Using Analysts' Forecasts for Stock Predictions - An Entropic Tilting Approach*

Christoph Frey†

Erasmus University Rotterdam

July 24, 2017

Abstract

In this paper, we combine predictive density forecasts for US stock returns from Bayesian vector autoregressions with financial analysts' forecasts via entropic tilting. In particular, we modify the predictive density of the asset returns to match moment conditions that are formed based on average analysts' forecasts. The advantage of this approach is that we can combine model-based time-series information with external, forward-looking information in a parsimonious way using closed-form solutions. Using a model with time-varying coefficients and stochastic volatility, we show that tilting the predictive asset return distribution towards the mean and variance of target price implied expected returns increase prediction accuracy for both point and density forecasts.

Keywords: Bayesian vector autoregression, return prediction, tilting

JEL classification: C11, C58, G11

*I would like to thank Dick van Dijk and Winfried Pohlmeier for helpful comments. Financial support by the German Science Foundation (DFG, Project "Development and application of new methodologies of combining time series and expert survey data for economic forecasting.") and the Graduate School of Decision Sciences (University of Konstanz) is gratefully acknowledged. All remaining errors are ours.

†Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, Room H11-32, P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands. Phone (office): (+31) 10 408 1253, fax: (+31) 10 4089162, email: christoph.frey@gmail.com.

Using Analysts' Forecasts for Stock Predictions - An Entropic Tilting Approach

Abstract

In this paper, we combine predictive density forecasts for US stock returns from Bayesian vector autoregressions with financial analysts' forecasts via entropic tilting. In particular, we modify the predictive density of the asset returns to match moment conditions that are formed based on average analysts' forecasts. The advantage of this approach is that we can combine model-based time-series information with external, forward-looking information in a parsimonious way using closed-form solutions. Using a model with time-varying coefficients and stochastic volatility, we show that tilting the predictive asset return distribution towards the mean and variance of target price implied expected returns increase prediction accuracy for both point and density forecasts.

1 Introduction

Predicting stock returns is a popular exercise for professionals and academics in finance. It is a challenging task because of estimation and model uncertainty, because of a substantial unpredictable component (shocks) in future stock returns and because successful forecasting models may be exploited by other market participants causing trading behavior that destroys the prediction power of the model. Recent empirical findings suggest that the equity premium is predictable when accounting for model and estimation uncertainty through time-varying parameters, stochastic volatility and by using Bayesian predictive regressions (Zellner and Chetty, 1965; Barberis, 2000; Pettenuzzo and Ravazzolo, 2016). While the amount of predictability from variables such as the dividend-earnings ratio is limited in terms of out-of-sample R^2, it can translate into substantial portfolio gains for an expected utility investor (Johannes et al., 2014).[1]

Most prediction models relate the future stock returns to past observations of asset specific variables such as the dividend yield or to general macroeconomic indicators such as inflation. However, the market stock price is forward-looking as it reflects the expectations of market participants about the future cash flows of the company. It is natural to expect that using forward-looking information should also be beneficial for stock return predictions. A recent example is Metaxoglou et al. (2016), who improve equity premium forecasts by using option prices. In this paper, we use professional analyst's forecasts to improve asset return predictions. We do so by reweighting the predictive return distributions by a method called *entropic (exponential) tilting* to incorporate the information contained in analysts' forecasts, such as target prices. The idea of this method is to modify the predictive density of the asset returns to match certain moment conditions that are formed based on average analysts' forecasts. The advantage of this approach is that we can combine model-based time-series information with external, forward-looking information in a parsimonious way using closed-form solutions.

Professional security analysts provide market analyses, make earnings forecasts and give investment advise by providing twelve months forward target prices and stock recommendations. Their reports and opinions can have major short- and long-run impact on stock prices (e.g. Irvine, 2003) and provide a powerful way to disseminate financial information to market

[1]See Rapach and Zhou (2013) for a recent overview on return predictability.

participants.[2] And a vast literature exists about the shortcomings of analyst forecasts revealing skewed incentives (conflicts of interest), herding and biases to please clients that may create market inefficiencies.[3]

In this study, we do not evaluate the accuracy of the analysts' forecasts, but we use the (dis-)agreement in the analysts' forecast to regularize the predictive return distribution. In particular, we restrict the mean and variance of the predictive distribution to coincide with the mean and the variance of monthly target price implied expected returns, i.e. simple returns between the current spot and the target price. While we find that restricting the variance of the asset returns is beneficial in terms of out-of-sample performance, as it provides a forward looking measure for (un-)certainty in the market, only restricting the mean has no particular forecasting power. Target prices are usually higher than current spot prices and so there is an upward bias in the target price implied expected returns which is not beneficial for the forecast performance.

Of course we could also include the analyst information simply as another predictor variable in the predictive regressions, but this would add further parameters and more estimation noise to the prediction problem. Using the analyst information instead in a tilting framework, we only change the shape of the predictive distribution by reweighting it and hence do not require the data to formalize the relationship between asset returns and target prices.

Another econometric contribution of this paper is the use of a large Bayesian vector autoregressive system to formalize the predictive relationship between asset returns and a great number of predictor variables. To overcome the computational burden that arises in a recursive forecasting exercise, we adopt the so-called *forgetting factors* approach of Koop and Korobilis (2013) which allows for all the features recently found to be important to find significant return predictability: Time-varying parameters, stochastic volatility, parameter shrinkage as well as dynamic model averaging and variable selection. The idea of forgetting factors is to approximate the conditional error term variance in a Kalman filter type estimation of the time-varying model, reducing the computational burden significantly. While Dangl and Halling (2012) used a similar forgetting factor approach for return prediction, we will combine this method with a Minnesota-type prior that restricts the parameter matrix to deal with the case of many predictors.

It is also worth mentioning that market excess returns, such as the S&P500, might not be

[2]The US Bureaus of Labor statistics report over 275,000 thousand financial analysts jobs in 2014 (see `http://www.bls.gov/ooh/business-and-financial/financial-analysts.htm`, checked on 19.08.2015).

[3]See Ramnath et al. (2008) for an in-depth review of the analyst forecast literature.

the ideal candidate to find predictability, because of their aggregation over many sectors and industries. In this paper instead, we will investigate the predictability of individual assets by looking at a cross-section of Dow Jones index constituents. We will investigate if these returns are predictable using company-specific characteristics like the book-to-market ratio or by using market and economic indicators such as inflation. Out-of-sample studies for cross-sectional returns are limited and a few exceptions among others are Avramov (2002) and Rapach et al. (2015). While these studies use factor and industry portfolio returns, we fill the gap and go down to the level of individual equity asset predictability.

Figure 1 serves as an illustrative example. It shows the IBM spot price, the mean twelve months forward target price and the percentage of buy recommendations from all recommendations (buy, sell, hold) of the IBM stock. While the correlation between the spot and the target price is almost perfect (0.9721), spot price and buy recommendations are negatively correlated (-0.6470). In the plot the target price is almost always higher than the current spot price, indicating an upward bias in these forecasts. The only times when the two plots coincide is after price drops when the spot price starts to increase again, i.e. in late 2002, 2007 and 2014. That is, financial analyst might not be able to predict trend changes (regime switches) but they are able to forecast the price direction. This is in line with the more volatile buy recommendations, which only on average indicate the stock price direction.

While this pattern is similar also for the other 19 stocks considered here, comparing the target price with the observed spot price twelve months ahead gives a mixed picture. Table 1 provides the root mean squared forecast errors (RMSFE) between the two for the 20 stocks and compares it to the two year historical mean. Only for 11 out of the 20 stocks the target prices were better forecasts than the historical mean, e.g. for the IBM stock.

Figure 2 gives an idea about the predictive power of individual predictors against a benchmark intercept only model by plotting three out-of-sample performance measures for the IBM stock. Considering the first panel showing the cumulative sum of squared forecast errors differentials, we see that mostly all predictors fail to outperform the intercept only model, especially in the financial crisis. The only exception is the log return of the mean analyst twelve months forward target prices. Note that values above zero indicate that a given predictor has better forecast performance than the benchmark model, while negative values suggest the opposite. The target price itself only has some predictive power between 2008 and 2010 (square markers). This might suggest that not the level of the predicted target price matters, but that the timing of the shifts, represented by the target price implied returns, have predictive power. Also, we see that the average analyst recommendations have no particular predictive power for the IBM stock.

The paper is now organized as follows. Section 2 gives an overview about the findings from other authors trying to exploit analysts' forecasts and reviews the state of the art on return predictability. Section 3 describes the applied Bayesian vector autoregression model, which relies on the forgetting factor approach by Koop and Korobilis (2013) for large systems. We then introduce the concept of entropic tilting and explain how we translate the analyst forecasts into moment conditions for the predictive return distribution. Section 4 summarizes the set-up of the empirical study and presents the results. Section 5 concludes and gives an outlook on further generalizations.

2 Predictive powers

2.1 Analysts' forecasts

A great number of studies document the investment value of sell-side analyst research: Barber et al. (2001); Green (2006) report portfolio benefits and abnormal returns from simple trading strategies that go long in stocks with favorable analyst recommendations and short in stocks with unfavorable ratings. Jegadeesh et al. (2004) instead finds that the predictive power of the level of consensus analyst recommendations varies substantially across assets and proposes to use the changes or revisions in consensus analyst recommendations instead. Jegadeesh and Kim (2006) applies this to build portfolio strategies that are long (short) in stocks with positive (negative) recommendation revisions yielding significant positive abnormal risk-adjusted returns against the market. Cvitanić et al. (2006) further shows that such simple long-short portfolios can be outperformed by more complicated multi-period expected utility maximizing strategies that are estimated using consensus analysts recommendations. This is not the case for He et al. (2013), who find no significant portfolio returns from a Black and Litterman (1992) strategy incorporating analyst recommendations for the Australian stock market.

The evidence for return predictability from target prices is even more mixed: While Brav and

Lehavy (2003) show that target prices are generally informative about future stock prices and that there are substantial short-term market reactions in the stock price to target changes, Bonini et al. (2010) find little evidence for target price accuracy using different metrics measuring the prediction error between the target price and the current, twelve months forward and in-between spot prices. This is in line with Bradshaw et al. (2012) who find an average target price premium over the spot price of 15 percent and who report that only two thirds of the target prices are met by the stock price at some time during the forecast horizon in their sample. Lin et al. (2016) instead consider changes in target prices, but also find no evidence that institutional trading activity following the direction of the target price revisions yields higher risk-adjusted out-of-sample returns. Only Da and Schaumburg (2011) find that aggregating stocks across sectors according to their twelve month forward target price implied expected return, i.e. simple return between the current and the target price, yields significant risk-adjusted abnormal returns for different long-short portfolio.

However, despite this discouraging evidence, none of the studies above operate in a Bayesian framework using predictive distributions and none consider a tilting framework in which the analyst's forecasts serve as a shrinkage target instead of just being another predictor variable.

2.2 Predictive regressions

Kandel and Stambaugh (1987) propose a vector autoregression formulation to jointly model the dynamics of asset returns and its predictor:

$$r_{t+1} = a_r + b_r x_t + \varepsilon_{r,t+1}, \tag{1}$$

$$x_{t+1} = a_x + b_x x_t + \varepsilon_{x,t+1}, \tag{2}$$

where x_t is the explanatory variable (for example the dividend yield) and r_t is the asset return and $\varepsilon_t = (\varepsilon_{r,t}, \varepsilon_{x,t})' \overset{iid}{\sim} \mathrm{N}(0, \Sigma)$. Here, equation (2) is needed to model the times-series dynamics of the predictor variable trough an autoregressive process. This model has been used by various authors to investigate the predictability of the equity premium and to build portfolios using the predictive return distribution.

Kandel and Stambaugh (1996) estimate the model using Bayesian priors that center the mean of the slope coefficient in (1) to zero, implying no predictability and a weak form of market efficiency. Using a one-period model with power utility, they find that not accounting for the variability in the regressor can decrease the annual certainty equivalent up to 3%.

Pástor (2000) extends the model in (1) - (2) to a multivariate regression for N risky assets and apply a Normal inverse Wishart prior for the model mean and variance of the asset returns. They interpret the prior as a form to describe the disbelief of the investor in the given prediction/asset-pricing model. Importantly, they find that the optimal portfolio weights tend to be less extreme when using the joint predictive distribution of the asset returns by integrating out the estimated model parameters.

While these studies are concerned with investments for one period, Barberis (2000) investigates the impact of changes in the prediction variables to the asset allocation in the long-term. As the variance of the cumulative log returns grows less than linear in the investment horizon, stocks appear less risky and so an investor should allocate more funds to stocks when the investment horizon increases. However such an over-allocating to stocks disappears when accounting for estimation uncertainty and possible structural breaks in the parameters of the prediction model (Pettenuzzo and Timmermann, 2011).

Stock returns may be predicted by company specific characteristics such as the dividend-price ratio or they may be macroeconomic indicators. Given the great number of potential predictor variables, shrinkage and model averaging methods are natural candidates to reduce model risk in predictive regressions. For example, Avramov (2002) considers 14 predictor variables and extend the predictive system in (1) - (2). The author averages over all $2^{14} = 16384$ distinct model to predict the asset returns by weighting each model by its posterior weight. Similar examples for Bayesian model averaging (BMA) in the literature include Cremers (2002) and Wachter and Warusawitharana (2015). Furthermore, Pettenuzzo and Ravazzolo (2016) propose a Bayesian density combination approach over 15 predictor variables that, instead of minimizing a statistical loss function, maximizes the certainty equivalent of a power utility investor and report robust out-of-sample predictability across various performance measures.

Studies using time-varying coefficients in prediction models have also become numerous in recent years: While Welch and Goyal (2008) find no evidence for significant in-sample and out-of-sample predictability of 14 variables on the equity premium using a rolling window estimation, Dangl and Halling (2012) find predictability in one-step-ahead forecasts of monthly

S&P 500 returns by using Bayesian predictive regressions with time-varying parameters and model averaging over a range of predictor variables. Using time-varying coefficients as well as stochastic volatility but only two possible predictor variables, Johannes et al. (2014) document statistically and economically significant portfolio benefits for a dynamic rebalancing power utility investor using Bayesian predictive regressions. Very recently, Feng and Polson (2016) propose to perform a *predictive* cross-validation approach to select the right amount of shrinkage that provides optimal posterior point estimates instead of marginalizing out the model parameters to obtain a full predictive return distribution.

Very related to this paper is Metaxoglou et al. (2016), who are the first to apply entropic tilting to predict the US equity premium. For this, the authors rely on option data to form corresponding moment conditions. In particular, the authors use the second moment of the risk-neutral density implied by option prices on the S&P500 to tilt predictive distributions obtained from OLS model combinations with individual predictors, where the combination weights are chosen such that the model minimizes a discounted mean squared forecast error (Rapach and Zhou, 2013). For monthly returns and one-step ahead forecasts, they find that the tilted forecasts outperform an historical average based on various performance measures and for various sub-samples.

