WORKING PAPER 10/2021 (ECONOMICS AND STATISTICS)

Predicting returns and dividend growth - the role of non-Gaussian innovations

Tamas Kiss, Stepan Mazur and Hoang Nguyen

ISSN 1403-0586

Örebro University School of Business
SE-701 82 Örebro, SWEDEN

Predicting returns and dividend growth - the role of non-Gaussian innovations

Tamás Kiss[(a)], Stepan Mazur[(a)], Hoang Nguyen[(a)]

[(a)] School of Business - Örebro University, Sweden

May 24, 2021

Abstract

In this paper we assess whether flexible modelling of innovations impact the predictive performance of the dividend price ratio for returns and dividend growth. Using Bayesian vector autoregressions we allow for stochastic volatility, heavy tails and skewness in the innovations. Our results suggest that point forecasts are barely affected by these features, suggesting that workhorse models on predictability are sufficient. For density forecasts, however, we find that stochastic volatility substantially improves the forecasting performance.

JEL Classification: C11, C58, G12

Keywords: Bayesian VAR; Dividend Growth Predictability; Predictive Regression; Return Predictability

1 Introduction

Since the seminal contribution of Campbell and Shiller (1988), we know that variation of the dividend-price ratio can be attributed either to variation in expected returns or variation in dividend growth. This observation has spurred a large empirical literature analysing jointly return and dividend-growth predictability using the dividend-price ratio as a predictor. The empirical investigation has been rich, yet, no conclusive evidence has been reached about the role of the dividend-price ratio as a predictor of returns and dividend growth.[1]

We contribute to this literature by analysing the role of non-Gaussianity of the variables. This is particularly important in the current context as both returns and dividend growth exhibit non-Gaussian behaviour (Kon, 1984; Harvey and Siddique, 2000; Jondeau and Rockinger, 2003, 2012; Adcock et al., 2015). Using the standard dataset on the stock market of the United States, we estimate Bayesian VAR models which can allow for stochastic volatility and non-Gaussian innovations, and assess the importance of these features both by in-sample and out-of-sample performance evaluation.

Our in-sample results suggest that accounting for more flexible distributions somewhat strengthens the evidence for return predictability while weakens that for dividend-growth. However, the effect is relatively small and suggest that for point forecasting purposes, the workhorse Gaussian, homoscedastic model is sufficient. In contrast, assessing out-of-sample density forecasting performance highlights a clear importance of stochastic volatility.

The paper is organised as follows. Section 2 presents the econometric framework. Section 3 discusses the data and the empirical results. Finally, Section 4 concludes.

2 The econometric framework

Let r_t denote the log-returns from period $t-1$ to t, Δd_t the change in log-dividends at time t, and dp_t the corresponding log dividend-price ratio. Defining $\mathbf{y}_t = (r_t, \Delta d_t, dp_t)'$, we can write the joint

[1] The techniques for empirical investigation range from predictive regressions Fama and French (1988); Lettau and Ludvigson (2005); Cochrane (2008); Chen (2009); Golez and Koudijs (2018); Golez (2014) to more complicated models Van Binsbergen and Koijen (2010); Koijen and Van Nieuwerburgh (2011), including vector autoregressions (VARs) (Campbell, 1991; Engsted et al., 2012).

model of return and dividend-growth predictability as a VAR(1) system

$$\mathbf{y}_t = \boldsymbol{\alpha} + \mathbf{B}\mathbf{y}_{t-1} + \mathbf{A}^{-1}\boldsymbol{\epsilon}_t,$$

with $\boldsymbol{\alpha} = (\alpha_r, \alpha_d, \alpha_{dp})'$, coefficient matrix $\mathbf{B}$ with elements b_{ij}, and innovation terms $\boldsymbol{\epsilon}_t = \left(\epsilon_t^r, \epsilon_t^d, \epsilon_t^{dp}\right)'$. $\mathbf{A}$ is a lower triangular matrix with ones on the diagonal absorbing the contemporaneous interaction of the endogenous variables. This is a standard specification used by Campbell (1991). However, earlier literature points out that because of the Campbell and Shiller (1988) identity, one equation in the above VAR is redundant.[2] Therefore, we use two bivariate systems with $\mathbf{y}_t = (r_t, dp_t)'$ and $\mathbf{y}_t = (\Delta d_t, dp_t)'$ for return and dividend-growth predictability, respectively. These bivariate systems contain the same information as the full trivariate system and avoid potential issues with multicollinearity between the variables (Cochrane, 2008; Engsted et al., 2012; Hjalmarsson and Kiss, 2021).

