0018)
Months	1	2	3	4	5	6	9	12		
Structural model implied										
$\beta_{r,dp}(h)$	0.0166 (0.0135)	0.0157 (0.0127)	0.0150 (0.0122)	0.0145 (0.0117)	0.0141 (0.0114)	0.0138 (0.0111)	0.0131 (0.0104)	0.0126 (0.0099)		
$\beta_{\Delta d,dp}(h)$	-0.0022 (0.0023)	-0.0017 (0.0027)	-0.0016 (0.0028)	-0.0015 (0.0029)	-0.0015 (0.0029)	-0.0014 (0.0029)	-0.0014 (0.0029)	-0.0013 (0.0028)		
Univariate regression										
$\beta_{r,dp}(h)$	0.0112 (0.0089)	0.0119 (0.0083)	0.0121 (0.0078)	0.0123 (0.0076)	0.0129 (0.0074)	0.0135 (0.0072)	0.0153 (0.0069)	0.0161 (0.0069)		
$\beta_{\Delta d,dp}(h)$	-0.0029 (0.0030)	-0.0011 (0.0015)	-0.0004 (0.0012)	-0.0001 (0.0008)	-0.0000 (0.0007)	0.0001 (0.0006)	0.0002 (0.0005)	0.0001 (0.0004)		

shows that the only “structural” shock that enters the return and VRP equations with the opposite sign is  $\tilde{\epsilon}_{q,t}$ . Indeed, excluding the impact of the economic uncertainty shock from both equations changes the monthly conditional correlation, or “leverage effect”, from a negative -0.09 to a positive 0.66, again reinforcing the importance of jointly modeling all of the elements in the  $X_t$  state vector.

### 4.3. Model-implied predictability relations

The VAR-based formula for the slope coefficients presented above allows for a direct assessment of the statistical significance of the different predictor variables across different forecast horizons. The formula also allows us to directly assess the enhanced efficiency afforded by the “structural” factor GARCH model compared to the reduced form VAR and simple univariate regression procedures traditionally used in the literature.

To begin, the top panel in Table 3 reports the implied slope coefficients for forecasting returns and cash flows by the dividend-price ratio  $dp_t$  over long 1- to 10-year horizons, as previously analyzed in the literature. Although the patterns in the estimated coefficients are generally in line with the estimates reported in the existing literature based on longer calendar time spans of data, taken as a whole there is little evidence for any predictability over these long multi-year horizons in the data analyzed here.<sup>31</sup> The results for the shorter within year “structural” and simply unconstrained univariate regressions reported in the lower panel of the table tell a similar story.

The lack of predictability for the long multi-year horizons, is, of course, not too surprising. With only slightly more than twenty years worth of monthly observations any suggestions about statistically significant long-run predictability should be taken with a grain of salt. For the remainder of this section, we will consequently restrict our attention to within-year horizons only.<sup>32</sup>

Turning to our key empirical findings pertaining to the “new” variance related forecasting variables, Fig. 3 shows the regression slope coefficients for the variance risk premium  $VRP_t$  implied by the “structural” factor GARCH model (indicated by dots) along with the corresponding 95% confidence intervals (indicated by the shaded area). For comparison purposes, we also include the estimated slope coefficients from simple univariate predictive regressions based on the variance risk premium (indicated by the stars) along with their 95% confidence intervals (indicated by the dashed lines). Focusing on the top panel for the returns, both procedures result in significant estimates for up to eight months. It is noteworthy that even though the model-implied point estimates are systematically lower than the unrestricted OLS estimates, they are also less erratic, and the confidence intervals much smaller. Indeed, looking at the numbers in Table 4, the  $t$ -statistics for testing the null hypothesis of no return predictability are uniformly larger for the “structural” approach.

This discrepancy in the results across the two approaches is even stronger for the cash flow predictability regressions reported in the bottom panel in Fig. 3. Whereas the estimated slope coefficients from the univariate regressions are all insignificant, the  $t$ -statistics associated with the VAR-based model-implied coefficients are all negative and exceed conventional significance levels for up to six months. Hence, not only are higher variance

<sup>31</sup> We also experimented with a traditional two-variable homoskedastic VAR for the dividend-price ratio and the dividend growth rate, as in Cochrane (2008), resulting in similar coefficient estimates, but typically larger standard errors, thus highlighting the more accurate inference afforded by explicitly incorporating the equilibrium-based restrictions and the strong heteroskedasticity inherent in the data. Further details concerning these results are available upon request.