3 Methodology

3.1 Prediction model

In this study we follow the literature and use vector autoregressions (VARs) to model the relationship between asset returns and predictor variables. Extending the VAR given in equation (1) - (2) to a system with K predictors reads as follows:

$$\begin{bmatrix} r_t \\ x_t \end{bmatrix} = a + \sum_{i=1}^{p} A_i \begin{bmatrix} r_{t-i} \\ x_{t-i} \end{bmatrix} + \varepsilon_t, \quad t = 1, \dots, T, \tag{3}$$

where r_t is the excess return of a particular stock, $x_t = [x_{1,t}, \dots, x_{K,t}]'$ is a $K \times 1$ vector of predictor variables and $\varepsilon_t \overset{iid}{\sim} \mathrm{N}(0, \Sigma)$.

The number of parameters in the system in (3) grows quickly with the number of included predictor variables. We therefore focus on two restrictions to reduce estimation noise. Since every VAR(p) system can be written in VAR(1) companion form, we restrict $p = 1$. Second, in (1) - (2) both, the excess return and the predictor variable, only depend on their own lag, but not on the lag of the other variable. We follow this and restrict the system such that r_t depends on the entire x_{t-1} vector but $x_{k,t}$, $1 \leq k \leq K$, only depends on its own lag $x_{k,t-1}$. Compactly, the resulting model is of the form

$$y_t = (r_t, x_t)' = a + A_1 y_{t-1} + \varepsilon_t, \tag{4}$$

where $a = (a_r, a_{x_1}, \ldots, a_{x_K})'$ and $A_1 = \begin{pmatrix} 0 & A_1^{1,2} & A_1^{1,3} & \cdots & A_1^{1,K+1} \\ 0 & A_1^{2,2} & 0 & \cdots & 0 \\ \vdots & \ddots & A_1^{3,3} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & A_1^{K+1,K+1} \end{pmatrix}$. The zero restriction follows from (1) - (2) using multiple predictor variables. Usually, the correlation between the return and its first lag is very low, supporting the restriction $A_1^{1,1} = 0$. All other variables are supposed to follow an autoregressive process of order 1. To implement these restrictions *softly* on the slope coefficient matrix A_1we use a variant of the Minnesota prior (Doan et al., 1984). Further specifying independent marginal normal priors for each parameter yields the joint prior distribution through multiplication of the independent marginals. That is

$$\begin{aligned} p(a) &\sim \mathrm{N}\big(0, \zeta \times \mathrm{I}_{(K+1 \times K+1)}\big)\,, && (5) \\ p(A_1^{1,1}) &\sim \mathrm{N}(0, \varrho \times 1)\,, && (6) \\ p(A_1^{1,k}) &\sim \mathrm{N}\left(0, \zeta \times \frac{\sigma_r^2}{\sigma_{x_k}^2}\right), \quad k = 1, \ldots, K && (7) \\ p(A_1^{k,1}) &\sim \mathrm{N}\left(0, \varrho \times \frac{\sigma_{x_k}^2}{\sigma_r^2}\right), \quad k = 2, \ldots, K, && (8) \\ p(A_1^{k,l}) &\sim \mathrm{N}\left(\underline{A}_1^{k,l}, \varrho \times \frac{\sigma_{x_k}^2}{\sigma_{x_l}^2}\right), \quad k = 2, \ldots, K, \quad l = 2, \ldots, K && (9) \end{aligned}$$

with $\underline{A}_1^{k,l} = d_k$ if $k = l$, and $\underline{A}_1^{k,l} = 0$ otherwise.

Following Frey and Mokinski (2016), we set $d_k = 0$ for each real variable, and $d_k = 0.8$ for the nominal variables. Further we fix $\varrho = 10^{-4}$ and $\zeta = 0.2$, a common choice for the tightness parameter of the Minnesota prior in the Bayesian VAR forecasting literature according to Carriero et al. (2015). Note that the prior in (7) is centered around zero implying no predictability. Finally, the ratios $\sigma^2_{x_k}/\sigma^2_r\ \forall k$ and $\sigma^2_{x_k}/\sigma^2_{x_l}\ \forall k, l$ account for differences in the scale and variability of the different predictor variables. $\sigma^2_{x_k}\ \forall k$ and also σ^2_r are approximated by the residual variances of an AR(1) regression for k-th variable and the asset return. The specification is completed by assuming an independent diffuse prior for Σ, $p(\Sigma) \propto |\Sigma|^{-2(2M+1)/2}$.[4]

3.1.1 Time-varying Bayesian VAR and stochastic volatility (TVP-BVAR with SV)

The literature provides various examples favoring equity prediction models with time-varying parameters (TVP) (Dangl and Halling, 2012), stochastic volatility (SV) (Johannes et al., 2014) and Bayesian model averaging techniques (Pettenuzzo and Ravazzolo, 2016). To evaluate the predictive performance, for example marginal likelihoods for individual models have to be easily available without great computational costs at each point of the forecasting period. While this may be so for simple constant parameter models through the use of conjugate priors, they are almost infeasible to obtain for large VAR models such as given in (3) with many parameters. The latter require informative priors to reduce estimation noise which rely on Markov Chain Monte Carlo (MCMC) methods for estimation at each point in time with typical tens of thousands of simulation draws to ensure convergence.

The same is true for time-varying parameter models with stochastic volatility that not only require Kalman filtering for the regression coefficients but also computational costly sampling methods for the error term variances. To overcome the computational burden that arises in a recursive forecasting exercise, we adopt the so-called *forgetting factors* approach of Koop and Korobilis (2013) which also allows for all the features to model return predictability: Time-varying parameters, stochastic volatility, parameter shrinkage as well as dynamic model averaging and variable selection. Forgetting factors are used in state space models to allow for a moderate variation of the predictive variance over time. Let us consider a time-varying VAR version of (4)

[4] Posterior results for the full model are obtained in a standard fashion and are omitted here for parsimony.The interested reader is referred to Koop and Korobilis (2010).

with stochastic volatility which can be expressed as follows:

$$y_t = a_t + A_{1,t}\, y_{t-1} + \varepsilon_t, \tag{10}$$

$$A_t = \phi A_{t-1} + (1-\phi)\underline{A}_0 + u_t, \tag{11}$$

where $A_t = [a_t \;\; A_{1,t}]$ is time-indexed for every parameter, $\varepsilon_t \overset{iid}{\sim} \mathrm{N}(0, \Sigma_t)$, $u_t \overset{iid}{\sim} \mathrm{N}(0, \Omega_t)$ and ε_t and u_s are independent of each other for all t and s. Here, ϕ is an unknown parameter governing the mean of A_t. While $\phi = 1$ implies a random walk behavior, $\phi = 0$ implies a random behavior of each A_t around $\underline{A}_0$. Here, we will use the means of the Minnesota prior described in the previous section to specify $\underline{A}_0$. Since ϕ adds another layer to the prediction model, the restrictions imposed on the coefficient matrix are relaxed compared to the constant coefficient model.

Typically, the estimation of the system (10) - (11) relies on MCMC techniques. Given the initial conditions A_0, Σ_0 and Ω_0, it involves drawing A_t conditional on Σ_t and Ω_t (e.g. through a Kalman filter), then drawing Σ_t conditional on A_t and Ω_t, the sampling Ω_t given A_t and Σ_t and eventually drawing further parameters given conditional on A_t, Σ_t, and Ω_t for all t. This is computationally demanding as it involves simulating Σ_t, and Ω_t for every $t = 1, \ldots, T$. The idea of the forgetting factors here is to avoid simulating Ω_t recursively for each t. Instead, we avoid using Ω_t in the Kalman filter by approximating the one-step ahead predictor variance of $A_t|y^{t-1} \sim \mathrm{N}\big(A_{t|t-1}, P_{t|t-1}\big)$, i.e. $P_{t|t-1}$, by the variance of the filtered estimator $A_{t-1}|y^{t-1} \sim \mathrm{N}\big(A_{t-1|t-1}, P_{t-1|t-1}\big)$, i.e. $P_{t-1|t-1}$, divided by a *forgetting factor* $\lambda \in [0,1]$. That is $P_{t|t-1} = P_{t-1|t-1}/\lambda$.[5] Then, Ω_t is approximated by $(\lambda^{-1} - 1)P_{t-1|t-1}$. From this we can see that $\lambda = 1$ implies a constant coefficient model. Eventually, Σ_t is estimated recursively through an exponential weighted moving average using a decay factor κ between $\hat{\Sigma}_{t-1}$ and the variance-covariance matrix of filtered Kalman residuals, i.e. $\hat{\Sigma}_t = \kappa\hat{\Sigma}_{t-1} + (1-\kappa)\hat{\varepsilon}_t\hat{\varepsilon}_t'$, where $\hat{\varepsilon}_t = y_t - A_{t|t}[1 \;\; y_{t-1}]$ is obtained in the Kalman filter.[6]

The specification of the model involves a set of parameters, namely λ, κ and ϕ, that have to be defined by the prior, either through an hierarchical hyperprior, an empirical Bayes estimator or a search over a grid of possible values. Here, we estimate the model for every parameter

[5] For textbook explanations of the Kalman filtering technique the reader is referred to for example Durbin and Koopman (2012).

[6] The details of the estimation of the model can be found in the Appendix A.1.

combination over a grid and then choose the model with the highest predictive density over the recent past. We also consider an average over all models with different hyperparameter values.

Similar to Koop and Korobilis (2013), the *dynamic model selection and averaging* technique is performed over different priors and not different sets of predictor variables. The idea follows Raftery et al. (2010). In particular, the weights for model j, which comes from the j-th combination of λ, κ and ϕ, at time t using all the information up to $t-1$ are given by

$$\omega_{t|t-1,j} = \omega^{\alpha}_{t-1|t-1,j} \Big/ \sum_{j=1}^{J} \omega^{\alpha}_{t-1|t-1,j}, \text{ and} \tag{12}$$

$$\omega_{t|t,j} = \omega_{t|t-1,j} p_j(y_t|y^{t-1}) \Big/ \sum_{j=1}^{J} \omega_{t|t-1,j} p_j(y_t|y^{t-1}), \tag{13}$$

where $p_j(y_t|y^{t-1})$ is the predictive likelihood of model j evaluated at y_t and $\alpha = 0.99$ is a decay factor governing the weighting of past observations. For monthly data, this value implies that the observations from about two years ago only receive approximately 80 percent of the weight of the observation in $t-1$. We note that dynamic model weights imply a different treatment of every model in each period leading to different averaging results and also may lead to a different forecasting model selection in each period. Following Koop and Korobilis (2013), we perform model averaging across different prior parameter values. That is, $\lambda \in \{0.97, 0.98.0.99, 1\}$, $\kappa \in \{0.94, 0.96.0.98\}$ and $\phi \in \{0, 0.5, 0.75, 1\}$. This results in 48 models based on different model parameters from which we either select the best performing one or average across all of them.[7]

Eventually, we are interested in the marginal predictive distribution of the asset return r_t. This is a main advantage of the Bayesian approach (Klein and Bawa, 1976; Barberis, 2000). The predictive distribution is obtained from the joint predictive density function of r_{t+1} and $\Theta_t = [A_t, \Sigma_t, \Omega_t]$ by integrating over all values of Θ_t. This is

$$f(r_{t+1}|y^t) = \int f(r_{t+1}, \Theta_t|y^t)\, d\Theta_t = \int f(r_{t+1}|y^t, \Theta_t) p(\Theta_t|y^t)\, d\Theta_t, \tag{14}$$

where $y^t = \{y_1, \ldots, y_t\}$ is the collection of all past observations used for estimation. This function is independent of the unknown parameters and is in fact something like the average over all possible values for Θ_t. Numerically, it is obtained by simulating I draws from the posterior

[7]The reader is referred to Koop and Korobilis (2013) for more details about the forecasting set-up and model selection.

distribution and making a prediction $\hat{r}_{t+1}$ for every posterior draw.

3.2 Entropic tilting

In addition to traditional point forecasts, the recent literature has considered probabilistic (or "density") forecasts of macroeconomic and financial variables. In contrast to point forecasts, the latter provide information on various possible scenarios and thus quantify the uncertainty surrounding the future. In the Bayesian methodology, predictive distributions for the variables of interest are easily obtained by integrating out the parameter uncertainty from the likelihood function (evaluated at a future (predicted) realization) times the posterior distribution.

Similar to point forecasts, density forecasts can also be combined to form a merged model that upholds the strengths of each of its components. This can be achieved for example by mixing two densities or by reweighting a forecasted density according to another model; ensuring that the new mixture model is well defined. Entropic tilting is a non-parametric method to combine time-series model forecasts with information from other origins. We now explain the method in more detail.

Suppose at time t we want to make a forecast h periods ahead for a $N \times 1$ vector of interest r_{t+h}, in our case a vector of out-of-sample excess stock returns. Denote by $f_{t,h} := \{r_{t+h,i}\}_{i=1}^{I}$, where $r_{t+h} \in \mathbb{R}^N$ and $N \geq 1$, a baseline sample from the predictive return distribution $p(r_{t+h}|r^t)$, i.e. a discrete sample of I (MCMC) draws of the h-step ahead forecasts. These draws can either come from a closed-from analytical expression of the predictive density $f_{t,h}$ or might be simulated. It also may depend on estimated parameters.

We now want to incorporate additional information about the return r_{t+h}, which was not used to generate the base sample, in the form of M moment conditions on the function $g(r_{t+h}) : \mathbb{R}^N \to \mathbb{R}^M$ in the following sense:

$$\mathbb{E}\left[g(r_{t+h})\right] = \bar{g}_t, \tag{15}$$

where $\bar{g}_t \in \mathbb{R}^M$ and $M, N \geq 1$. For example $g(r_{t+h}) = r_{t+h}$ imposes that the mean of r_{t+h} is equal to $\bar{g}_t$ and $g(r_{t+h}) = \left(r_{t+h} - \mathbb{E}\left(r_{t+h}\right)\right)^2$ sets the variance equal to it. $\bar{g}_t$ can be formed from various origins: Giacomini and Ragusa (2014) use an Euler equation to specify $\bar{g}_t$, Altavilla et al. (2014); Krüger et al. (2015) use survey forecasts and Metaxoglou et al. (2016) adopt option-implied information for $\bar{g}_t$.

In general under the base density $f_{t,h}$, the moments of $g(r_{t+h})$ are not equal to $\bar{g}_t$:

$$\mathbb{E}_{f_{t,h}}\left[g(r_{t+h})\right] = \int g(r_{t+h}) f_{t,h}(r_{t+h})\, dr_{t+h} \neq \bar{g}_t. \tag{16}$$

Instead, entropic tilting describes finding the density $\tilde{f}_{t,h}$ out of the set of densities that fulfill the moment condition in (15) that is closest to the base density in terms of the Kullback-Leibler divergence measure. This is formalized in the following proposition.