In the vast majority of empirical applications, the above VAR models are estimated assuming Gaussian, homoscedastic innovation terms. Most prominently, the features such as stochastic volatility, skewness or heavy tails of the innovation term are disregarded when it comes to either in-sample or out-of-sample evaluation of the prediction performance of the dividend-price ratio. In this paper, we relax these assumptions by using the framework of Karlsson et al. (2021) where the non-Gaussian innovations are derived as a Gaussian variance-mean mixture. Also stochastic volatility is allowed for to capture potential time variation in the second moment. In particular, we model the innovation term $\boldsymbol{\epsilon}_t$ as a vector of orthogonal skew-t (OST) distributions,

$$\boldsymbol{\epsilon}_t = \mathbf{W}_t\boldsymbol{\gamma} + \mathbf{W}_t^{1/2}\mathbf{H}_t^{1/2}\boldsymbol{e}_t, \quad t = 1, \ldots, T, \tag{2}$$

where $\boldsymbol{\gamma} = (\gamma_1, \gamma_2)'$ contains skewness parameters; $\mathbf{W}_t = diag(\xi_{1t}, \xi_{2t})$ is a diagonal matrix of mixing variables ξ_{it} that are mutually independent and follow an inverse gamma distribution with the same shape and rate parameters equal to $\nu_i/2$, i.e. $\xi_{it} \sim \mathcal{IG}(\frac{\nu_i}{2}, \frac{\nu_i}{2})$, $i = 1, 2$; $\mathbf{H}_t = diag(h_{1t}, h_{2t})$

[2] By Campbell and Shiller (1988), the following (approximate) present value identity holds between returns, dividend growth and the dividend price ratio:

$$r_t = const - \rho dp_t + \Delta d_t + dp_{t-1}, \tag{1}$$

where $\rho \approx 1$ is a linearization constant with exact values depending on the exact dataset.

is a diagonal matrix that captures the heteroskedastic volatility; and $\boldsymbol{e}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. Moreover, $\mathbf{W}_t$, $\mathbf{H}_t$, and $\boldsymbol{e}_t$ are mutually independent, then the marginal distribution of $\boldsymbol{\epsilon}_t$ is a vector of independent OST distributions (Aas and Haff, 2006; Karlsson et al., 2021).

We also assume a random walk process of the log volatility,

$$\log h_{it} = \log h_{it-1} + \sigma_i \eta_{it}, \quad i = 1, 2, \ t = 1, \ldots, T, \tag{3}$$

where $\eta_{it} \sim \mathcal{N}(0, 1)$. Note that the above model nests simpler models that we consider while assessing the importance of heteroscedasticity, heavy tails and skewness. In particular, the conditional distribution of vector $\mathbf{y}_t$ reduces to an orthogonal Student's t-distribution (OT.SV) when $\gamma_1 = \gamma_2 = 0$. Additionally as $\nu_i \to \infty$ for $i = 1, 2$, we approach the Gaussian VAR model with stochastic volatility (Gaussian.SV). Finally, the VAR models without stochastic volatility can be achieved by imposing $\sigma_i^2 = 0$ for $i = 1, 2$, with h_{i0} being the constant volatility. More details on the prior distributions used in the Bayesian inference are discussed in Appendix A, while further details about the inference method can be found in Karlsson et al. (2021).

We aim to assess the question whether returns or dividend growth are predictable by the dividend price ratio, and whether the relationship is affected by the presence of heavy tails and skewness. For this, we evaluate the above VAR models both in-sample (using the posterior distribution of parameters) and out-of-sample. For the latter, we use Mean Squared Forecasting Error (MSFE) for point forecast, and Log Predictive Score (LPS) as well as Continuously Ranked Probability Score (CRPS) for density forecast, see Gneiting and Raftery (2007). In all cases we identify significant improvements based on the one-sided Diebold and Mariano (1995) test with Newey-West standard innovations (Clark, 2011).