<sup>32</sup> The univariate return regressions reported in Bollerslev et al. (2009) and Drechsler and Yaron (2011) that in part motivate our analysis also suggest that the return predictability inherent in the variance risk premium is confined to relatively short horizons.

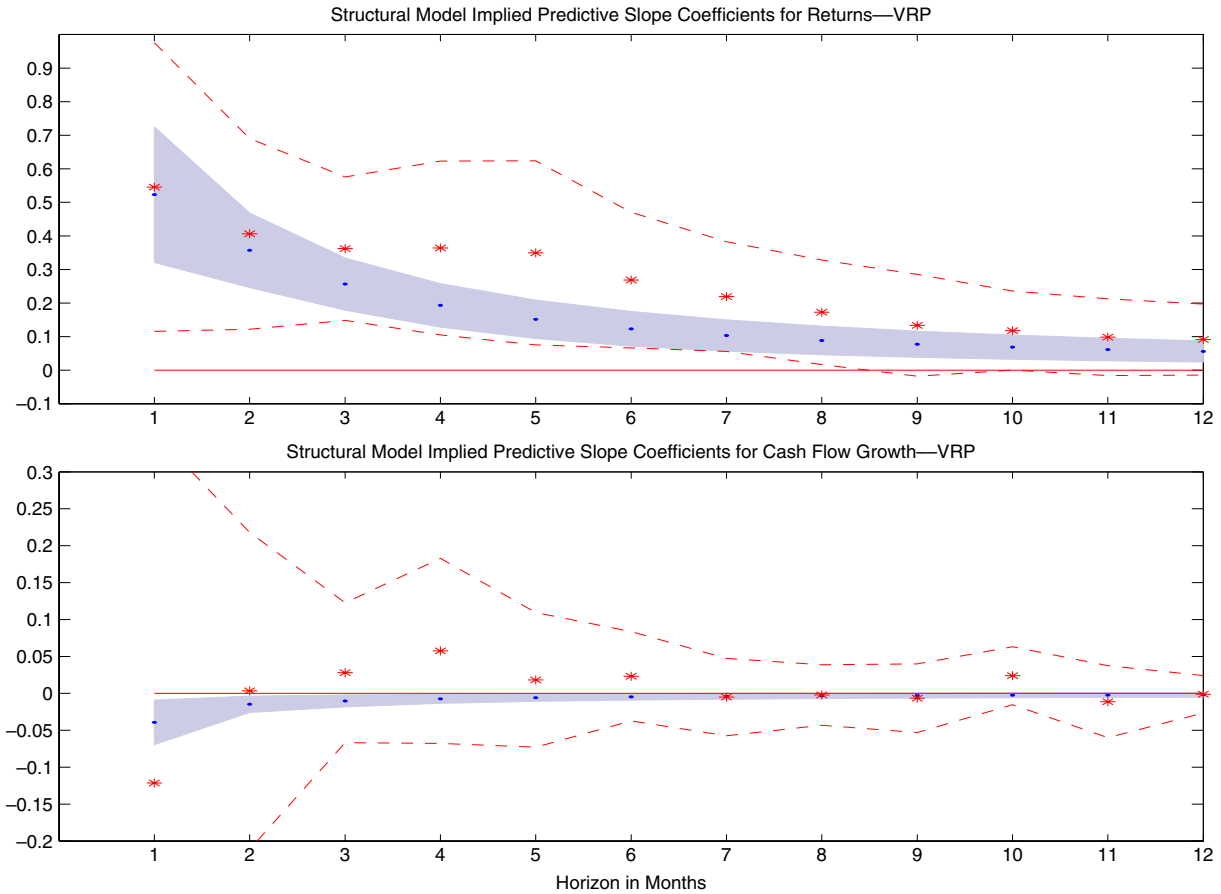


Fig. 3. Predictive regressions based on the variance risk premium.