Proposition 3.1

If a solution $\tilde{f}_{t,h}(r)$ to the constrained minimization

$$\min_{\tilde{f}_{t,h}\in\mathcal{F}} \mathbb{E}_{\tilde{f}_{t,h}}\left[\log \frac{\tilde{f}_{t,h}(r)}{f_{t,h}(r)}\right] = \int \log \frac{\tilde{f}_{t,h}(r)}{f_{t,h}(r)} \tilde{f}_{t,h}(r)\, dr, \tag{17}$$

$$\text{s.t. } \mathbb{E}_{\tilde{f}_{t,h}}\left[g(r)\right] = \int g(r)\tilde{f}_{t,h}(r)\, dr = \bar{g}_t, \tag{18}$$

exists, then it is unique and it is given by

$$\tilde{f}^*_{t,h}(r) = f_{t,h}(r)\exp\left(\gamma^{*\prime}_{t,h} g(r)\right) \Big/ \int \exp\left(\gamma^{*\prime}_{t,h} g(r)\right) f_{t,h}(r)\, dr, \tag{19}$$

$$\gamma^*_{t,h} = \arg\min_{\gamma_{t,h}} \int f_{t,h}(r)\exp\left(\gamma'_{t,h}(g(r)-\bar{g}_t)\right) dr. \tag{20}$$

Proof. The proof is given in Giacomini and Ragusa (2014) Proposition 1 on page 147. □

While $\tilde{f}^*_{t,h}(r)$ is generally not of a known form, the entropic tilting problem can also be interpreted as finding a new sets of weights $\pi^*_{t,h}$ in t for the base h-step ahead density $f_{t,h}(r)$ that satisfy the moment condition. For a sample of I draws from the base predictive density, the expectation in (17) is

$$\mathbb{E}_{\tilde{f}_{t,h}}\left[\log \frac{\tilde{f}_{t,h}(r)}{f_{t,h}(r)}\right] = \sum_{i=1}^{I} \tilde{\pi}_i \log\left(\frac{\tilde{\pi}_i}{\pi_i}\right) \overset{\pi_i=1/I}{=} \log I + \sum_{i=1}^{I} \tilde{\pi}_i \log\left(\tilde{\pi}_i\right), \tag{21}$$

where π_i, $i=1,\ldots,I$, are the original weights for the base density usually equal to $1/I$. Following Robertson et al. (2005), imposing the condition (18) via $\mathbb{E}_{\tilde{f}_{t,h}}\left[g(r)\right] = \sum_{i=1}^{I}\tilde{\pi}_i g(r_{t,i})$ yields the tilting solution from (19) and (20) as

$$\pi_i^* = \frac{\exp\left(\gamma_{t,h}^{*\prime} g(r_{t+h,i})\right)}{\sum_{i=1}^{I} \exp\left(\gamma_{t,h}^{*\prime} g(r_{t+h,i})\right)}, \tag{22}$$

$$\gamma_{t,h}^* = \arg\min_{\gamma_{t,h}} \sum_{i=1}^{I} \exp\left(\gamma_{t,h}^{\prime}(g(r_{t+h,i}) - \bar{g}_t)\right). \tag{23}$$

Equation (22) ensures that all elements of the new weight vector $\pi_{t,h}^*$ are positive and sum up to one. $\gamma_{t,h}^*$ in (23) has dimension M (the number of moment conditions) and can easily be found by a Lagrangian optimization.

The moment condition in (15) restricts the set of possible candidate densities. Hence, the usefulness or uncertainty about the additional information is not measured. Moreover, the more moment conditions exists, the smaller is the set of candidate distributions.

Entropic tilting also has a shrinkage interpretation (Robertson et al., 2005, p. 394): Given a certain mean condition on the target random variable, imposing higher moment conditions that *shrink* the variance of the target variable to zero, sets the mean automatically to the imposed target mean. In other words, $\bar{g}_t$ can be interpreted as a shrinkage target for the entire predictive return distribution, achieved through re-weighting every single draw of it and that changes its moments. This can also be seen from considering the following example: Let y follow a bivariate normal distribution with $f(y) = N(\theta, \Sigma)$ and impose the restriction that the mean of the second variable y_2 is μ_2 and its variance is Ω_{22}. Then it follows that the tilted distribution is also normal $\tilde{f}^*(y) = N(\mu, \Omega)$ and the mean of y_1 is given by

$$\mu_1 = \theta_1 + \Sigma_{22}^{-1}\Sigma_{12}(\mu_2 - \theta_2) = \lambda\theta_1 + (1-\lambda)\underbrace{\left(\theta_1 + \frac{\Sigma_{22}^{-1}\Sigma_{12}(\mu_2 - \theta_2)}{1-\lambda}\right)}_{=:\tilde{\theta}_1}. \tag{24}$$

Here, $\tilde{\theta}_1$ is the shrinkage target that depends implicitly on the distance between the mean condition for second variable μ_2 and the true mean θ_2.

Entropic tilting only changes the location and shape of predictive return distribution, but it does not foster better parameter estimates for the underlying prediction model. It is therefore not equivalent to putting an informative prior on the mean and variance of the prediction model centered at the analysts' forecasts. Also, it does not imply a structural relationship between the asset returns and the analysts' forecasts as the approach of Frey and Mokinski (2016), who augment a Bayesian VAR system by survey nowcasts and impose parameter restrictions between the original variables in the VAR and the added equations for the nowcasts through the prior.

4 Empirical Application

4.1 Data and set-up

All data used in the empirical exercise was obtained from Thomson Reuters Datastream. We will investigate the forecast performance of various prediction models for 20 Dow Jones constituents for which the Institutional Brokers Estimate System (I/B/E/S) database provide target prices and analyst recommendations. We use monthly observations because the I/B/E/S summary data is aggregated on this frequency. Historical target price data from I/B/E/S is available from April 1999 to October 2014. The initial estimation window has size $h = 60$ and hence pseudo out-of-sample evaluation period starts in April 2004. We only consider one-step ahead predictions in this study. For each stock, we compute logarithmic returns (including dividends) and subtract the 3-month T-bill rate to obtain excess returns. For the choice of the predictor variables, we partly follow Welch and Goyal (2008); Pettenuzzo and Ravazzolo (2016) and (i) consider firm specific *fundamentals* such as the log dividend yield, the log earnings price ratio, the log dividend-payout ratio, and the book-to-market ratio, (ii) market and economic measures such as the 3-month T-bill rate, the yield on long-term government bonds, the market excess return and CPI inflation.

Thomson Reuters I/B/E/S database provides detailed and consensus estimates featuring up 26 forecast measures for more than 70,000 companies in more than 90 countries worldwide. While, the summary files contain a monthly snapshot of each company followed by sell-side analysts whose brokerage firm provides data to I/B/E/S, the detail files offer forecasts from individual analysts. We are mainly interested in two variables:

1. **Price targets:** The mean and standard deviation of the projected price level forecasted by professional analysts with a 12-month time horizons. At each point in time, we also consider the entire vector of target prices from individual analysts.

2. **Recommendations summary:** The mean and standard deviation of analysts' recom-

mendation based on a five point standardized scale (strong buy = 1, buy, hold, sell, strong sell = 5) as well as the total number of recommendations, the number of up- and downgrade revisions and the percentage of buy, hold and sell recommendations.

We will use the target price returns and variance as well as the consensus analyst recommendations and their revisions as predictor variables. In particular, target prices will be used to calculate (i) monthly forward target price implied expected return, i.e. simple returns between the spot and the twelve months forward target price at each point t divided by twelve, and (ii) monthly target price implied expected return variances. For this, we used the detail history I/B/E/S files that contain target prices of individual analysts. We first calculate monthly forward target price implied expected returns for every individual analyst and then use the mean and variance of these returns as first and second moment restrictions for the entropic tilting exercise.

Table 2 reports the descriptive statistics on the returns, expected target returns and recommendations for the 20 Dow Jones constituents used in this study. While we note that the mean and standard deviations of the logarithmic returns differ substantially across assets, the target price and recommendation characteristics are very similar: All mean forward target price implied expected returns are positive, indicating an upward bias in the target prices compared to the spot prices. The standard deviation of the expected target price returns is generally slightly lower in scale than the standard deviation of the logarithmic stock returns. Moreover, the upward bias can also be seen in the mean recommendations, which is always smaller than three, given the five point scale. A lower score indicates more buy recommendations. The company with the least number target prices of 11 and number of recommendations with 16 is DuPont (DD) and the company with the greatest number of target prices, 28, and recommendations, 39, is Intel (INTC).

4.2 Competing models

We consider a number of different models to distinguish the effects, considering different sets of predictor variables, models with constant and time-varying coefficients as well as mean and variance tilted models. These are

1. [AR1] Autoregressive model of order one for the return process of each asset.

2. [VAR-Full] Bayesian vector autoregressive model of order one with an uninformative prior

for all parameters (full specification).

3. [VAR-Minnesota] Bayesian vector autoregressive model of order one with the Minnesota prior given in (5) - (9) to impose the full model for the return and an autoregressive model for all other variables.
4. [TVPVAR-SV-DMA] Time-varying parameter model with stochastic volatility and using forgetting factors and dynamic model averaging over different prior parameters as described in Section 3.1.1.
5. [TVPVAR-SV-DMS] Time-varying parameter model with stochastic volatility and using forgetting factors and dynamic model selection over different prior parameters as described in Section 3.1.1.
6. [TVPVAR-SV-DMAm] TVP-BVAR with SV using dynamic model averaging with mean tilting using the monthly target price implied expected returns.
7. [TVPVAR-SV-DMAm/v] TVP-BVAR with SV using dynamic model averaging with mean and variance tilting using the monthly target price implied expected returns.
8. [TVPVAR-SV-DMSm] TVP-BVAR with SV using dynamic model selection with mean tilting using the monthly target price implied expected returns.
9. [TVPVAR-SV-DMSm/v] TVP-BVAR with SV using dynamic model selection with mean and variance tilting using the monthly target price implied expected returns.
10. [Bayesian lasso] The Bayesian lasso (Park and Casella, 2008) is a shrinkage for univariate regressions based on the Laplace prior that can be used to impose an L_1-norm penalization on the regression coefficients to shrink them to zero. It has been shown to be a strong forecasting device (Korobilis, 2013) and is used here as a benchmark model.

4.3 Evaluation criteria

We consider two evaluation criteria in this study. To evaluate the entire predictive accuracy of point forecasts we use the out-of-sample R^2. For model j, it is given by

$$\mathrm{R}^2_{OoS,j} = 1 - \frac{\sum_{i=h+1}^{T} e_{j,i}^2}{\sum_{i=h+1}^{T} e_{0,i}^2}, \tag{25}$$

where $e_{0,i} = y_i^1 - \hat{y}_{0,i}^1 = r_i - \hat{r}_{0,i}$ denotes the forecast error of a simple mean or intercept only model ($r_i = a + \varepsilon_i$) and $e_{j,i}$ the forecast error in the returns of model j at time i and h denotes the end-point of the initial estimation period. Note that we only evaluate the forecast error with respect to the asset return and not with regard to all predictor variables. All errors are obtained by averaging over the (marginal) predictive density function of the asset returns. Values above zero indicate that model j produces lower forecasts error than the intercept only model. Second, to evaluate the predictive distribution, we consider the average log score differential (LSD) given by

$$\mathrm{LSD}_{j,t} = \frac{\sum_{i=h+1}^{t} (\mathrm{LS}_{j,i} - \mathrm{LS}_{0,i})}{\sum_{i=h+1}^{t} \mathrm{LS}_{0,i}}, \tag{26}$$

where where $\mathrm{LS}_{j,i}$ is log predictive score of model j at time i. Again, values above zero indicate that a given model j shows better forecast performance than the benchmark model, while negative values suggest the opposite.

4.4 Individual predictor performance

Before heading to the main analysis using all predictors, we start by looking at models with only a single predictor in order to investigate which ones have particular forecasting power. For this, we apply a constant parameter model and a time-varying model with stochastic volatility to estimate the system given in (1) - (2).[8] Tables 3 to 6 provide the results for the out-of-sample R^2 and Tables 7 to 10 provide the results in terms of the log predictive scores. Bold numbers show positive performance measures, indicating that the one predictor model outperforms the simple intercept only model. We test statistical significance using the Diebold and Mariano (1995) t-tests with a null hypothesis of equal average forecasting ability.

The results can be summarized as follows: (i) No predictor variable shows significant forecast performance using a constant parameter model (Tables 3 and 7). (ii) Using the time-varying parameter model with stochastic volatility increases forecast performance for all predictors, most profoundly for the general economic and market indicators (Tables 4 and 8). This may be due to the model flexibility which can reflect the nature of asset returns (e.g. heteroskedasticity) much better (see also Johannes et al., 2014). (iii) Tilting the predictive distribution from the TVP-BVAR towards the target price implied expected returns does not increase forecast

[8] For the latter, the underlying model parameters are set to $\lambda = 0.99$, $\kappa = 0.96$ and $\phi = 0.5$.

performance (Tables 5 and 9) significantly. (iv) However, tilting the predictive distribution from the TVP-BVAR to posses the mean and variance of the target price implied expected returns does indeed increase forecast performance, but only significantly for a couple of assets and predictors, especially for the models using the target price return and variance as the predictor variables (Tables 6 and 10).

To illustrate the effect of the tilting, we plot the baseline and the tilted predictive densities for the IBM stock returns at different times. Figure 3 compares the baselines density against the mean tilted density for the TVPVAR(1) model. It is obvious that the baseline and tilted densities are very similar to each other, only in 2010 and 2012 the mode of tilted distribution seems to be closer to the actual outcome. The similarity in the two distributions thus explains the lack of forecast improvements from mean tilting. The picture changes when looking at the tilted density of the mean and variance tilted models. In Figure 4, we see that the tilted distribution is much more concentrated around the actual outcome in calm market times (2006, 2012) and more flat and has fatter tails than the baseline density in crises times between 2008 and 2010. In the entire sample, the tilted densities are more often stronger concentrated around the true observation than the baseline. This concentration in the density comes from the agreement of the analyst about the future target price and does not imply that the analysts make unbiased forecasts.

4.5 Complete model performance

None of the individual predictor models outperformed the intercept only model consistently across all assets. Instead, the individual predictor exercise revealed that especially the time-varying parameters and the stochastic volatility in connecting with the mean and variance tilting can produce significantly better forecasts. Therefore, we now put all predictors in the model and perform dynamic model averaging and selection using the full VAR system for the 13 variables (12 predictors plus the asset return). This leads to the main results of this paper which can be found in Tables 11 and 12. While the AR(1) model cannot outperform the mean model, the full BVAR system with uninformative priors clearly gives worst forecasts. This deterioration is likely to reflect over-fitting from the great number of estimable parameters. The Minnesota prior (shrinking many parameters to zero here) instead improves forecast performance, although not significantly. Its performance is similar to the Bayesian lasso. The time-varying models with dynamic model averaging and selection clearly improve the forecast performance further. However, the difference between model averaging and selection is only marginal. Again, tilting the mean of the predictive distribution towards the target price implied return does not improve the forecast further. This is only achieved when also tilting the variance towards the target price implied return variance, giving significant better forecasts for a wide range of assets.