3 Empirical analysis

For both return and dividend-growth predictability, we analyse four specifications. The baseline is the VAR model where homoscedastic Gaussian innovations are assumed. We then allow sequentially for stochastic volatility, heavy tails and skewness in the innovation distribution. The analysis is based on quarterly returns, dividend growth and dividend-price ratio data from the first quarter of 1927 to the last quarter of 2019. The stock returns are captured by the return on the S&P500 index. Dividends are four-quarter moving sums of the dividends paid out by companies in the S&P500 index year to avoid calendar effects in dividend payouts. For calculating the dividend-price ratio, the S&P500 index is used as a price index. Data is obtained from the updated Welch and Goyal (2008) dataset. We use quarterly data because it provides us enough observation for a reliable Bayesian estimation and, at the same time, it is less noisy than the monthly data.[3] Figure 1 presents the data. The histograms in Figure 1 show clearly that the unconditional distribution of both returns and dividend growth are heavily leptokurtic and skewed, while the dividend-price ratio is closer to a normally distributed. This suggests that capturing non-Gaussianity in returns and the dividend-growth is potentially important in modelling these series.

3.1 In-sample results

Results from Bayesian estimation are collected in Table 1. Posterior means of the main parameters of interest are shown, along with their 90% credible intervals. Looking at the baseline Gaussian model, we see no evidence for return predictability by the dividend-price ratio (the credible interval for the posterior of the $b_{r,dp}$ coefficient is fairly wide and contains zero) and the dividend-price ratio is highly persistent ($b_{dp,dp}$ is close to unity). However, dividend growth appears predictable with a sizeable slope coefficient ($b_{d,dp}$ in the table) and a credible interval well in the negative range. The picture changes somewhat, once we allow for stochastic volatility. In particular, the coefficient capturing return predictability increases in magnitude and the credible interval no longer includes zero, which means evidence in favour of return predictability. At the same time, the posterior mean of the dividend growth coefficient shrinks in absolute terms. Also, the log marginal likelihoods (calculated as in Karlsson et al., 2021) show a strong improvement in both bivariate VARs once we allow for stochastic volatility.

The results are mostly unchanged once we allow for heavy tails and skewness: the return and dividend-growth coefficients retain the same sign, but their credible intervals suggest some (albeit weak) in-sample evidence for predictability. In general, more general modelling of higher order moments does not seem to impact the mean equations substantially. Looking at the credible

[3] Using quarterly data in a similar context is somewhat unusual, but it is certainly present in the literature, see Campbell and Viceira (2002); Pástor and Stambaugh (2009, 2012); Welch and Goyal (2008).