Table 4

Predictive Regressions based on the Variance Risk Premium. The table reports the slope coefficients in the return and cash flow predictability regressions,

$$\frac{1}{h} \sum_{i=1}^h r_{t,t+i} = \alpha_{r,VRP} + \beta_{r,VRP}(h) \cdot VRP_t + \zeta_{t,t+h}$$

$$\frac{1}{h} \sum_{i=1}^h \Delta d_{t,t+i} = \alpha_{\Delta d,VRP} + \beta_{\Delta d,VRP}(h) \cdot VRP_t + \zeta_{t,t+h}$$

implied by the parameter estimates for the “structural” factor GARCH model discussed in the main text, with asymptotic standard errors in parentheses. The table also reports the slope coefficients from simple univariate predictive regressions. The time horizon  $h$  runs from one to twelve months. All of the results are based on monthly data from February 1990 to November 2011.

Months	1	2	3	4	5	6	9	12
Structural model implied								
$\beta_{r,VRP}(h)$	0.5228 (0.1031)	0.3571 (0.0566)	0.2564 (0.0396)	0.1929 (0.0330)	0.1514 (0.0292)	0.1231 (0.0263)	0.0772 (0.0200)	0.0557 (0.0161)
$\beta_{\Delta d,VRP}(h)$	-0.0393 (0.0154)	-0.0147 (0.0057)	-0.0103 (0.0041)	-0.0074 (0.0031)	-0.0058 (0.0026)	-0.0047 (0.0023)	-0.0029 (0.0018)	-0.0020 (0.0016)
Univariate regression								
$\beta_{r,VRP}(h)$	0.5454 (0.2194)	0.4060 (0.1450)	0.3620 (0.1090)	0.3640 (0.1321)	0.3494 (0.1400)	0.2683 (0.1031)	0.1335 (0.0773)	0.0911 (0.0540)
$\beta_{\Delta d,VRP}(h)$	-0.1215 (0.2420)	0.0035 (0.1094)	0.0280 (0.0483)	0.0576 (0.0640)	0.0183 (0.0465)	0.0232 (0.0309)	-0.0066 (0.0237)	-0.0012 (0.0130)

risk premia positively related to future returns, as previously documented in the literature, they also predict lower near-term future cash flows.<sup>33</sup> This, of course, contrasts with the view

commonly expressed in the literature that dividend growth rates are largely unpredictable over short within-year horizons.

Of course, the much-studied classical risk-return trade-off is not based on the variance risk premium, but rather the return variation itself. In spite of the intuitively appealing idea behind such a relationship, empirical attempts at establishing a significant risk-return tradeoff have largely proven futile; see, e.g., the discussion in Bollerslev and Zhou (2006) and Guo and Whitelaw

<sup>33</sup> This is also related to the observation by Bloom (2009) that an increase in economic uncertainty causes firms to temporarily reduce their investment and hiring, in turn resulting in a short-term productivity drop.