5 Concluding Remarks

In this paper we demonstrate that financial analyst forecasts can have predictive power for equity assets of the Dow Jones Industrial index using a novel entropic tilting approach to combine time-series forecasts with analysts' information. We find that the extent of predictability varies across assets and that using Bayesian vector autoregressions with time-varying coefficients, stochastic volatility and model averaging and selection among priors improves return predictions across all assets. While tilting the mean of the predictive distribution towards the target price implied expected returns did not improve forecast performance significantly, tilting the variance of the predictive return distribution towards the implied expected target return variance produced forecasts outperforming a simple intercept only model. This may be explained by the fact that the tilted densities are more often stronger concentrated around the true outcome than the baseline density. In other words, the agreement among analysts reduces the predictive variance of the asset returns. Contrary, the disagreement among the analysts may be an indicator for future market uncertainties.

Using entropic tilting has several advantages: It can incorporate any kind of (forward-looking) information into a predictive regression framework in a parsimonious way without increasing estimation noise. Hence, it is a regularization approach that is suitable in high-dimensional settings, even portfolio problems. Notably, by using the analyst information in a tilting framework we only change the simulated draws of the predictive distribution and hence we do not require the data to formalize and estimate the true relationship between asset returns and analyst target prices. However, while in this way the dimensionality of the forecasting model, in this case of a Bayesian vector autoregression, is unchanged, we do not account for the stochastic nature of the tilting information.

This opens the door for possible extensions: While one could develop a tilting framework that not only considers the set of predictive distributions that strictly fulfill the moment conditions, one could try to search among all possible predictive distributions and then minimize the distance of the target moments given a statistic or economic criteria. Second, one could also apply the tilting framework to predictive portfolio weight regressions. For example, the framework of (Frey and Pohlmeier, 2015) may be used to incorporate external knowledge directly about the portfolio weights, i.e. a certain target allocation, instead of tilting the underlying return process. Finally, the tilting approach is also applicable for panel VAR systems that model various assets simultaneously. Tilting the joint predictive return distributions may then be used in a portfolio allocation problem of an expected utility maximizing investor.

References

Altavilla, C., R. Giacomini, and G. Ragusa (2014): "Anchoring the Yield Curve Using Survey Expectations," ECB Working Paper No. 1632, ECB.

Avramov, D. (2002): "Stock return predictability and model uncertainty," *Journal of Financial Economics*, 64, 423–458.

Barber, B., R. Lehavy, M. McNichols, and B. Trueman (2001): "Can Investors Profit from the Prophets? Security Analyst Recommendations and Stock Returns," *The Journal of Finance*, 56, 531–563.

Barberis, N. (2000): "Investing for the Long Run when Returns are Predictable," *The Journal of Finance*, 55, 225–264.

Black, F. and R. Litterman (1992): "Global Portfolio Optimization," *Financial Analysts Journal*, 48, 28–43.

Bonini, S., L. Zanetti, R. Bianchini, and A. Salvi (2010): "Target Price Accuracy in Equity Research," *Journal of Business Finance & Accounting*, 37, 1177–1217.

Bradshaw, M. T., L. D. Brown, and K. Huang (2012): "Do Sell-Side Analysts Exhibit Differential Target Price Forecasting Ability?" *Review of Accounting Studies*, 18, 930–955.

Brav, A. and R. Lehavy (2003): "An Empirical Analysis of Analysts' Target Prices: Short-Term Informativeness and Long-Term Dynamics," *The Journal of Finance*, 58, 1933–1967.

Carriero, A., T. E. Clark, and M. Marcellino (2015): "Bayesian VARs: Specification Choices and Forecast Accuracy," *Journal of Applied Econometrics*, 30, 46–73.

Cremers, K. (2002): "Stock return predictability: a Bayesian model selection perspective," *The Review of Financial Studies*, 15, 1223–1249.

Cvitanić, J., A. Lazrak, L. Martellini, and F. Zapatero (2006): "Dynamic Portfolio Choice with Parameter Uncertainty and the Economic Value of Analysts' Recommendations," *Review of Financial Studies*, 19, 1113–1156.

Da, Z. and E. Schaumburg (2011): "Relative Valuation and Analyst Target Price Forecasts," *Journal of Financial Markets*, 14, 161–192.

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

Diebold, F. X. and R. S. Mariano (1995): "Comparing Predictive Accuracy," *Journal of*

Business & Economic Statistics, 20, 134–144.

Doan, T., R. Litterman, and C. Sims (1984): "Forecasting and Conditional Projection using Realistic Prior Distributions," *Econometric Reviews*, 3, 1–100.

Durbin, J. and S. Koopman (2012): *Time Series Analysis by State Space Methods*, Oxford Statistical Science, Oxford University Press, 2 ed.

Feng, G. and N. G. Polson (2016): "Regularizing Bayesian Predictive Regressions," *arXiv:1606.01701 [stat]*.

Frey, C. and F. Mokinski (2016): "Forecasting with Bayesian Vector Autoregressions Estimated Using Professional Forecasts," *Journal of Applied Econometrics*, 31, 1083–1099.

Frey, C. and W. Pohlmeier (2015): "Bayesian Regularization of Portfolio Weights," Tech. rep., Department of Economics, University of Konstanz, http://ssrn.com/abstract=2730475.

Giacomini, R. and G. Ragusa (2014): "Theory-Coherent Forecasting," *Journal of Econometrics*, 182, 145–155.

Giacomini, R. and H. White (2006): "Tests of Conditional Predictive Ability," *Econometrica*, 74, 1545–1578.

Green, T. C. (2006): "The Value of Client Access to Analyst Recommendations," *Journal of Financial and Quantitative Analysis*, 41, 1–24.

He, P. W., A. Grant, and J. Fabre (2013): "Economic value of analyst recommendations in Australia: an application of the Black–Litterman asset allocation model," *Accounting & Finance*, 53, 441–470.

Irvine, P. J. (2003): "The Incremental Impact of Analyst Initiation of Coverage," *Journal of Corporate Finance*, 9, 431–451.

Jegadeesh, N., J. Kim, S. D. Krische, and C. M. C. Lee (2004): "Analyzing the Analysts: When Do Recommendations Add Value?" *The Journal of Finance*, 59, 1083–1124.

Jegadeesh, N. and W. Kim (2006): "Value of Analyst Recommendations: International Evidence," *Journal of Financial Markets*, 9, 274–309.

Johannes, M., A. Korteweg, and N. Polson (2014): "Sequential Learning, Predictability, and Optimal Portfolio Returns," *The Journal of Finance*, 69, 611–644.

Kandel, S. and R. Stambaugh (1996): "On the Predictability of Stock Returns: An Asset-Allocation Perspective," *Journal of Finance*, 51, 385–424.

Kandel, S. and R. F. Stambaugh (1987): "Long-horizon returns and short-horizon models," CRSP Working paper 222, University of Chicago.

Klein, R. and V. Bawa (1976): "The Effect of Estimation Risk on Optimal Portfolio Choice," *Journal of Financial Economics*, 3, 215–231.

Koop, G. and D. Korobilis (2010): "Bayesian Multivariate Time Series Methods for Empirical Macroeconomics," *Foundations and Trends® in Econometrics*, 3, 267–358.

——— (2013): "Large time-varying parameter VARs," *Journal of Econometrics*, 177, 185–198.

Korobilis, D. (2013): "Hierarchical shrinkage priors for dynamic regressions with many predictors," *International Journal of Forecasting*, 29, 43–59.

Krüger, F., T. E. Clark, and F. Ravazzolo (2015): "Using Entropic Tilting to Combine BVAR Forecasts with External Nowcasts," *Journal of Business & Economic Statistics*, 1–48.

Lin, S., H. Tan, and Z. Zhang (2016): "When Analysts Talk, Do Institutional Investors Listen? Evidence from Target Price Changes," *Financial Review*, 51, 191–223.

Metaxoglou, K., D. Pettenuzzo, and A. Smith (2016): "Option-Implied Equity Premium Predictions via Entropic Tilting," *Available at SSRN 2754246.*

Park, T. and G. Casella (2008): "The Bayesian Lasso," *Journal of the American Statistical Association*, 103, 681–686.

Pástor, Ľ. (2000): "Portfolio Selection and Asset Pricing Models," *The Journal of Finance*, 55, 179–223.

Pettenuzzo, D. and F. Ravazzolo (2016): "Optimal Portfolio Choice Under Decision-Based Model Combinations," *Journal of Applied Econometrics*, forthcoming.

Pettenuzzo, D. and A. Timmermann (2011): "Predictability of stock returns and asset allocation under structural breaks," *Journal of Econometrics*, 164, 60–78.

Raftery, A. E., M. Kárný, and P. Ettler (2010): "Online Prediction Under Model Uncertainty via Dynamic Model Averaging: Application to a Cold Rolling Mill," *Technometrics*, 52, 52–66.

Ramnath, S., S. Rock, and P. Shane (2008): "The financial analyst forecasting literature: A taxonomy with suggestions for further research," *International Journal of Forecasting*, 24, 34–75.

Rapach, D., J. Strauss, J. Tu, and G. Zhou (2015): "Industry interdependencies and cross-industry return predictability," *Available at SSRN 2566541*.

RAPACH, D. AND G. ZHOU (2013): "Chapter 6 - Forecasting Stock Returns," Elsevier, vol. 2, Part A of *Handbook of Economic Forecasting*, 328–383.

ROBERTSON, J. C., E. W. TALLMAN, AND C. H. WHITEMAN (2005): "Forecasting Using Relative Entropy," *Journal of Money, Credit and Banking*, 37, 383–401.

WACHTER, J. A. AND M. WARUSAWITHARANA (2015): "What is the chance that the equity premium varies over time? Evidence from regressions on the dividend-price ratio," *Journal of Econometrics*, 186, 74–93.

WELCH, I. AND A. GOYAL (2008): "A Comprehensive Look at The Empirical Performance of Equity Premium Prediction," *Review of Financial Studies*, 21, 1455–1508.

ZELLNER, A. AND V. CHETTY (1965): "Prediction and Decision Problems in Regression Models from the Bayesian Point of View," *Journal of the American Statistical Association*, 60, 608–616.

A Appendix

A.1 Estimation of the TVP-BVAR with SV using forgetting factors

We consider the model

$$y_t = a_t + A_{1,t}\, y_{t-1} + \varepsilon_t, \tag{27}$$

$$A_t = \phi A_{t-1} + (1-\phi)\underline{A}_0 + u_t, \tag{28}$$

where $A_t = [a_t \;\; A_{1,t}]$ is an unknown state vector, $\underline{A}_0$ is some initial condition for each t and also $A_0 = \underline{A}_0$, $\varepsilon_t \overset{iid}{\sim} \mathrm{N}(0, \Sigma_t)$ with initial condition Σ_0, $u_t \overset{iid}{\sim} \mathrm{N}(0, \Omega_t)$ with initial condition Ω_0 and ε_t and u_s are independent of each other for all t and s. To estimate the mode, we run a Kalman filter for $t = 1, \ldots, T$ as follows:[9]

I. Prediction step:

1. Set $A_{t|t-1} = \phi A_{t-1|t-1} + (1-\phi)\underline{A}_0$.
2. Set $P_{t|t-1} = \frac{1}{\lambda} P_{t-1|t-1}$
   where for $t = 1$ we set $A_{0|0} = \underline{A}_0$ and $P_{0|0} = \underline{P}_0$.

II. Update step:

1. Calculate $\tilde{\varepsilon}_t = y_t - a_{t|t-1} + A_{t|t-1}\, y_{t-1}$.
2. Calculate $\hat{\Sigma}_t = \kappa\hat{\Sigma}_{t-1} + (1-\kappa)\tilde{\varepsilon}_t'\tilde{\varepsilon}_t$ with $\hat{\Sigma}_1 = \kappa\Sigma_0$.
3. Estimate $A_{t|t} = A_{t|t-1} + P_{t|t-1}[1 y_{t-1}]' \left(\hat{\Sigma}_t + [1 y_{t-1}]P_{t|t-1}[1 y_{t-1}]'\right)^{-1} \tilde{\varepsilon}_t$.
4. Calculate $P_{t|t} = P_{t|t-1} + P_{t|t-1}[1 y_{t-1}]' \left(\hat{\Sigma}_t + [1 y_{t-1}]P_{t|t-1}[1 y_{t-1}]'\right)^{-1} P_{t|t-1}$.

The one-step ahead predictive density of the VAR model is then analytically available from the Kalman filter as

$$p(y_t|y^t) \sim \mathrm{N}\Big([1 \;\, y_{t+1}]A_{t+1|t}, \hat{\Sigma}_{t+1} + [1 \;\, y_{t+1}]A_{t+1|t}[1 \;\, y_{t+1}]'\Big)\,. \tag{29}$$

[9] The algorithm is taken and amended from the technical appendix of the working paper version of Koop and Korobilis (2013).

A.2 Figures

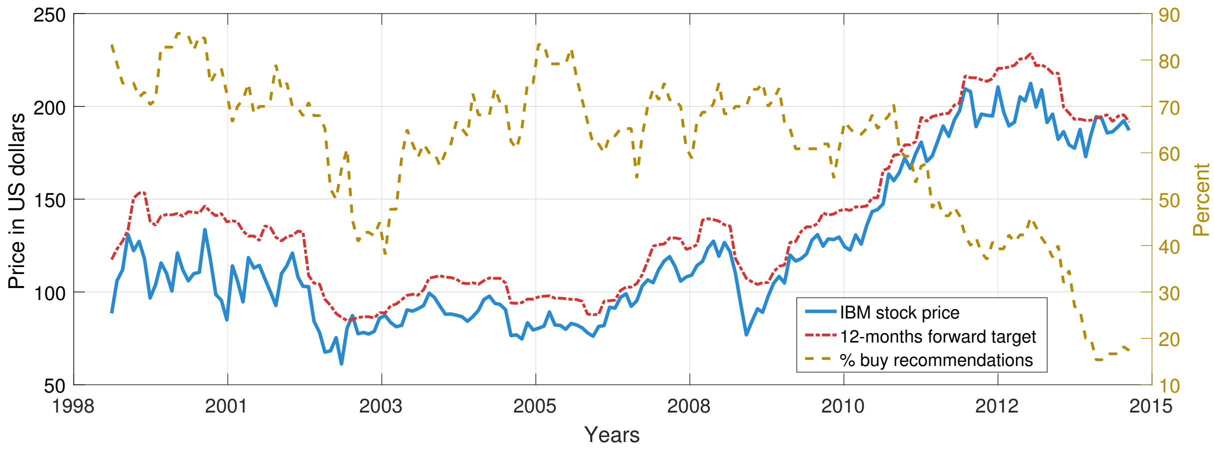

Figure 1: The figure shows IBM spot price, the mean 12 months forward target price and the percentage of buy recommendations from all recommendations (buy, sell, hold) of the IBM stock between 1999 and 2015.