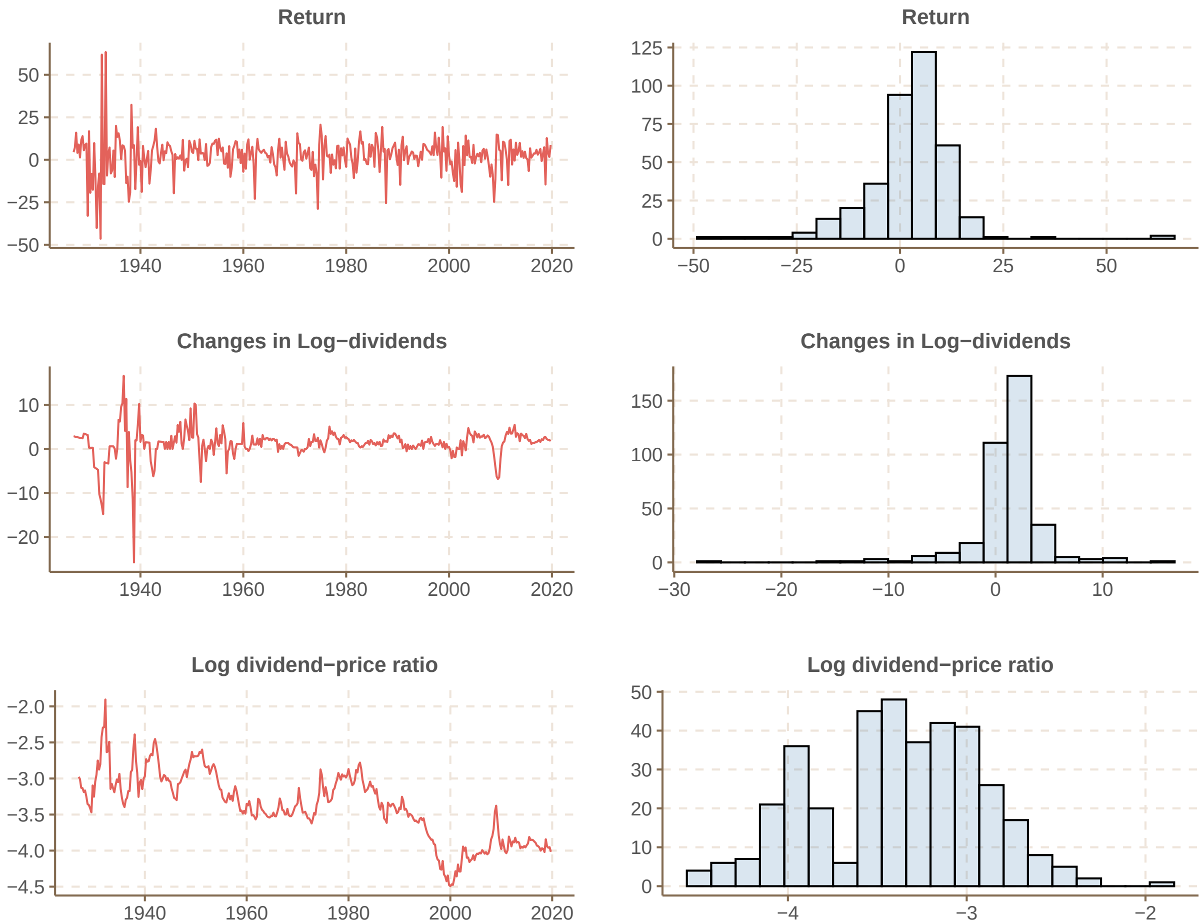

Figure 1: The plots show the quarterly log returns, log dividend-growth and log dividend-price ratio series used in the analysis. The left panels include the time series of the variables and the right panel shows histograms (with number of observations on the y-axis).

interval of the parameters governing heavy tails (ν_r) and asymmetry (γ_r) in returns, the results reveal that innovations to returns are characterized by both heavier-than-normal tails and negative skewness, even when stochastic volatility is allowed for. This is not the case for innovations to dividend-growth and the dividend-price ratio: the heavy tail parameters (ν_d and ν_{dp}) are large and the skewness parameters (γ_d and γ_{dp}) are close to zero. This suggests that these variables can be very well described by having Gaussian innovations, as long as time variation in the second moment is taken care of. The log marginal likelihood calculations suggest that the best model for returns is the one with stochastic volatility and Gaussian innovations. For the dividend-growth VAR, there is some small improvement in log marginal likelihoods when allowing for non-Gaussianity, and the model with orthogonal skew-t (OST) distributions turns out to be the best.

3.2 Forecasting performance

The out-of-sample exercise is mostly meant to confirm our in-sample evidence for return and dividend-growth predictability. To assess forecasting performance, we use return and dividend-growth predictions from 233 recursive forecasts between the first quarter of 1960 and the last quarter of 2017. We compare the forecasts to a benchmark Gaussian AR(1) model for both returns and dividend growth. For dividend-growth, an AR(1) specification is necessary due to the serial correlation in the series, while the benchmark is basically equal to the historical mean for the return series due to the low serial correlation[4]. We present results for both for the short run, and also for long-horizon (cumulative) return and dividend-growth forecasts from one to eight quarters ahead.[5]

For the point forecasts, one does not really benefit from abandoning Gaussianity. For returns we see an interesting pattern where the Gaussian model actually underperforms the benchmark, and it is overturned, at least in the one quarter horizon. This improvement disappears though on longer horizons, and even in the short horizon it is not statistically significant. For dividend-

[4] The AR(1) coefficient of return is -0.043(0.052) and that of log dividend is 0.599(0.042) where the numbers in the brackets are the standard errors.