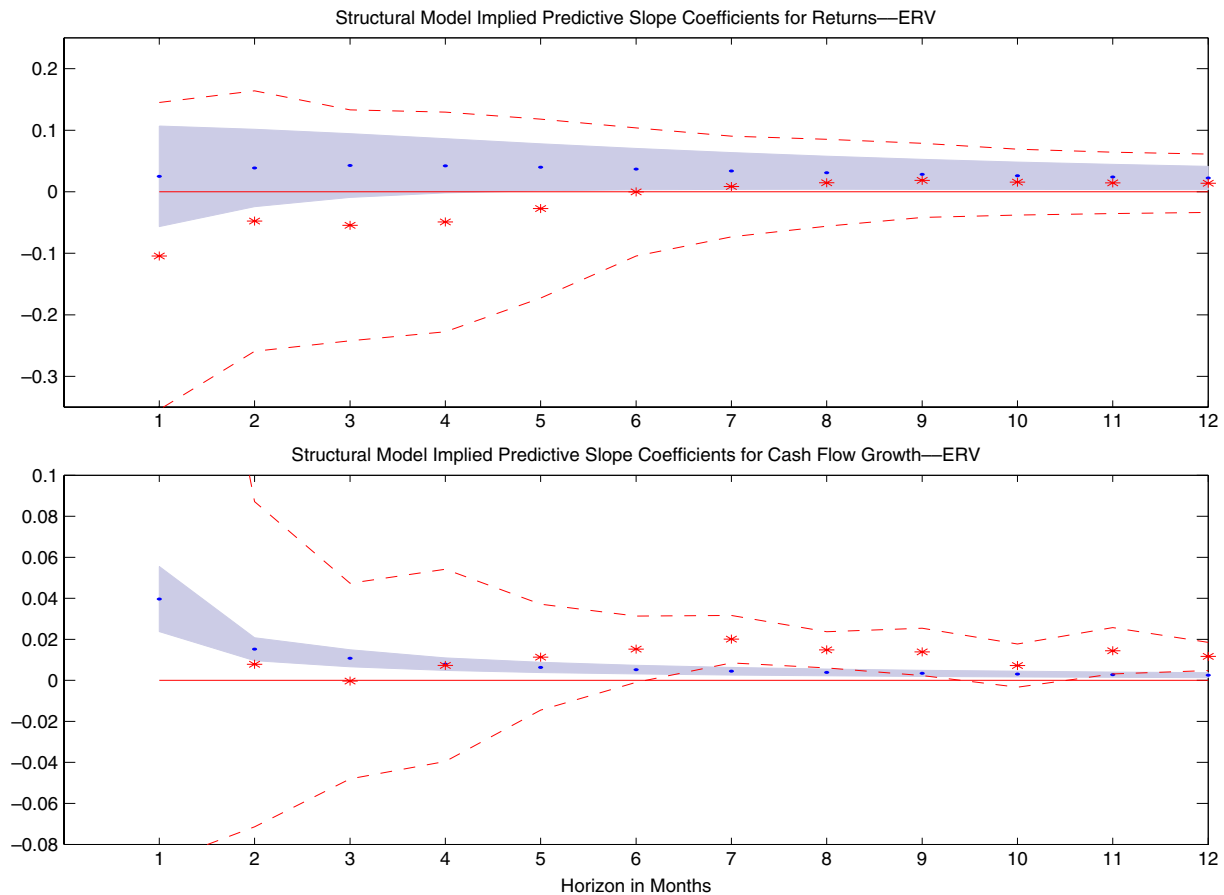


Fig. 4. Predictive regressions based on the expected variation.

(2006), and the many other references therein. The result for the univariate return regressions based on  $ERV_t$  reported in the top panel in Fig. 4 and Table 5 underscore the elusive nature of a simple linear relationship between the expected returns and the expected variation in the data analyzed here. None of the regression coefficients are significant, and most have the “wrong” sign. By contrast, the VAR-based estimates implied by the “structural” model are all positive and marginally significant for return horizons in excess of 4 months.<sup>34</sup>

The difference in the quality of the inference afforded by standard univariate regression-based procedures traditionally employed in the literature and the “structural” approach advocated here is even more dramatic for the cash flow predictions reported in the bottom panel in Fig. 4. While the simple univariate regressions suggest that the 1–6 months dividend growth rate is unpredictable, the regression coefficients implied by the “structural” model are all highly significant. Interestingly, whereas an increase in  $VRP_t$  predicts lower future cash flows, and increase in  $ERV_t$  is associated with significantly higher future cash flows. Again, this strong empirical evidence for short-run within-year cash flow predictability stands in sharp contrast to the results reported in the existing literature based on other more traditional predictor variables and valuation ratios.

<sup>34</sup> The use of  $IV_t = VRP_t + ERV_t$  results in qualitatively similar patterns, but slightly more significant coefficient estimates, compared to the ones reported for  $ERV_t$ , thus confirming earlier empirical findings in Bollerslev and Zhou (2006) and Guo and Whitelaw (2006) of a stronger risk-return trade-off when using implied as opposed to realized variation. Still, none of the univariate return regressions based on  $IV_t$  result in any significant predictability. Further details of these results are available upon request.

At a more general level, the results for the two different approaches reported in Tables 3–5 and Figs. 3–4 may also be seen as providing indirect support for the equilibrium-based “structural” model, in that the more accurate model-implied predictive relations systematically fall within the wider standard error bands associated with the unrestricted regressions. This, of course, would not necessarily be the case if the assumptions underlying the “structural” model were violated.