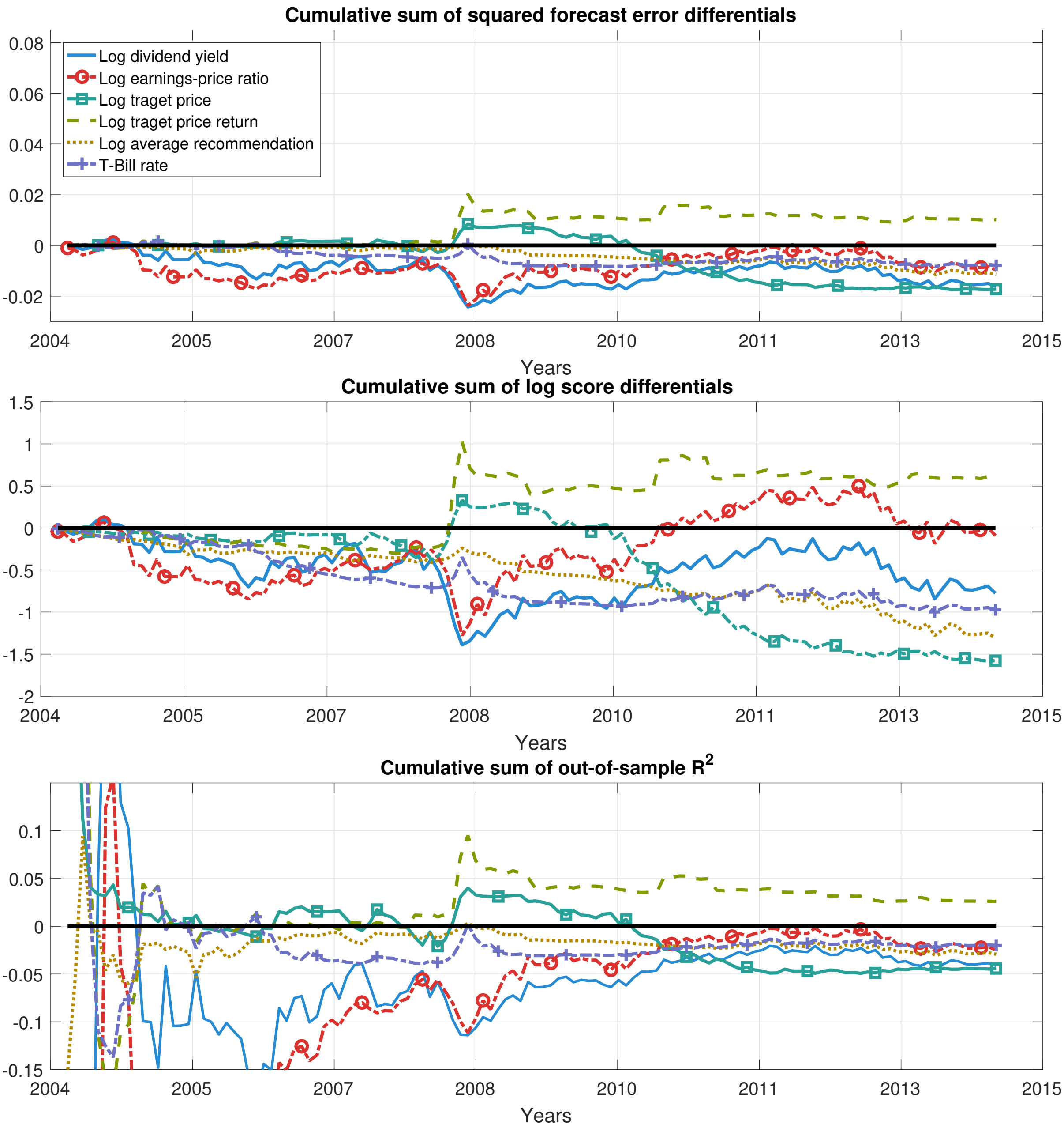

Figure 2: The figure provides out-of-sample forecast performance results for different univariate models for the IBM stock for 2004 to 2015. The top panel shows the cumulative sum of squared forecast errors of the benchmark mean model model minus the sum of squared forecast errors for six univariate models with different regressors, i.e. for model m this is $\text{CSSED}_{m,t} = \sum_{i=S+1}^{t} \left(e_{0,i}^2 - e_{m,i}^2 \right)$. Each model is estimated from a linear regression of monthly excess returns on an intercept and a lagged predictor variable, i.e. $r_t = \alpha + \beta x_{t-1} + \varepsilon_t$. The middle panel shows the cumulative sum of log predictive scores of the six models minus the sum of log predictive scores of the benchmark mean model. The bottom panel shows the cumulative sum of out-of-sample R^2 values of each of the six univariate models. For all three panels it holds that values above zero indicate that a given predictor has better forecast performance than the benchmark model, while negative values suggest the opposite.

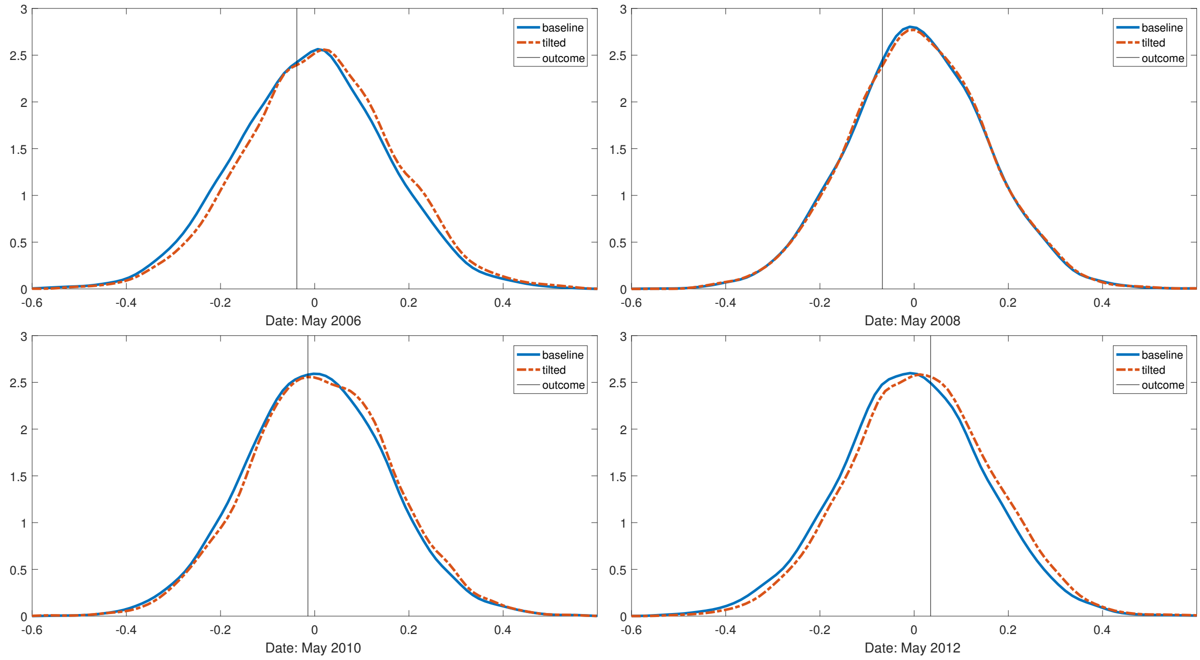

Figure 3: The figure shows the kernel of the predictive density of the IBM returns from the TVPVAR(1) model with dynamic model averaging and mean tilting towards the target price implied expected return at different times. The black horizontal line indicates the actual outcome return)

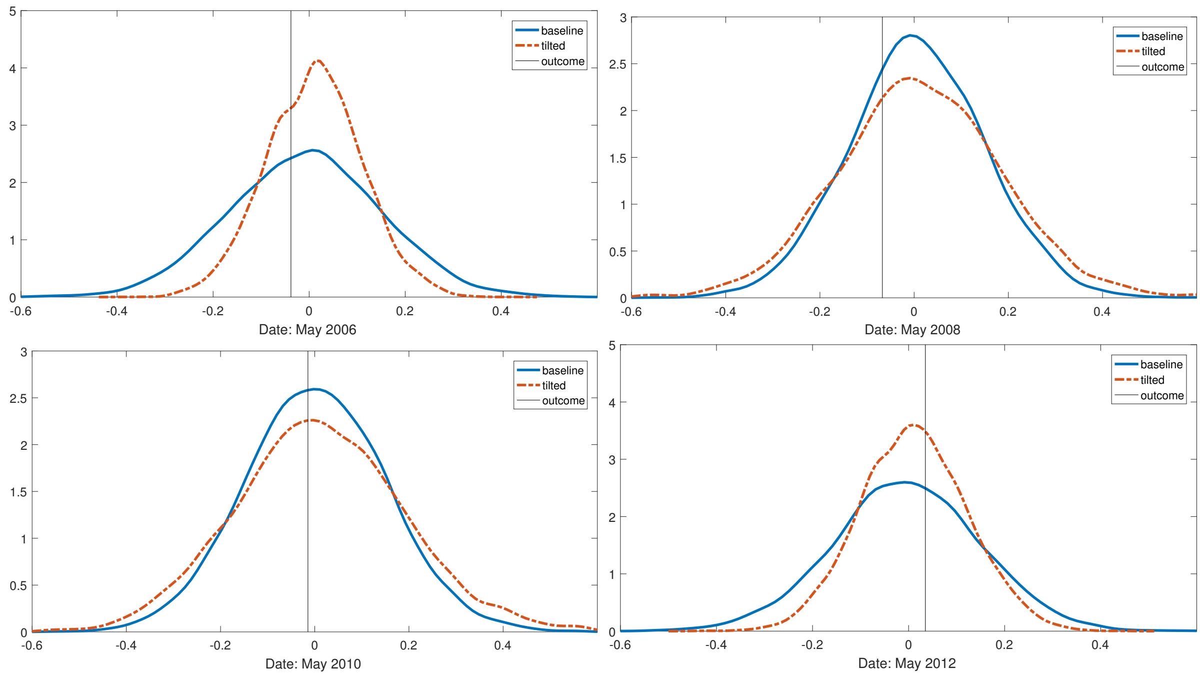

Figure 4: The figure shows the kernel of the predictive density of the IBM returns from the TVPVAR(1) model with dynamic model averaging with tilting towards the mean and variance of the target price implied expected returns at different times. The black horizontal line indicates the actual outcome return)

A.3 Tables

Table 1: Relative root mean squared errors between forecasted and observed spot prices for 20 Dow Jones constituents (sample: 1999 - 2015)

Stock	AA	AAPL	AIG	AXP	BA	CAT	KO	DD	GE	HD
rRMSFE	1.33**	**0.74**	1.45***	**0.93**	**0.87**	**0.96**	1.09	1.24**	1.23	**0.83**
Stock	INTC	IBM	JNJ	MCD	MRK	MSFT	PG	UTX	WMT	DIS
rRMSFE	1.36	**0.94**	**0.85**	**0.69*****	1.25*	1.65***	1.01	**0.82**	1.24	**0.84**

Note: The table displays relative root mean squared errors between observed spot price twelve months ahead and the mean 12-month forward target price as well as the two year historical average for 20 Dow Jones constituents between 1999 and 2015. Values lower than one indicate that the target price generates superior forecast performance. For each stock, we test whether the target price forecast has lower MSFE than the average price forecast by the test proposed by Giacomini and White (2006). One/two/three asterisks denote rejection of the null hypothesis of equal predictive ability at the ten/five/one percent test level.

Table 2: Descriptive statistics on the returns, target prices and recommendations for 20 Dow Jones constituents (sample: 1999 - 2015)

Stock	AA	AAPL	AIG	AXP	BA	CAT	KO	DD	GE	HD
Mean log ret	-0.52	2.01	-1.79	0.26	0.41	0.43	-0.05	-0.18	-0.35	0.27
Std log return	12.00	14.13	21.28	9.52	8.95	10.13	6.03	8.27	8.74	8.22
# price tragets	14.38	25.19	13.17	16.61	16.18	13.99	13.11	11.81	13.33	18.39
Mean exp ret	1.44	1.44	2.52	1.12	1.03	1.08	0.95	1.30	1.31	1.10
Std exp ret	10.72	7.49	25.95	6.08	7.04	7.20	4.57	5.79	6.07	6.37
# RECs	18.40	32.33	20.19	21.36	23.18	20.09	17.73	16.48	17.07	25.76
Mean RECs	2.35	2.13	2.19	2.36	2.26	2.29	2.11	2.40	2.02	2.11
Std RECs	0.38	0.38	0.58	0.35	0.38	0.25	0.25	0.28	0.34	0.25

Stock	INTC	IBM	JNJ	MCD	MRK	MSFT	PG	UTX	WMT	DIS
Mean log ret	-0.13	0.14	0.24	0.27	-0.24	-0.09	0.14	0.39	0.10	0.40
Std log return	11.57	7.76	5.26	6.54	7.88	8.90	5.78	7.25	5.76	7.91
# price tragets	28.66	17.20	14.66	14.92	15.53	23.65	13.34	14.89	17.98	20.19
Mean exp ret	1.34	0.95	0.71	1.10	0.97	1.53	0.86	0.95	1.06	1.26
Std exp ret	7.91	5.29	3.68	5.40	5.91	6.91	3.80	4.52	4.16	6.62
# RECs	39.53	23.10	23.78	20.60	24.28	32.73	18.77	19.98	26.02	27.22
Mean RECs	2.18	2.18	2.10	2.16	2.42	1.91	2.09	1.95	2.05	2.29
Std RECs	0.29	0.26	0.25	0.26	0.36	0.26	0.21	0.24	0.25	0.25

Note: The table reports descriptive statistics on the returns, expected target returns and recommendations for 20 Dow Jones constituents. It reports the mean and standard deviation of the logarithmic monthly returns, the mean number of available target prices, the mean and variance of the monthly forward target price implied expected return, i.e. simple returns between the spot and the twelve month forward target price at each point t divided by 12, constructed from individual analyst data, the mean number of recommendations as well as the mean and standard deviation of the recommendations based on the 1 (strong buy) to 5 (strong sell) scale. Mean returns and standard deviations are multiplied by 100. Target prices and recommendations are obtained from I/B/E/S Datastream.