[5] Since the VAR model is dynamically complete, we can form forecasts several periods ahead by rolling the system forward, computing values for $\hat{r}_{t+1|t}, \hat{r}_{t+2|t}, \ldots, \hat{r}_{t+h|t}$ for returns and $\widehat{\Delta d}_{t+1|t}, \widehat{\Delta d}_{t+2|t}, \ldots, \widehat{\Delta d}_{t+h|t}$. The long-horizon forecasts of the model are then given by

$$\hat{r}_{t\to t+h|t} = \hat{r}_{t+1|t} + \hat{r}_{t+2|t} + \cdots + \hat{r}_{t+h|t},$$
$$\widehat{\Delta d}_{t\to t+h|t} = \widehat{\Delta d}_{t+1|t} + \widehat{\Delta d}_{t+2|t} + \cdots + \widehat{\Delta d}_{t+h|t}.$$

Table 1: Summary of the posterior samples from VAR models for the quarterly SP data (1927-2019)

	Gaussian	Gaussian.SV	OT.SV	OST.SV
		Bivariate VAR - Return - DP Ratio		
$b_{r,dp}$	0.442	0.823	0.794	0.725
	(-0.56;1.502)	(0.061;1.667)	(0.031;1.641)	(-0.033;1.565)
$b_{dp,dp}$	0.988	0.992	0.993	0.993
	(0.974;0.999)	(0.982;0.999)	(0.983;0.999)	(0.983;0.999)
ν_r	-	-	16.117	9.506
	-	-	(5.907;39.615)	(4.552;23.428)
ν_{dp}	-	-	38.806	40.707
	-	-	(17.829;68.66)	(19.623;71.224)
γ_r	-	-	-	-1.115
	-	-	-	(-2.333;0.292)
γ_{dp}	-	-	-	-0.004
	-	-	-	(-0.022;0.013)
LML	-908.351	-457.476	-459.006	-467.689
		Bivariate VAR - Dividend - DP Ratio		
$b_{d,dp}$	-0.72	-0.222	-0.216	-0.225
	(-1.118;-0.303)	(-0.409;-0.039)	(-0.403;-0.035)	(-0.414;-0.04)
$b_{dp,dp}$	0.979	0.986	0.986	0.985
	(0.958;0.996)	(0.971;0.998)	(0.972;0.998)	(0.97;0.998)
ν_d	-	-	25.512	30.185
	-	-	(8.844;53.142)	(12.22;55.53)
ν_{dp}	-	-	23.848	20.792
	-	-	(10.195;48.759)	(10.071;40.495)
γ_d	-	-	-	-0.094
	-	-	-	(-1.022;0.686)
γ_{dp}	-	-	-	0.05
	-	-	-	(0.013;0.1)
LML	-637.364	-318.389	-317.692	-314.561

The table compares the estimation of the parameters using the four specifications we consider: the Gaussian VAR (Gaussian), the Gaussian VAR with stochastic volatility (Gaussian.SV), the VAR with stochastic volatility and orthogonal t-distributed innovations (OT.SV), and the VAR with stochastic volatility and orthogonal skew-t distributed innovations (OST.SV). Estimation is based on the quarterly data using the S&P500 data between 1927 and 2019. $b_{r,dp}$, $b_{d,dp}$ are the predictive coefficients and $b_{dp,dp}$ is the persistence of the dividend-price ratio obtained from the respective $\mathbf{B}$ matrices. ν_r, ν_d, ν_{dp} are heavy tail parameters of the returns, dividend-growth and the dividend-price ratio respectively. Similarly, γ_r, γ_d, γ_{dp} are asymmetry parameters of these variables. The numbers in the brackets are the 90% credible intervals. The table also presents log marginal likelihoods (LML) for all the model specifications, evaluated as in Karlsson et al. (2021).

growth predictability, the homoscedastic, Gaussian model performs substantially worse than the benchmark. Accounting for stochastic volatility, heavy tails and skewness improves in terms of MSFE for almost all horizons, these changes are not statistically significant though.