#### 4.4. Further discussion and interpretation

The contrast between the long-run predictability inherent in the dividend-price ratio, and the variance variables ability to predict both return and cash flow over shorter within-year horizons is intimately related to our equilibrium-based long-run risk model, and the way in which the fundamental risk factors affect the state variables.

In particular, while the dividend-price ratio  $dp_t$  loads on the long-run risk factor  $x_t$  and both of the volatility factors  $\sigma_t^2$  and  $q_t$ , the expected variation  $ERV_t$  depends only on the two volatility factors  $\sigma_t^2$  and  $q_t$ , and the variance risk premium  $VRP_t$  is exclusively determined by the volatility-of-volatility factor  $q_t$ . Consistent with earlier less formal model calibrations reported in the literature, our GMM-based estimates imply that the long-run risk factor is highly persistent with AR(1) coefficient equal to  $\rho_x = 0.988$ , while the consumption volatility factor is moderately persistent with AR(1) coefficient equal to  $\rho_\sigma = 0.64$ , and the consumption volatility-of-volatility factor is quickly mean-reverting with AR(1) coefficient equal to  $\rho_q = 0.46$ .

In light of these estimates for the underlying systematic risk factors, it is therefore not surprising that the “structural” model

**Table 5**

Predictive regressions based on the expected variation. The table reports the slope coefficients in the return and cash flow predictability regressions,

$$\frac{1}{h} \sum_{i=1}^h r_{t,t+i} = \alpha_{r,ERV} + \beta_{r,ERV}(h) \cdot ERV_t + \zeta_{t,t+h}$$

$$\frac{1}{h} \sum_{i=1}^h \Delta d_{t,t+i} = \alpha_{\Delta d,ERV} + \beta_{\Delta d,ERV}(h) \cdot ERV_t + \zeta_{t,t+h}$$

implied by the parameter estimates for the “structural” factor GARCH model discussed in the main text, with asymptotic standard errors in parentheses. The table also reports the slope coefficients from simple univariate predictive regressions. The time horizon  $h$  runs from one to twelve months. All of the results are based on monthly data from February 1990 to November 2011.

Months	1	2	3	4	5	6	9	12
Structural model implied								
$\beta_{r,ERV}(h)$	0.0251 (0.0417)	0.0387 (0.0321)	0.0426 (0.0264)	0.0422 (0.0225)	0.0398 (0.0195)	0.0368 (0.0171)	0.0283 (0.0124)	0.0223 (0.0096)
$\beta_{\Delta d,ERV}(h)$	0.0396 (0.0081)	0.0152 (0.0029)	0.0107 (0.0021)	0.0079 (0.0016)	0.0063 (0.0013)	0.0052 (0.0011)	0.0034 (0.0008)	0.0025 (0.0006)
Univariate regression								
$\beta_{r,ERV}(h)$	-0.1044 (0.1273)	-0.0477 (0.1080)	-0.0546 (0.0957)	-0.0491 (0.0911)	-0.0274 (0.0742)	-0.0003 (0.0531)	0.0185 (0.0307)	0.0137 (0.0242)
$\beta_{\Delta d,ERV}(h)$	0.1173 (0.1051)	0.0079 (0.0404)	-0.0004 (0.0243)	0.0073 (0.0239)	0.0113 (0.0132)	0.0152 (0.0082)	0.0139 (0.0059)	0.0117 (0.0035)

implied return predictability regressions based on  $VRP_t$ , which depends solely on  $q_t$ , result in the most significant coefficients over relatively short 1–6 months horizon. Meanwhile, the regressions based on  $dp_t$ , which loads heavily on  $x_t$ , should show the greatest explanatory power over longer multi-year horizons, which, of course is difficult to detect statistically with the limited time span of data analyzed here. Also, whereas the variance risk premium is most significant over horizons less than 6 months, the expected variation  $ERV_t$  displays the most significant predictability over 6–12 months horizons, as the more persistent  $\sigma_t^2$  process “shifts” the predictable forward.