Table 3: Forecast performance in terms of out-of-sample R^2 for 20 Dow Jones constituents (sample: 2004 - 2015) using a Bayesian VAR(1)

Stock	AA	AAPL	AIG	AXP	BA	CAT	KO	DD	GE	HD
Log DY	-6.90	**0.20**	-3.46	-0.68	-1.24	-0.51	-0.04	-0.20	-0.44	-0.13
Log EPR	-1.97	**0.13**	-0.82	-0.30	-5.98	-0.06	-0.22	-0.16	-0.32	-0.62
Log DPR	-0.13	-0.04	-0.05	-0.09	-1.09	**0.13**	-0.10	**0.25**	-0.13	-0.18
BMR	-0.15	**0.21**	-1.18	-0.14	-0.13	-0.13	-0.73	-0.24	**0.00**	-0.06
3M Tbill rate	-0.10	**0.09**	-0.07	**0.33**	**0.06**	-0.18	-0.36	-0.28	-0.11	-0.17
Market return	-0.09	**0.22**	-0.02	-0.06	-0.05	-0.12	-0.09	-0.21	-0.11	-0.57
LT yield	-0.22	**0.03**	**0.24**	-0.15	-0.35	-0.27	-0.42	-0.28	-0.09	-0.29
CPI inflation	-0.15	**0.18**	-0.10	-0.14	-0.13	-0.19	-0.30	-0.19	-0.20	**0.14**
Log TPR	-0.05	-0.16	-0.58	-0.45	-0.10	-0.13	-0.30	-0.17	-0.50	-0.20
Log TPV	-0.24	**0.15**	-0.12	-0.17	-0.12	-0.53	-0.16	-0.19	-0.28	**0.28**
Log REC	-0.54	**0.01**	-0.08	-0.48	-1.10	-0.31	-0.52	-0.72	-0.22	-0.08
Log REC return	**8.06*****	-0.36	-0.09	**0.09**	-0.32	-0.01	**8.06*****	**8.06*****	-0.11	-0.02

Stock	INTC	IBM	JNJ	MCD	MRK	MSFT	PG	UTX	WMT	DIS
Log DY	-0.35	-0.20	-0.46	-0.39	-0.31	-0.22	-0.43	-0.06	-0.23	-0.01
Log EPR	-0.23	-0.19	-0.05	-0.51	-0.51	-0.11	**0.02**	-0.03	-0.44	-0.08
Log DPR	-0.10	-0.45	-0.45	-0.50	-0.12	**0.09**	-1.70	**0.35**	-0.26	-0.03
BMR	-0.12	-0.35	-0.15	-0.48	-0.12	-0.28	**0.03**	-0.06	-0.20	-0.06
3M Tbill rate	-0.36	-0.41	-0.64	-0.62	-0.32	-0.32	-0.13	-0.09	-0.17	**0.27**
Market return	-0.39	-0.23	-0.07	-0.37	-0.16	-0.18	-0.52	-0.19	-0.23	-0.20
LT yield	-0.39	-0.38	-0.25	-0.46	-0.40	-0.39	-0.73	-0.19	-0.32	-0.28
CPI inflation	-0.14	-0.16	-0.06	-0.33	-0.22	-0.15	-0.40	-0.10	-0.14	-0.02
Log TPR	-0.02	-0.24	-0.03	**0.03**	-0.21	-0.00	-0.16	**0.05**	-0.26	**0.01**
Log TPV	-0.18	-0.15	-0.07	-0.52	-0.18	-0.12	-0.48	-0.24	-0.24	**0.11**
Log REC	-0.88	-0.38	-0.25	-1.07	-0.38	-0.38	-1.83	-1.16	-0.23	-0.34
Log REC return	-0.59	-0.20	-0.05	-0.06	-0.03	-0.19	-0.14	-0.60	-0.15	-0.12

Note: The table provides forecast performance results in terms of mean out-of-sample R^2 for 20 Dow Jones constituents (sample: 2004 - 2015) with a one month forecast horizon. The benchmark model is a simple mean model. For each asset, we estimate a Bayesian VAR system with constant coefficients using the Minnesota prior outlined in section 3 for the monthly excess returns on an intercept and a lagged predictor variable, i.e. $\begin{bmatrix} r_t \\ x_t \end{bmatrix} = a + A_1 \begin{bmatrix} r_{t-i} \\ x_{t-i} \end{bmatrix} + \varepsilon_t$, $t = 1, \dots, T$. Further, DY is the dividend yield, PR is the earnings-price ratio, DPR is the dividend-price-ratio, BMR is the book-to-market ratio, LT is longterm yield, TPR is the target price return, TPV the target price variance and REC stands for recommendations. Values above zero indicate that a given predictor has better forecast performance than the benchmark model, while negative values suggest the opposite. All values are multiplied by 100. We test statistical significance in the average loss between the each model and a simple mean model using the Diebold and Mariano (1995) test. One/two/three asterisks denote rejection of the null hypothesis of equal predictive ability at the ten/five/one percent test level.

Table 4: Forecast performance in terms of out-of-sample R^2 for 20 Dow Jones constituents (sample: 2004 - 2015) using a TVP-BVAR(1) with stochastic volatility

Stock	AA	AAPL	AIG	AXP	BA	CAT	KO	DD	GE	HD
Log DY	-6.49	**0.68**	-3.12	-0.33	-0.89	-0.14	**0.36**	**0.26**	-0.15	**0.21**
Log EPR	-1.52	**0.37**	-0.44	-0.14	-5.60	**0.06**	-0.10	-0.01	-0.08	-0.25
Log DPR	-0.07	**0.36**	**0.32**	**0.39**	-0.95	**0.39**	**0.36**	**0.63**	-0.12	**0.04**
BMR	**0.30**	**0.28**	-0.99	-0.12	**0.21**	**0.22**	-0.56	**0.14**	**0.17**	-0.02
3M Tbill rate	**0.21**	**0.30**	**0.26**	**0.55**	**0.39**	**0.27**	-0.26	-0.09	-0.03	-0.06
Market return	-0.04	**0.67**	**0.06**	**0.13**	**0.03**	**0.36**	**0.03**	**0.07**	**0.28**	-0.11
LT yield	-0.08	**0.42**	**0.59**	**0.23**	-0.29	**0.00**	-0.11	-0.24	**0.06**	-0.22
CPI inflation	**0.13**	**0.66**	-0.09	**0.25**	**0.12**	-0.12	-0.07	-0.16	**0.07**	**0.55**
Log TPR	**0.47**	**0.35**	-0.41	-0.26	**0.38**	**0.24**	**0.31**	**0.36**	**0.18**	**0.01**
Log TPV	**0.30**	**0.90**	**0.44**	-0.08	**0.24**	-0.00	**0.12**	**0.15**	**0.39**	**0.77**
Log REC	-0.46	**0.43**	-0.03	-0.25	-0.81	**0.11**	-0.23	-0.25	-0.09	-0.05
Log REC return	**8.55*****	**0.10**	**0.32**	**0.41**	-0.21	**0.11**	**8.34*****	**8.13*****	**0.22**	**0.20**

Stock	INTC	IBM	JNJ	MCD	MRK	MSFT	PG	UTX	WMT	DIS
Log DY	-0.30	-0.11	-0.26	-0.22	-0.03	**0.06**	-0.18	**0.11**	-0.11	**0.12**
Log EPR	**0.26**	-0.06	-0.02	-0.06	-0.48	**0.04**	**0.24**	**0.44**	-0.35	**0.12**
Log DPR	-0.10	-0.38	-0.33	-0.31	-0.00	**0.46**	-1.48	**0.79**	-0.14	**0.27**
BMR	**0.27**	-0.28	-0.08	-0.42	**0.06**	-0.19	**0.19**	**0.21**	**0.01**	**0.07**
3M Tbill rate	**0.05**	**0.02**	-0.55	-0.23	**0.09**	**0.02**	**0.12**	**0.22**	-0.02	**0.58**
Market return	**0.04**	**0.06**	**0.05**	-0.18	-0.15	-0.09	-0.27	**0.10**	**0.23**	**0.16**
LT yield	-0.35	-0.11	-0.04	-0.34	-0.38	-0.20	-0.32	-0.09	-0.11	-0.17
CPI inflation	**0.06**	-0.09	-0.03	-0.13	-0.14	**0.16**	-0.01	**0.05**	-0.04	**0.04**
Log TPR	**0.44**	**0.46**	**0.23**	**0.78**	**0.17**	**0.43**	**0.28**	**0.47**	-0.09	**0.10**
Log TPV	**0.26**	**0.06**	**0.02**	-0.01	**0.33**	**0.39**	**0.10**	**0.56**	-0.16	**0.16**
Log REC	-0.66	-0.20	-0.01	-0.60	-0.05	**0.09**	-1.42	-0.74	-0.01	-0.13
Log REC return	-0.13	**0.06**	**0.19**	**0.42**	**0.20**	**0.20**	**0.13**	-0.51	-0.10	**0.13**

Note: The table provides forecast performance results in terms of mean out-of-sample R^2 for 20 Dow Jones constituents (sample: 2004 - 2015) with a one month forecast horizon. The benchmark model is a simple mean model. For each asset, we estimate a Bayesian VAR system with time-varying coefficients and stochastic volatility for the monthly excess returns on an intercept and a lagged predictor variable, i.e. $\begin{bmatrix} r_t \\ x_t \end{bmatrix} = a_t + A_{1,t} \begin{bmatrix} r_{t-i} \\ x_{t-i} \end{bmatrix} + \varepsilon_t$, $t = 1, \ldots, T$, $A_t = \phi A_{t-1} + (1-\phi)\underline{A}_0 + u_t$, where $A_t = [a_t \;\; A_{1,t}]$ is time-index for every single parameter, $\varepsilon_t \overset{iid}{\sim} \mathrm{N}(0, \Sigma_t)$, $u_t \overset{iid}{\sim} \mathrm{N}(0, \Omega_t)$ and ε_t and u_s are independent of one each other for all t and s. We estimate the model using forgetting factors with the following parameter values: $\lambda = 0.99$, $\kappa = 0.96$ and $\phi = 0.5$. Further, DY is the dividend yield, PR is the earnings-price ratio, DPR is the dividend-price-ratio, BMR is the book-to-market ratio, LT is longterm yield, TPR is the target price return, TPV the target price variance and REC stands for recommendations. Values above zero indicate that a given predictor has better forecast performance than the benchmark model, while negative values suggest the opposite. All values are multiplied by 100. We test statistical significance in the average loss between the each model and a simple mean model using the Diebold and Mariano (1995) test. One/two/three asterisks denote rejection of the null hypothesis of equal predictive ability at the ten/five/one percent test level.

Table 5: Forecast performance in terms of out-of-sample R^2 for 20 Dow Jones constituents (sample: 2004 - 2015) using a TVP-BVAR(1) with stochastic volatility and entropic tilting towards the mean of monthly target price implied expected returns

Stock	AA	AAPL	AIG	AXP	BA	CAT	KO	DD	GE	HD
Log DY	-6.45	**0.38**	-3.35	-0.57	-1.24	-0.11	**0.47**	**0.00**	-0.05	**0.10**
Log EPR	-1.74	**0.50**	-0.60	-0.03	-5.50	**0.26**	**0.10**	-0.13	-0.04	-0.51
Log DPR	**0.36**	**0.20**	**0.37**	-0.01	-0.75	**0.27**	-0.09	**0.52**	**0.17**	**0.34**
BMR	**0.06**	**0.67**	-0.73	-0.11	**0.41**	**0.24**	-0.67	-0.13	**0.25**	-0.01
3M Tbill rate	**0.32**	**0.51**	**0.37**	**0.80**	**0.35**	-0.13	**0.12**	-0.21	-0.04	-0.12
Market return	**0.13**	**0.31**	**0.15**	**0.25**	**0.21**	**0.23**	**0.18**	-0.10	**0.16**	-0.49
LT yield	**0.23**	**0.50**	**0.53**	**0.36**	**0.09**	**0.09**	**0.05**	-0.20	**0.38**	-0.20
CPI inflation	**0.27**	**0.72**	-0.05	**0.24**	-0.01	**0.21**	-0.19	-0.08	**0.28**	**0.48**
Log TPR	**0.20**	**0.25**	**0.15**	-0.42	-0.03	**0.76**	**0.54**	**0.45**	-0.02	**0.61**
Log TPV	-0.02	**0.35**	**0.76**	**0.31**	**0.60**	-0.47	-0.14	**0.52**	**0.52**	**0.84**
Log REC	-0.10	**0.33**	**0.29**	**0.01**	-0.79	**0.11**	-0.51	-0.56	**0.09**	**0.43**
Log REC return	**8.59*****	-0.28	**0.18**	**0.63**	**0.15**	**0.30**	**8.40*****	**8.36*****	**0.24**	**0.38**
Stock	INTC	IBM	JNJ	MCD	MRK	MSFT	PG	UTX	WMT	DIS
Log DY	**0.05**	**0.32**	-0.43	-0.14	**0.06**	**0.28**	**0.12**	**0.43**	**0.17**	**0.05**
Log EPR	-0.19	-0.03	**0.04**	-0.09	-0.25	-0.07	**0.20**	**0.26**	-0.05	-0.01
Log DPR	**0.38**	-0.29	-0.08	-0.05	**0.38**	**0.22**	-1.54	**0.41**	**0.17**	**0.27**
BMR	**0.40**	-0.16	**0.04**	-0.42	-0.06	-0.25	**0.07**	**0.39**	-0.04	**0.21**
3M Tbill rate	**0.18**	-0.16	-0.14	-0.52	**0.09**	-0.08	**0.03**	**0.10**	**0.21**	**0.76**
Market return	**0.08**	**0.12**	-0.00	-0.17	**0.24**	-0.17	-0.50	-0.03	**0.08**	**0.24**
LT yield	**0.04**	-0.37	**0.30**	-0.43	-0.09	**0.11**	-0.45	**0.22**	-0.10	**0.13**
CPI inflation	**0.14**	**0.30**	**0.24**	-0.04	-0.12	-0.04	**0.01**	-0.10	-0.10	**0.01**
Log TPR	**0.10**	-0.07	**0.07**	**0.88**	**0.16**	**0.85**	**0.73**	**0.35**	-0.04	**0.62**
Log TPV	**0.02**	-0.11	**0.48**	-0.20	**0.71**	**0.48**	**0.21**	**0.36**	**0.03**	**0.59**
Log REC	-0.81	-0.19	-0.10	-0.96	-0.30	-0.13	-1.79	-0.83	**0.10**	**0.10**
Log REC return	-0.57	**0.05**	**0.18**	**0.44**	**0.09**	-0.14	**0.29**	-0.32	**0.26**	**0.40**

Note: The table provides forecast performance results in terms of mean out-of-sample R^2 for 20 Dow Jones constituents (sample: 2004 - 2015) with a one month forecast horizon. The benchmark model is a simple mean model. For each asset, we estimate a Bayesian VAR system with time-varying coefficients and stochastic volatility for the monthly excess returns on an intercept and a lagged predictor variable, i.e. $\begin{bmatrix} r_t \\ x_t \end{bmatrix} = a_t + A_{1,t} \begin{bmatrix} r_{t-i} \\ x_{t-i} \end{bmatrix} + \varepsilon_t$, $t = 1, \ldots, T$, $A_t = \phi A_{t-1} + (1-\phi)\underline{A}_0 + u_t$, where $A_t = [a_t \ \ A_{1,t}]$ is time-index for every single parameter, $\varepsilon_t \overset{iid}{\sim} \mathrm{N}(0, \Sigma_t)$, $u_t \overset{iid}{\sim} \mathrm{N}(0, \Omega_t)$ and ε_t and u_s are independent of one each other for all t and s. We estimate the model using forgetting factors with the following parameter values: $\lambda = 0.99$, $\kappa = 0.96$ and $\phi = 0.5$. The mean of the predictive distribtion is tilted towards the mean of the monthly forward target price implied expected returns. Further, DY is the dividend yield, PR is the earnings-price ratio, DPR is the dividend-price-ratio, BMR is the book-to-market ratio, LT is longterm yield, TPR is the target price return, TPV the target price variance and REC stands for recommendations. Values above zero indicate that a given predictor has better forecast performance than the benchmark model, while negative values suggest the opposite. All values are multiplied by 100. We test statistical significance in the average loss between the each model and a simple mean model using the Diebold and Mariano (1995) test. One/two/three asterisks denote rejection of the null hypothesis of equal predictive ability at the ten/five/one percent test level.