Table 2: MSFEs and relative MSFEs [relative to the Gaussian AR(1) model]

	MSFEs and relative MSFEs				
	1Q	2Q	3Q	4Q	8Q
(a) Returns					
AR(1)	63.525	136.547	200.784	261.241	490.984
Gaussian	1.007	1.001	1.001	1.001	1.009
Gaussian.SV	0.984	1.002	1.002	1.000	1.025
OT.SV	0.986	1.003	1.006	1.009	1.045
OST.SV	0.985	1.003	1.006	1.007	1.037
(b) Dividends					
AR(1)	1.368	5.109	12.708	24.443	97.446
Gaussian	1.108	1.202	1.229	1.241	1.304
Gaussian.SV	0.947	0.918	0.921	0.943	1.029
OT.SV	0.944	0.915	0.920	0.944	1.033
OST.SV	0.943	0.914	0.917	0.938	1.016

The first line reports the MSFE of the benchmark Gaussian VAR(1) model without stochastic volatility during 1960:Q1-2017:Q4 (233 recursive estimations). The relative improvement is computed as the ratio of the MSFE of alternative specifications over the benchmark. The entries less than 1 indicate that the given model is better. ***,**,* denote that the corresponding model significantly outperforms the benchmark at 1%, 5%, 10% level based on the one-sided Diebold and Mariano (1995) test where the standard errors of the test statistics are computed with the Newey–West estimator (Clark, 2011).

The picture is more interesting when it comes to density forecasts. Results in Table 3 based on the LPS measure show that there is already an improvement in terms of density forecast of returns if the dividend-price ratio is used as a predictor variable. Furthermore, allowing for stochastic volatility improves substantially over the the Gaussian model, across all horizons, while further flexibility of the innovation terms is not supported by our results. For dividend growth, we find no evidence for the dividend-price ratio to improve density forecasts, however, stochastic volatility helps for this variable as well, in particular at within-year horizons. Results based on the CRSP measure are collected in Table 4. They show an overall similar picture with the only main exception that the dividend-price ratio in itself does not help improve density forecasts.

Table 3: LPS and relative LPSs [relative to the Gaussian VAR(1) model]

	LP and relative LPs				
	1Q	2Q	3Q	4Q	8Q
(a) Returns					
AR(1)	-4.734	-5.327	-5.427	-5.450	-5.481
Gaussian	0.037**	0.055*	0.097**	0.148**	0.292**
Gaussian.SV	1.242***	1.418***	1.275***	1.076***	0.654*
OT.SV	1.246***	1.419***	1.281***	1.088***	0.698*
OST.SV	1.243***	1.417***	1.272***	1.086***	0.711*
(b) Dividends					
AR(1)	-1.729	-2.379	-2.794	-3.094	-3.758
Gaussian	-0.021	-0.039	-0.059	-0.080	-0.141
Gaussian.SV	0.179***	0.199***	0.101	-0.010	-0.283
OT.SV	0.175***	0.191***	0.100	-0.004	-0.258
OST.SV	0.190***	0.207***	0.119*	0.018	-0.223

The first line reports the LP of the benchmark Gaussian VAR(1) model without stochastic volatility during 1960:Q1-2017:Q4 (233 recursive estimations). The relative improvement is computed as the difference of the LPS of alternative specifications over the benchmark. The entries greater than 0 indicate that the given model is better. ***,**,* denote that the corresponding model significantly outperforms the benchmark at 1%, 5%, 10% level based on the one-sided Diebold and Mariano (1995) test where the standard errors of the test statistics are computed with the Newey–West estimator (Clark, 2011).