The documented differences in the degree of cash flow predictability are most easily understood in terms of the correlations among the “structural” shocks. From the model estimates the cash flow shock is more strongly negatively correlated with the contemporaneous variance shock ( $s_{d,\sigma} \propto \tilde{S}_{3,1} = -0.36$ ), than it is with the uncertainty shock ( $s_{d,q} \propto \tilde{S}_{3,2} = -0.09$ ) or the long-run risk shock ( $s_{d,x} \propto -\tilde{S}_{3,4} = -0.15$ ). Since the expected variation loads more heavily on  $\sigma_t^2$  than  $q_t$ , while the dividend-price ratio and the variance risk premium are mostly determined by  $x_t$  and  $q_t$ , respectively,  $ERV_t$  will be more strongly negatively related to  $\Delta d_t$  than either  $dp_t$  or  $VRP_t$ . Because of the negative autocorrelation in  $\Delta d_t$  ( $\rho_d = -0.23 < 0$ ), this in turn translates into the strongest positive short-run cash flow predictability results for the  $ERV_t$  predictor variable implied by the “structural” VAR.

## 5. Conclusion

We examine the joint predictability of return and dividend growth rates within a present value framework, explicitly imposing the economic equilibrium-based constraints from a long-run risk model with time-varying consumption volatility and volatility-of-volatility risk. The model clearly differentiate the long-run predictability channels associated with the dividend-price ratio from the economic mechanisms responsible for the short-run predictability inherent in the variance risk premium and the expected return variation.

Consistent with [Bansal and Yaron \(2004\)](#), our GMM-based estimates of the “structural” factor GARCH model point to a highly persistent latent long-run risk factor. Our estimates also corroborate the calibrations in [Bollerslev et al. \(2009\)](#), and the notion that consumption volatility is more persistent than consumption

volatility-of-volatility. In addition, the “structural” shocks identified within the model reveal that cash flow respond negatively to contemporaneous long-run growth shocks, while consumption volatility decreases with shocks to the long-run growth factor, and volatility uncertainty increases with long-run growth shocks. A new “leverage effect” whereby shocks to consumption volatility is negatively related to volatility-of-volatility also emerges from our “structural” estimation.

By allowing for much sharper and accurate inference than the procedures traditionally employed in the literature, the VAR implied by the “structural” model also provides striking new evidence on the return and cash flow predictability inherent in the data. Specifically, we find that the variance risk premium, and to a lesser extend the expected return variation, significantly predicts short-run within-year returns. On the other hand, the expected return variation, and to a lesser extend the variance risk premium, strongly predicts short-run within-year dividend growth rates. This latter finding stands in sharp contrast to the view expressed by a number of studies in the literature that cash flows are largely unpredictable.

## Appendix A. Supplementary data

Supplementary material related to this article can be found online at <http://dx.doi.org/10.1016/j.jeconom.2015.02.031>.

## References

- Aboura, Sofiane, Niklas, Wagner, 2012, Extreme Asymmetric Volatility, Leverage, Feedback and Asset Prices, Working Paper, University of Paris Dauphine and Passau University.
- Andersen, Torben G., Tim, Bollerslev, Diebold, Francis X., 2007. *Roughing it up: including jump components in the measurement, modeling, and forecasting of return volatility*. *Rev. Econ. Statist.* 89, 701–720.
- Bagwell, Laurie S., Shoven, John B., 1989. *Cash distributions to shareholders*. *J. Econ. Perspect.* 3, 129–140.
- Baker, Malcolm, Wurgler, Jefferey, 2000. *The equity share in new issues and aggregate stock returns*. *J. Financ. Econometrics* 55, 2219–2257.
- Bansal, Ravi, Gallant, A. Ronald, Tauchen, George, 2007a. *Rational pessimism, rational exuberance, and asset pricing models*. *Rev. Econom. Stud.* 74, 1005–1033.
- Bansal, Ravi, Kiku, Dana, Shaliastovich, Ivan, Yaron, Amir, 2014. *Volatility, the macroeconomy, and asset prices*. *J. Financ.* 2471–2511.
- Bansal, Ravi, Kiku, Dana, Yaron, Amir, 2007b. *A Note on the Economics and Statistics of Predictability: A Long Run Risks Perspective*, Working Paper, Duke University and University of Pennsylvania.
- Bansal, Ravi, Yaron, Amir, 2004. *Risks for the long run: A potential resolution of asset pricing puzzles*. *J. Financ.* 59, 1481–1509.