Table 6: Forecast performance in terms of out-of-sample R^2 for 20 Dow Jones constituents (sample: 2004 - 2015) using a TVP-BVAR(1) with stochastic volatility and entropic tilting towards the mean and variance of monthly target price implied expected returns

Stock	AA	AAPL	AIG	AXP	BA	CAT	KO	DD	GE	HD
Log DY	-6.77	**0.60**	-2.88	**0.10**	-0.33	-0.17	**0.30**	**0.34**	-0.34	**0.93**
Log EPR	-1.76	**0.38**	**0.25**	**0.70**	-5.48	**0.21**	**0.66**	**1.00**	**0.85**	-0.51
Log DPR	-0.01	**0.57**	**0.12**	**0.39**	-0.50	**0.99**	**0.06**	**0.30**	**0.66**	**0.93**
BMR	**0.43**	**1.30****	-0.72	**0.76**	**0.70**	**0.62**	**0.27**	**0.93**	**0.28**	**0.42**
3M Tbill rate	**0.13**	**0.84**	**1.05***	**1.33****	**1.23****	**0.53**	-0.19	-0.05	**0.38**	-0.12
Market return	**0.99**	**0.34**	**1.08***	**0.33**	**0.34**	**0.68**	**0.61**	**0.59**	**0.03**	-0.16
LT yield	-0.10	**0.49**	**1.09***	**0.51**	**0.66**	-0.21	**0.02**	**0.42**	**0.23**	**0.59**
CPI inflation	-0.09	**0.24**	**0.64**	**1.03***	**0.75**	**0.23**	**0.67**	**0.62**	**0.11**	**1.09***
Log TPR	**0.96**	**0.88**	-0.56	-0.03	**1.26****	-0.09	**0.95**	-0.08	**0.38**	**0.88**
Log TPV	**0.40**	**0.53**	**0.67**	**1.24****	**0.46**	**0.47**	**1.30****	**0.49**	**0.75**	**1.25****
Log REC	-0.16	**1.20****	**0.07**	**0.27**	-0.67	**0.54**	**0.53**	**0.25**	**0.20**	**0.99**
Log REC return	**8.28*****	**0.61**	**0.79**	**0.52**	**0.48**	**1.01***	**8.49*****	**8.09*****	**0.04**	**0.04**

Stock	INTC	IBM	JNJ	MCD	MRK	MSFT	PG	UTX	WMT	DIS
Log DY	**0.01**	**0.98**	**0.11**	**0.57**	**0.13**	**0.27**	-0.13	**0.99**	**0.37**	**0.52**
Log EPR	-0.17	**0.67**	**0.38**	**0.57**	**0.61**	**0.69**	**0.96**	**0.31**	**0.39**	**0.44**
Log DPR	**0.14**	**0.56**	**0.50**	**0.22**	**0.87**	**1.21****	-0.64	**1.16***	**0.75**	**0.11**
BMR	**0.75**	**0.17**	**0.79**	**0.58**	**0.90**	**0.69**	**1.13***	**0.74**	**0.53**	**0.92**
3M Tbill rate	**0.51**	**0.15**	**0.16**	**0.51**	**0.13**	**0.26**	**0.54**	**0.06**	**0.52**	**0.66**
Market return	**0.66**	**0.44**	**0.09**	**0.29**	**0.55**	**0.73**	**0.20**	**0.30**	**0.16**	**0.10**
LT yield	**0.31**	-0.06	-0.22	**0.42**	**0.65**	**0.12**	-0.55	**0.14**	**0.23**	**0.13**
CPI inflation	-0.05	**0.74**	**0.61**	**0.36**	**0.90**	**1.01***	**0.68**	**0.76**	**0.72**	**0.43**
Log TPR	**1.07***	**0.64**	**0.06**	**0.46**	**1.23****	**0.29**	**1.24****	**0.46**	**0.33**	**0.20**
Log TPV	**0.38**	**0.03**	**1.40****	**0.37**	**0.10**	**0.39**	**0.11**	-0.01	**0.32**	**0.76**
Log REC	-0.54	-0.01	**0.92**	-0.30	**0.41**	**0.09**	-0.75	-0.17	-0.21	**0.13**
Log REC return	**0.06**	-0.03	**0.29**	**0.56**	**0.06**	**0.35**	**0.78**	-0.14	**0.66**	**0.36**

Note: The table provides forecast performance results in terms of mean out-of-sample R^2 for 20 Dow Jones constituents (sample: 2004 - 2015) with a one month forecast horizon. The benchmark model is a simple mean model. For each asset, we estimate a Bayesian VAR system with time-varying coefficients and stochastic volatility for the monthly excess returns on an intercept and a lagged predictor variable, i.e. $\begin{bmatrix} r_t \\ x_t \end{bmatrix} = a_t + A_{1,t} \begin{bmatrix} r_{t-i} \\ x_{t-i} \end{bmatrix} + \varepsilon_t$, $t = 1, \ldots, T$, $A_t = \phi A_{t-1} + (1-\phi)\underline{A}_0 + u_t$, where $A_t = [a_t \;\; A_{1,t}]$ is time-index for every single parameter, $\varepsilon_t \overset{iid}{\sim} \mathrm{N}(0, \Sigma_t)$, $u_t \overset{iid}{\sim} \mathrm{N}(0, \Omega_t)$ and ε_t and u_s are independent of one each other for all t and s. We estimate the model using forgetting factors with the following parameter values: $\lambda = 0.99$, $\kappa = 0.96$ and $\phi = 0.5$. The mean and variance of the predictive distribution are tilted towards the mean and variance of the monthly forward target price implied expected returns. Further, DY is the dividend yield, PR is the earnings-price ratio, DPR is the dividend-price-ratio, BMR is the book-to-market ratio, LT is longterm yield, TPR is the target price return, TPV the target price variance and REC stands for recommendations. Values above zero indicate that a given predictor has better forecast performance than the benchmark model, while negative values suggest the opposite. All values are multiplied by 100. We test statistical significance in the average loss between the each model and a simple mean model using the Diebold and Mariano (1995) test. One/two/three asterisks denote rejection of the null hypothesis of equal predictive ability at the ten/five/one percent test level.

Table 7: Forecast performance in terms of average log predictive score differentials for 20 Dow Jones constituents (sample: 2004 - 2015) using a Bayesian VAR(1)

Stock	AA	AAPL	AIG	AXP	BA	CAT	KO	DD	GE	HD
Log DY	-2.02	**0.04**	**7.14*****	-0.44	-0.23	-0.39	-0.00	-0.08	-0.25	-0.05
Log EPR	-0.10	-0.01	**0.55***	-0.16	-0.26	-0.06	-0.07	-0.05	-0.13	-0.15
Log DPR	-0.09	-0.03	-0.28	-0.05	-0.33	**0.05**	**0.02**	**0.16**	-0.03	-0.06
BMR	-0.03	**0.03**	**0.91***	-0.07	-0.04	-0.06	-0.11	-0.07	**0.06**	-0.03
3M Tbill rate	**0.01**	-0.01	-0.46	**0.33**	**0.12**	-0.09	-0.06	-0.01	**0.05**	-0.05
Market return	-0.06	**0.04**	-0.20	-0.03	-0.03	-0.09	**0.01**	-0.11	-0.03	-0.08
LT yield	-0.09	**0.00**	-1.00	-0.03	-0.10	-0.18	-0.09	-0.08	**0.01**	-0.08
CPI inflation	-0.12	**0.04**	-0.20	-0.10	-0.07	-0.17	-0.06	-0.09	-0.09	**0.02**
Log TPR	-0.02	-0.05	**0.67***	-0.26	-0.05	-0.10	-0.08	-0.06	-0.26	-0.07
Log TPV	-0.19	**0.01**	-0.20	-0.09	-0.06	-0.48	-0.04	-0.09	-0.14	**0.07**
Log REC	-0.44	-0.04	-0.23	-0.27	-0.28	-0.26	-0.11	-0.29	-0.09	-0.03
Log REC return	**0.00**	-0.16	-0.28	**0.08**	-0.12	-0.02	**0.00**	**0.00**	-0.03	**0.01**

Stock	INTC	IBM	JNJ	MCD	MRK	MSFT	PG	UTX	WMT	DIS
Log DY	-0.08	-0.00	-0.08	-0.07	-0.11	-0.06	-0.05	-0.03	-0.05	-0.00
Log EPR	-0.08	-0.04	-0.03	-0.01	-0.17	-0.03	-0.01	-0.03	-0.08	-0.05
Log DPR	-0.01	-0.08	-0.06	-0.07	-0.03	**0.02**	-0.19	**0.13**	-0.04	**0.01**
BMR	-0.06	-0.05	-0.01	-0.07	-0.04	-0.08	**0.01**	-0.03	-0.04	-0.02
3M Tbill rate	-0.09	-0.07	-0.08	-0.05	-0.09	-0.08	-0.04	-0.01	-0.01	**0.18**
Market return	-0.08	-0.04	**0.03**	-0.05	-0.06	-0.06	**0.05**	-0.04	-0.04	-0.04
LT yield	-0.09	-0.08	-0.06	-0.08	-0.11	-0.11	-0.11	-0.05	-0.07	-0.08
CPI inflation	-0.06	-0.05	-0.03	-0.05	-0.09	-0.05	-0.06	-0.05	-0.04	-0.01
Log TPR	-0.03	-0.07	-0.02	-0.01	-0.08	-0.02	-0.04	-0.00	-0.06	**0.02**
Log TPV	-0.07	-0.02	-0.02	-0.07	-0.06	-0.03	-0.07	-0.07	-0.05	**0.04**
Log REC	-0.18	-0.07	-0.06	-0.14	-0.11	-0.10	-0.21	-0.23	-0.06	-0.10
Log REC return	-0.02	-0.06	-0.03	-0.02	**0.00**	-0.07	-0.03	-0.14	-0.04	-0.04

Note: The table provides forecast performance results in terms of average log predictive score differentials between the benchmark mean model and a single regressor model for 20 Dow Jones constituents (sample: 2004 - 2015) with a one month forecast horizon. For each asset, we estimate a Bayesian VAR system with constant coefficients using the Minnesota prior outlined in section 3 for the monthly excess returns on an intercept and a lagged predictor variable, i.e. $\begin{bmatrix} r_t \\ x_t \end{bmatrix} = a + A_1 \begin{bmatrix} r_{t-i} \\ x_{t-i} \end{bmatrix} + \varepsilon_t$, $t = 1, \ldots, T$. Further, DY is the dividend yield, PR is the earnings-price ratio, DPR is the dividend-price-ratio, BMR is the book-to-market ratio, LT is longterm yield, TPR is the target price return, TPV the target price variance and REC stands for recommendations. Values above zero indicate that a given predictor has better forecast performance than the benchmark model, while negative values suggest the opposite. All values are multiplied by 100. We test statistical significance in the average loss between the each model and a simple mean model using the Diebold and Mariano (1995) test. One/two/three asterisks denote rejection of the null hypothesis of equal predictive ability at the ten/five/one percent test level.

Table 8: Forecast performance in terms of average log predictive score differentials for 20 Dow Jones constituents (sample: 2004 - 2015) using a TVP-BVAR(1) with stochastic volatility

Stock	AA	AAPL	AIG	AXP	BA	CAT	KO	DD	GE	HD
Log DY	-1.98	**0.31**	**7.19*****	**0.01**	**0.18**	-0.36	**0.48**	-0.05	-0.06	**0.16**
Log EPR	**0.03**	**0.11**	**0.68**	**0.00**	**0.03**	**0.28**	**0.26**	**0.15**	-0.03	**0.35**
Log DPR	**0.31**	**0.21**	-0.12	**0.30**	-0.23	**0.07**	**0.42**	**0.42**	**0.22**	**0.09**
BMR	-0.02	**0.34**	**1.25****	**0.03**	**0.08**	-0.02	**0.12**	**0.14**	**0.23**	**0.32**
3M Tbill rate	**0.48**	**0.33**	-0.39	**0.35**	**0.57**	**0.17**	**0.15**	**0.31**	**0.52**	**0.28**
Market return	**0.31**	**0.23**	**0.16**	**0.34**	-0.01	-0.04	**0.42**	**0.20**	**0.43**	**0.18**
LT yield	**0.15**	**0.19**	-0.95	**0.22**	**0.14**	**0.23**	-0.05	**0.07**	**0.04**	**0.27**
CPI inflation	**0.17**	**0.53**	**0.13**	**0.14**	**0.02**	**0.24**	**0.00**	**0.12**	**0.28**	**0.35**
Log TPR	**0.30**	**0.24**	**1.17***	**0.48**	**0.10**	**0.46**	**0.34**	**0.63**	**0.05**	**0.52**
Log TPV	**0.17**	**0.62**	**0.42**	**0.69**	**0.05**	-0.40	**0.38**	**0.30**	**0.40**	**0.49**
Log REC	**0.04**	**0.42**	**0.13**	**0.04**	-0.03	**0.07**	**0.31**	-0.21	**0.18**	**0.47**
Log REC return	**0.27**	**0.24**	**0.17**	**0.51**	**0.12**	**0.24**	**0.40**	**0.05**	**0.45**	**0.10**