Table 4: CRPS and relative CRPSs [relative to the Gaussian VAR(1) model]

	CRPS and relative CRPSs				
	1Q	2Q	3Q	4Q	8Q
(a) Returns					
AR(1)	-4.603	-6.877	-8.394	-9.627	-12.800
Gaussian	-0.016	-0.016	-0.020	-0.027	-0.055
Gaussian.SV	0.327***	0.499***	0.527**	0.446	-0.194
OT.SV	0.336***	0.492***	0.535**	0.469*	-0.196
OST.SV	0.336***	0.502***	0.548**	0.498*	-0.132
(b) Dividends					
AR(1)	-0.689	-1.306	-2.003	-2.737	-5.507
Gaussian	-0.025	-0.087	-0.160	-0.238	-0.631
Gaussian.SV	0.084***	0.192***	0.252***	0.261**	0.050
OT.SV	0.085***	0.193***	0.250***	0.260**	0.032
OST.SV	0.084***	0.195***	0.259***	0.275**	0.096

The first line reports the LPS of the benchmark Gaussian VAR(1) model without stochastic volatility during 1960:Q1-2017:Q4 (233 recursive estimations). The relative improvement is computed as the difference of the LPS of alternative specifications over the benchmark. The entries greater than 0 indicate that the given model is better. ***,**,* denote that the corresponding model significantly outperforms the benchmark at 1%, 5%, 10% level based on the one-sided Diebold and Mariano (1995) test where the standard errors of the test statistics are computed with the Newey–West estimator (Clark, 2011).

4 Concluding remarks

In this paper we have assessed whether the empirical evidence on return and dividend-growth predictability with the dividend-price ratio changes if we allow for a flexible modelling of the innovation to these variables. As a modelling framework we use a Bayesian VAR model with stochastic volatility, as well as heavy-tailed and skewed innovations. Our in-sample results suggest that adding heteroscedasticity and non-Gaussian innovations strengthens somewhat the evidence for return predictability, while weakens that of the dividend-growth, although changes are not substantial. Also, the innovations seem to be slightly heavy-tailed and negatively skewed based on the in-sample estimates. Since changes in the mean equation due to flexible modelling of innovations are not substantial for the slope coefficients, the above mentioned in-sample results do not translate into any significant out-of-sample gains in terms of point forecasts. For density forecasts however, we do see a significant improvement, as stochastic volatility appears to be the main driver behind density forecasts of both returns and dividend-growth.

Acknowledgment

Financial support from Jan Wallanders och Tom Hedelius stiftelse (grants number Bv18-0018, P18-0201 and W19-0021), and from the internal research grants provided by Örebro University, is gratefully acknowledged. The computations were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at HPC2N partially funded by the Swedish Research Council through grant agreement no. 2018-05973.

Appendix

A Prior distributions

The Bayesian approach is employed to make inferences for the set of the model parameters $\boldsymbol{\theta} = \{\boldsymbol{\alpha}', vec(\mathbf{B})', \mathbf{a}', \boldsymbol{\gamma}', \boldsymbol{\nu}', \boldsymbol{\sigma}^{2\prime}, \xi'_{1:2,1:T}, \mathbf{h}'_{1:2,0:T}\}'$, where $\mathbf{a} = (a_{21})$ is the stack vector of the elements in the lower triangular matrix $\mathbf{A}$, and $\boldsymbol{\sigma}^2 = (\sigma_1^2, \sigma_2^2)'$. Following Karlsson et al. (2021), the prior distribution of $\boldsymbol{\alpha}$ and $vec(\mathbf{B})$ follows the Minnesota prior distribution with the overall shrinkage $l_1 = 0.2$ and the cross-variable shrinkage $l_2 = 0.5$, see Koop and Korobilis (2010). The prior distribution $\mathbf{a} \sim \mathcal{N}(0, 10\mathbf{I})$ which imposes a weak assumption of no interaction among endogenous variables. The degree of freedom for each variable follows a gamma distribution, i.e. $\nu_i \sim \mathcal{G}(2, 0.1)$ so that the prior mean goes around 20. The asymmetry parameter follows a standard normal distribution, $\boldsymbol{\gamma}_i \sim \mathcal{N}(0, 1)$ for $i = 1, 2$. The mixing variables distribute as $\xi_{it}|\nu_i \sim \mathcal{IG}(\frac{\nu_i}{2}, \frac{\nu_i}{2})$ due to the model settings. The prior for the variance of shock to the volatility is $\sigma_i^2 \sim \mathcal{G}(\frac{1}{2}, \frac{1}{2})$, see Kastner and Frühwirth-Schnatter (2014). In all cases, $\log h_{i0} \sim \mathcal{N}\left(\log \hat{\Sigma}_{i,OLS}, 4\right)$ where $\hat{\Sigma}_{i,OLS}$ is the estimated variance of the AR(1) model using the ordinary least square method, see Clark and Ravazzolo (2015).