- Bekaert, Geert, Engstrom, Eric, Xing, Yuhang, 2009. Risk, uncertainty and asset prices. *J. Financ. Econ.* 91, 59–82.
- Binsbergen, Jules H. Van, Brandt, Michael W., Koijen, Ralph S.J., 2012. On the timing and pricing of dividends. *Amer. Econ. Rev.* 102, 1596–1618.
- Binsbergen, Jules H. Van, Koijen, Ralph S.J., 2010. Predictive regressions: A present-value approach. *J. Financ.* 65, 1439–1471.
- Bloom, Nicholas, 2009. The impact of uncertainty shocks. *Econometrica* 77, 623–685.
- Bollerslev, Tim, Hodrick, Robert J., 1995. Financial Market Efficiency Tests. In: Pesaran, Hashem, Wickens, Michael (Eds.), *Handbook of Applied Economics*, chap.9. Basil Blackwell, London, pp. 415–458. (Chapter 9).
- Bollerslev, Tim, Osterrieder, Daniela, Sizova, Natalia, Tauchen, George, 2013. Risk and return: Long-run relationships, fractional cointegration, and return predictability. *J. Financ. Econ.* 108, 409–424.
- Bollerslev, Tim, Sizova, Natalia, Tauchen, George, 2012. Volatility in equilibrium: Asymmetries and dynamic dependencies. *Rev. Financ.* 16, 31–80.
- Bollerslev, Tim, Tauchen, George, Zhou, Hao, 2009. Expected stock returns and variance risk premia. *Rev. Financ. Stud.* 16, 31–80.
- Bollerslev, Tim, Wooldridge, Jeffrey M., 1992. Quasi-maximum likelihood estimation and inference in dynamic models with time varying covariances. *Econometric Rev.* 11, 143–172.
- Bollerslev, Tim, Zhou, Hao, 2006. Volatility puzzles: A simple framework for gauging return-volatility regressions. *J. Econometrics* 131, 123–150.
- Boudoukh, Jacob, Richardson, Matthew, Roberts, Michael, 2007. On the importance of measuring payout yield: implications for empirical asset pricing. *Journal of Finance* 62, 877–915.
- Branger, Nicole, Rodriguez, Paulo, Schlag, Christian, 2011. The Role of Volatility Shocks and Rare Events in Long-Run Risk Models, Working Paper, University of Muenster and Goethe University Frankfurt.
- Branger, Nicole, Völkert, Clemens, 2012. What is the Equilibrium Price of Variance Risk? A Long-Run Risk Model with Two Volatility Factors, Working Paper, University of Muenster.
- Campbell, John Y., 1987. Stock returns and the term structure. *J. Financ. Econ.* 18, 373–399.
- Campbell, John Y., 1993. Intertemporal asset pricing without consumption data. *Amer. Econ. Rev.* 83, 487–515.
- Campbell, John Y., 1996. Understanding risk and return. *J. Polit. Econ.* 104, 298–345.
- Campbell, John Y., 2001. Why long horizons? A study of power against persistent alternatives. *J. Empir. Financ.* 8, 459–491.
- Campbell, John Y., Giglio, Stefano, Polk, Christopher, Turley, Robert, 2013. An Intertemporal CAPM with Stochastic Volatility, Working Paper, Harvard University, University of Chicago and London School of Economics.
- Campbell, John Y., Shiller, Robert J., 1988a. The dividend-price ratio and expectations of future dividends and discount factors. *Rev. Financ. Stud.* 1, 195–228.
- Campbell, John Y., Shiller, Robert J., 1988b. Stock prices, earnings, and expected dividends. *J. Financ.* 43, 661–676.
- Campbell, John Y., Vuolteenaho, Tuomo, 2004. Inflation illusion and stock prices. *Amer. Econ. Rev.* 94, 19–23.
- Cochrane, John H., 1991. Production-based asset pricing and the link between stock returns and economic fluctuations. *J. Financ.* 46, 209–237.
- Cochrane, John H., 2008. The dog that did not bark: A defense of return predictability. *Rev. Financ. Stud.* 21, 1533–1575.
- Cochrane, John H., 2011. Presidential address: discount rates. *J. Financ.* 66, 1047–1108.