References

Aas, K. and Haff, I. H. (2006). The generalized hyperbolic skew Student's t-distribution. *Journal of Financial Econometrics*, 4(2):275–309.

Adcock, C., Eling, M., and Loperfido, N. (2015). Skewed distributions in finance and actuarial science: A review. *The European Journal of Finance*, 21(13-14):1253–1281.

Campbell, J. Y. (1991). A variance decomposition for stock returns. *The Economic Journal*, 101(405):157–179.

Campbell, J. Y. and Shiller, R. J. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies*, 1(3):195–228.

Campbell, J. Y. and Viceira, L. M. (2002). Is the stock market safer for long-term investors? *Strategic Asset Allocation*, (33):88–120.

Chen, L. (2009). On the reversal of return and dividend growth predictability: A tale of two periods. *Journal of Financial Economics*, 92(1):128–151.

Clark, T. E. (2011). Real-time density forecasts from Bayesian vector autoregressions with stochastic volatility. *Journal of Business & Economic Statistics*, 29(3):327–341.

Clark, T. E. and Ravazzolo, F. (2015). Macroeconomic forecasting performance under alternative specifications of time-varying volatility. *Journal of Applied Econometrics*, 30(4):551–575.

Cochrane, J. H. (2008). The dog that did not bark: A defense of return predictability. *The Review of Financial Studies*, 21(4):1533–1575.

Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & economic statistics*, 13(3):134–144.

Engsted, T., Pedersen, T. Q., and Tanggaard, C. (2012). Pitfalls in var based return decompositions: A clarification. *Journal of Banking & Finance*, 36(5):1255–1265.

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

Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American statistical Association*, 102(477):359–378.

Golez, B. (2014). Expected returns and dividend growth rates implied by derivative markets. *The Review of Financial Studies*, 27(3):790–822.

Golez, B. and Koudijs, P. (2018). Four centuries of return predictability. *Journal of Financial Economics*, 127(2):248–263.

Harvey, C. R. and Siddique, A. (2000). Conditional skewness in asset pricing tests. *The Journal of finance*, 55(3):1263–1295.

Hjalmarsson, E. and Kiss, T. (2021). Dividend growth does not help predict returns compared to likelihood-based tests: An anatomy of the dog. *Critical Finance Review*, Forthcoming.

Jondeau, E. and Rockinger, M. (2003). Conditional volatility, skewness, and kurtosis: existence, persistence, and comovements. *Journal of Economic dynamics and Control*, 27(10):1699–1737.

Jondeau, E. and Rockinger, M. (2012). On the importance of time variability in higher moments for asset allocation. *Journal of Financial Econometrics*, 10(1):84–123.

Karlsson, S., Mazur, S., and Nguyen, H. (2021). Vector autoregression models with skewness and heavy tails. *Working Paper 8/2021*, School of Business, Örebro University.

Kastner, G. and Frühwirth-Schnatter, S. (2014). Ancillarity-sufficiency interweaving strategy (asis) for boosting mcmc estimation of stochastic volatility models. *Computational Statistics & Data Analysis*, 76:408–423.

Koijen, R. S. and Van Nieuwerburgh, S. (2011). Predictability of returns and cash flows. *Annu. Rev. Financ. Econ.*, 3(1):467–491.

Kon, S. J. (1984). Models of stock returns—a comparison. *The Journal of Finance*, 39(1):147–165.

Koop, G. and Korobilis, D. (2010). *Bayesian multivariate time series methods for empirical macroeconomics*. Now Publishers Inc.

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

Pástor, L. and Stambaugh, R. F. (2009). Predictive systems: Living with imperfect predictors. *The Journal of Finance*, 64(4):1583–1628.

Pástor, L. and Stambaugh, R. F. (2012). Are stocks really less volatile in the long run? *The Journal of Finance*, 67(2):431–478.

Van Binsbergen, J. H. and Koijen, R. S. (2010). Predictive regressions: A present-value approach. *The Journal of Finance*, 65(4):1439–1471.

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