- Corradi, Valentina, Distaso, Walter, Mele, Antonio, 2013. Macroeconomic determinants of stock volatility and volatility premiums. *J. Monet. Econ.* 60, 203–220.
- Corsi, Fulvio, 2009. A simple long memory model of realized volatility. *J. Financ. Econometrics* 7, 174–196.
- Drechsler, Itamar, Yaron, Amir, 2011. What's Vol got to do with it. *Rev. Financ. Stud.* 24, 1–45.
- Engle, Robert F., Mezrich, Joseph, 1996. GARCH for groups. *Risk* 9, 36–40.
- Engsted, Tom, Pedersen, Thomas Q., Tanggaard, Carsten, 2012. The log-linear return approximation, bubbles, and predictability. *J. Financ. Quant. Anal.* 47, 643–665.
- Epstein, Larry G., Zin, Stanley E., 1991. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis. *Journal of Political Economy* 99, 263–286.
- Fama, Eugene F., French, Kenneth R., 1988. Dividend yields and expected stock returns. *J. Financ. Econ.* 22, 3–25.
- Fama, Eugene F., French, Kenneth R., 1989. Business conditions and expected returns on stocks and bonds. *J. Financ. Econ.* 25, 23–49.
- Guo, Hui, Whitelaw, Robert F., 2006. Uncovering the risk-return relation in the stock market. *J. Financ.* 61, 1433–1463.
- Hamilton, James D., 1991. A Quasi-Bayesian approach to estimating parameters for mixtures of normal distributions. *J. Bus. Econom. Statist.* 9, 27–39.
- Hodrick, Robert J., 1992. Dividend yields and expected stock returns: Alternative procedures for inference and measurement. *Rev. Financ. Stud.* 5, 357–386.
- Kelly, Bryan T., 2011. Tail Risk and Asset Prices, Working Paper, University of Chicago Booth School of Business.
- Koijen, Ralph S.J., Nieuwerburgh, Stijn van, 2011. Predictability of returns and cash flows. *Annual Review of Financial Economics* 3, 467–491.
- Lamont, Owen, 1998. Earnings and expected returns. *J. Financ.* 53, 1563–1587.
- Lettau, Martin, Ludvigson, Sydney C., 2001. Consumption, aggregate wealth, and expected stock returns. *J. Financ.* 56, 815–849.
- Lettau, Martin, Ludvigson, Sydney C., 2005. Expected returns and expected dividend growth. *J. Financ. Econ.* 76, 583–626.
- Lettau, Martin, Ludvigson, Sydney C., 2013. Shocks and Crashes, Working Paper, UC Berkeley and New York University.
- Lewellen, Jonathan, 2004. Predicting returns with financial ratios. *J. Financ. Econ.* 74, 209–235.
- Maio, Paulo F., Santa-Clara, Pedro, 2013. Dividend yields, dividend growth, and return predictability in the cross-section of stocks. *J. Financ. Quant. Anal.* (forthcoming).
- Nakamura, Emi, Sergeyev, Dmitry, Steinsson, Jon, 2012. Growth-Rate and Uncertainty Shocks in Consumption: Cross-Country Evidence, Working Paper, Columbia University.
- Nieto, Belen, Rubio, Gonzalo, 2011. The volatility of consumption-based stochastic discount factors and economic cycles. *J. Bank. Financ.* 35, 2197–2216.
- Piatti, Ilaria, Trojani, Fabio, 2012. Predictable Risks and Predictive Regression in Present-Value Models, Working Paper, University of Lugano.
- Rigobon, Robert, 2003. Identification through heteroskedasticity. *Rev. Econ. Statist.* 85, 777–792.
- Rigobon, Roberto, Sack, Brian, 2003. Spillovers Across U.S. Financial Markets, Working Paper, NBER.
- Rozeff, M S., 1984. Dividend yields are equity risk premiums. *J. Portf. Manag.* 49, 141–160.
- Sentana, Enrique, Fiorentini, Gabriele, 2001. Identification, estimation and testing of conditionally heteroskedastic factor models. *J. Econometrics* 102, 143–164.
- Zhou, Guofu, Zhu, Yingzi, 2013. The Long-Run Risks Model: What Differences Can an Extra Volatility Factor Make? Working Paper, Olin School of Business, Washington University.