Stock	INTC	IBM	JNJ	MCD	MRK	MSFT	PG	UTX	WMT	DIS
Log DY	-0.06	**0.09**	**0.34**	**0.14**	-0.07	**0.33**	**0.27**	**0.14**	**0.11**	**0.12**
Log EPR	**0.21**	**0.17**	**0.46**	**0.14**	-0.01	**0.18**	**0.47**	**0.36**	**0.31**	**0.41**
Log DPR	**0.43**	**0.17**	**0.31**	**0.01**	**0.23**	**0.06**	-0.07	**0.46**	**0.19**	**0.14**
BMR	**0.27**	**0.01**	**0.16**	**0.02**	**0.29**	**0.05**	**0.35**	-0.03	-0.02	**0.36**
3M Tbill rate	**0.01**	**0.23**	**0.21**	**0.16**	**0.11**	**0.00**	**0.11**	**0.29**	**0.08**	**0.28**
Market return	**0.11**	**0.07**	**0.09**	-0.00	**0.35**	**0.08**	**0.38**	**0.15**	**0.32**	**0.11**
LT yield	**0.14**	**0.11**	**0.40**	**0.22**	**0.25**	**0.11**	**0.24**	**0.40**	**0.17**	-0.03
CPI inflation	**0.43**	**0.24**	**0.41**	**0.18**	**0.40**	**0.21**	-0.02	-0.05	**0.04**	**0.28**
Log TPR	**0.25**	**0.40**	**0.02**	**0.19**	**0.47**	**0.57**	**0.51**	**0.35**	**0.20**	**0.24**
Log TPV	**0.05**	**0.19**	**0.58**	**0.28**	**0.23**	**0.28**	**0.49**	-0.05	**0.29**	**0.19**
Log REC	**0.15**	**0.24**	**0.23**	**0.18**	-0.06	**0.15**	**0.12**	**0.00**	**0.04**	**0.11**
Log REC return	**0.17**	**0.07**	-0.01	**0.00**	**0.31**	**0.40**	**0.39**	**0.25**	**0.33**	**0.28**

Note: The table provides forecast performance results in terms of average log predictive score differentials between the benchmark mean model and a single regressor model for 20 Dow Jones constituents (sample: 2004 - 2015) with a one month forecast horizon. For each asset, we estimate a Bayesian VAR system with time-varying coefficients and stochastic volatility for the monthly excess returns on an intercept and a lagged predictor variable, i.e. $\begin{bmatrix} r_t \\ x_t \end{bmatrix} = a_t + A_{1,t} \begin{bmatrix} r_{t-i} \\ x_{t-i} \end{bmatrix} + \varepsilon_t$, $t = 1, \ldots, T$, $A_t = \phi A_{t-1} + (1-\phi)\underline{A}_0 + u_t$, where $A_t = [a_t \;\; A_{1,t}]$ is time-index for every single parameter, $\varepsilon_t \overset{iid}{\sim} \mathrm{N}(0, \Sigma_t)$, $u_t \overset{iid}{\sim} \mathrm{N}(0, \Omega_t)$ and ε_t and u_s are independent of one each other for all t and s. We estimate the model using forgetting factors with the following parameter values: $\lambda = 0.99$, $\kappa = 0.96$ and $\phi = 0.5$. Further, DY is the dividend yield, PR is the earnings-price ratio, DPR is the dividend-price-ratio, BMR is the book-to-market ratio, LT is longterm yield, TPR is the target price return, TPV the target price variance and REC stands for recommendations. Values above zero indicate that a given predictor has better forecast performance than the benchmark model, while negative values suggest the opposite. All values are multiplied by 100. We test statistical significance in the average loss between the each model and a simple mean model using the Diebold and Mariano (1995) test. One/two/three asterisks denote rejection of the null hypothesis of equal predictive ability at the ten/five/one percent test level.

Table 9: Forecast performance in terms of average log predictive score differentials for 20 Dow Jones constituents (sample: 2004 - 2015) using a TVP-BVAR(1) with stochastic volatility and entropic tilting towards the mean of monthly target price implied expected returns

Stock	AA	AAPL	AIG	AXP	BA	CAT	KO	DD	GE	HD
Log DY	-1.65	**0.45**	**7.57*****	-0.15	**0.01**	-0.02	**0.23**	**0.09**	**0.06**	**0.17**
Log EPR	-0.03	**0.52**	**0.66**	-0.11	**0.10**	**0.18**	**0.30**	**0.12**	**0.04**	**0.20**
Log DPR	**0.30**	**0.44**	-0.07	**0.17**	-0.26	**0.30**	**0.42**	**0.28**	**0.06**	**0.02**
BMR	**0.03**	**0.08**	**1.22****	-0.01	**0.47**	**0.28**	**0.18**	**0.07**	**0.40**	**0.18**
3M Tbill rate	**0.08**	**0.19**	-0.33	**0.39**	**0.23**	-0.05	-0.00	**0.48**	**0.59**	**0.04**
Market return	**0.29**	**0.24**	**0.16**	**0.40**	**0.12**	**0.08**	**0.36**	**0.27**	**0.07**	**0.33**
LT yield	**0.09**	**0.38**	-0.73	**0.13**	**0.34**	**0.24**	-0.02	**0.23**	**0.15**	**0.40**
CPI inflation	**0.24**	**0.37**	-0.12	**0.23**	**0.20**	**0.22**	**0.01**	**0.01**	**0.13**	**0.21**
Log TPR	**0.35**	**0.62**	**1.17***	**0.20**	**0.59**	**0.22**	**0.23**	**0.35**	-0.07	**0.21**
Log TPV	**0.36**	**0.53**	**0.32**	-0.01	**0.84**	**0.40**	**0.76**	**0.29**	-0.02	**0.72**
Log REC	-0.03	**0.08**	-0.07	**0.11**	-0.13	-0.21	-0.02	**0.21**	**0.13**	**0.26**
Log REC return	**0.13**	-0.11	-0.15	**0.50**	-0.10	-0.02	**0.11**	**0.39**	**0.51**	**0.47**

Stock	INTC	IBM	JNJ	MCD	MRK	MSFT	PG	UTX	WMT	DIS
Log DY	**0.25**	**0.16**	**0.30**	**0.45**	-0.01	**0.36**	**0.39**	**0.33**	**0.05**	**0.36**
Log EPR	**0.11**	**0.13**	**0.23**	**0.28**	-0.03	**0.10**	**0.46**	**0.31**	**0.45**	**0.16**
Log DPR	**0.15**	**0.28**	**0.06**	**0.31**	-0.02	**0.05**	**0.09**	**0.34**	**0.10**	**0.35**
BMR	**0.19**	**0.47**	**0.04**	-0.05	**0.47**	**0.34**	**0.36**	**0.05**	**0.47**	-0.01
3M Tbill rate	**0.14**	**0.45**	**0.37**	**0.40**	**0.27**	**0.29**	**0.48**	**0.00**	**0.11**	**0.68**
Market return	**0.12**	**0.21**	**0.13**	**0.36**	**0.45**	**0.33**	**0.29**	**0.19**	**0.17**	**0.40**
LT yield	**0.21**	**0.05**	**0.03**	-0.01	-0.02	**0.24**	-0.07	**0.05**	-0.02	**0.33**
CPI inflation	**0.35**	**0.37**	**0.34**	**0.24**	**0.42**	**0.18**	**0.42**	**0.35**	**0.31**	**0.43**
Log TPR	**0.68**	-0.06	**0.81**	**0.03**	**0.56**	**0.18**	**0.57**	**0.56**	**0.10**	**0.71**
Log TPV	**0.55**	**0.74**	**0.67**	**0.27**	**0.60**	**0.21**	**0.36**	**0.14**	**0.69**	**0.88**
Log REC	-0.11	**0.34**	**0.32**	**0.08**	**0.13**	**0.07**	**0.34**	**0.18**	**0.34**	**0.21**
Log REC return	-0.01	-0.00	**0.06**	**0.21**	**0.14**	**0.38**	**0.09**	**0.18**	**0.15**	**0.25**

Note: The table provides forecast performance results in terms of average log predictive score differentials between the benchmark mean model and a single regressor model for 20 Dow Jones constituents (sample: 2004 - 2015) with a one month forecast horizon. For each asset, we estimate a Bayesian VAR system with time-varying coefficients and stochastic volatility for the monthly excess returns on an intercept and a lagged predictor variable, i.e. $\begin{bmatrix} r_t \\ x_t \end{bmatrix} = a_t + A_{1,t} \begin{bmatrix} r_{t-i} \\ x_{t-i} \end{bmatrix} + \varepsilon_t$, $t = 1, \ldots, T$, $A_t = \phi A_{t-1} + (1-\phi)\underline{A}_0 + u_t$, where $A_t = [a_t \;\; A_{1,t}]$ is time-index for every single parameter, $\varepsilon_t \overset{iid}{\sim} \mathrm{N}(0, \Sigma_t)$, $u_t \overset{iid}{\sim} \mathrm{N}(0, \Omega_t)$ and ε_t and u_s are independent of one each other for all t and s. We estimate the model using forgetting factors with the following parameter values: $\lambda = 0.99$, $\kappa = 0.96$ and $\phi = 0.5$. The mean of the predictive distribtion is tilted towards the mean of the monthly forward target price implied expected returns. Further, DY is the dividend yield, PR is the earnings-price ratio, DPR is the dividend-price-ratio, BMR is the book-to-market ratio, LT is longterm yield, TPR is the target price return, TPV the target price variance and REC stands for recommendations. Values above zero indicate that a given predictor has better forecast performance than the benchmark model, while negative values suggest the opposite. All values are multiplied by 100. We test statistical significance in the average loss between the each model and a simple mean model using the Diebold and Mariano (1995) test. One/two/three asterisks denote rejection of the null hypothesis of equal predictive ability at the ten/five/one percent test level.

Table 10: Forecast performance in terms of average log predictive score differentials for 20 Dow Jones constituents (sample: 2004 - 2015) using a TVP-BVAR(1) with stochastic volatility and entropic tilting towards the mean and variance of monthly target price implied expected returns

Stock	AA	AAPL	AIG	AXP	BA	CAT	KO	DD	GE	HD
Log DY	-1.41	**0.06**	**7.66*****	-0.21	**0.52**	**0.04**	**0.70**	**0.06**	-0.13	**0.28**
Log EPR	**0.69**	**0.64**	**1.14***	**0.15**	**0.58**	**0.49**	**0.49**	**0.75**	**0.96**	-0.02
Log DPR	**1.05***	**0.08**	-0.20	**1.03***	**0.54**	**0.51**	**0.08**	**0.88**	**0.10**	**0.55**
BMR	**0.83**	**0.20**	**1.98*****	**0.65**	**0.37**	**0.87**	**0.16**	**0.00**	**0.68**	**0.68**
3M Tbill rate	**0.49**	**0.75**	-0.38	**0.93**	**0.74**	**0.79**	**0.94**	**0.05**	**0.22**	**0.86**
Market return	**0.94**	**1.07***	**0.33**	**0.70**	**0.64**	**0.42**	**0.03**	**0.07**	**0.64**	**0.01**
LT yield	**0.07**	**1.17***	-0.01	**0.96**	**0.09**	**0.65**	**0.95**	-0.05	**0.02**	**0.71**
CPI inflation	-0.05	**0.72**	**0.27**	**0.54**	**0.61**	**0.97**	**0.03**	**0.43**	**0.83**	**0.64**
Log TPR	**0.12**	-0.04	**1.86*****	**0.41**	**0.04**	**0.86**	**1.18***	**1.21****	**0.16**	**0.28**
Log TPV	**0.74**	**0.87**	**0.15**	**0.77**	**0.69**	-0.14	**1.42****	**0.67**	**0.98**	**1.51****
Log REC	-0.05	**0.58**	**0.84**	**0.24**	**0.73**	-0.12	**0.15**	**0.33**	**1.09***	**0.67**
Log REC return	**0.36**	**0.24**	**0.84**	**1.24****	**0.76**	**0.45**	**0.69**	**1.04***	**0.58**	**0.54**

Stock	INTC	IBM	JNJ	MCD	MRK	MSFT	PG	UTX	WMT	DIS
Log DY	**1.05***	**0.13**	**0.23**	-0.07	**0.52**	**0.79**	**1.05***	**0.19**	**0.93**	**0.60**
Log EPR	**0.71**	**0.49**	**0.33**	**0.21**	**0.15**	**0.74**	**1.08***	**0.36**	**0.80**	**0.87**
Log DPR	**0.53**	**0.28**	**0.45**	**0.10**	**0.05**	**0.68**	**0.52**	**0.61**	**0.17**	**0.92**
BMR	**0.95**	**0.43**	**0.13**	**0.25**	**0.48**	**0.18**	**0.41**	**0.63**	**0.39**	**0.67**
3M Tbill rate	**0.55**	**0.93**	**0.52**	**0.16**	**0.12**	**0.85**	**0.99**	**0.05**	**0.22**	**1.08***
Market return	**0.59**	**0.44**	**0.88**	**0.12**	-0.03	**0.21**	**0.58**	**0.62**	-0.04	**0.74**
LT yield	**0.72**	**0.39**	**0.24**	**0.64**	**1.04***	**0.34**	**0.98**	**0.28**	**0.31**	**0.07**
CPI inflation	**0.38**	**0.38**	**0.92**	**1.03***	**0.43**	**1.02***	-0.02	**0.24**	**0.80**	**0.59**
Log TPR	**0.90**	**0.19**	**0.37**	**1.36****	**1.01***	**0.84**	**0.57**	**0.13**	**0.70**	**1.11***
Log TPV	**0.83**	**0.11**	**1.26****	**0.98**	**0.29**	**1.18***	**1.41****	**0.41**	**0.04**	**0.87**
Log REC	**0.86**	**0.03**	-0.06	**0.44**	-0.10	**0.28**	**0.00**	**0.92**	**0.47**	**0.07**
Log REC return	**0.47**	**0.46**	**0.24**	**0.43**	**0.82**	**0.66**	**0.37**	**0.99**	**0.30**	**0.19**

Note: The table provides forecast performance results in terms of average log predictive score differentials between the benchmark mean model and a single regressor model for 20 Dow Jones constituents (sample: 2004 - 2015) with a one month forecast horizon. For each asset, we estimate a Bayesian VAR system with time-varying coefficients and stochastic volatility for the monthly excess returns on an intercept and a lagged predictor variable, i.e. $\begin{bmatrix} r_t \\ x_t \end{bmatrix} = a_t + A_{1,t} \begin{bmatrix} r_{t-i} \\ x_{t-i} \end{bmatrix} + \varepsilon_t$, $t = 1, \ldots, T$, $A_t = \phi A_{t-1} + (1-\phi)\underline{A}_0 + u_t$, where $A_t = [a_t \;\; A_{1,t}]$ is time-index for every single parameter, $\varepsilon_t \overset{iid}{\sim} \mathrm{N}(0, \Sigma_t)$, $u_t \overset{iid}{\sim} \mathrm{N}(0, \Omega_t)$ and ε_t and u_s are independent of one each other for all t and s. We estimate the model using forgetting factors with the following parameter values: $\lambda = 0.99$, $\kappa = 0.96$ and $\phi = 0.5$. The mean and variance of the predictive distribution are tilted towards the mean and variance of the monthly forward target price implied expected returns. Further, DY is the dividend yield, PR is the earnings-price ratio, DPR is the dividend-price-ratio, BMR is the book-to-market ratio, LT is longterm yield, TPR is the target price return, TPV the target price variance and REC stands for recommendations. Values above zero indicate that a given predictor has better forecast performance than the benchmark model, while negative values suggest the opposite. All values are multiplied by 100. We test statistical significance in the average loss between the each model and a simple mean model using the Diebold and Mariano (1995) test. One/two/three asterisks denote rejection of the null hypothesis of equal predictive ability at the ten/five/one percent test level.

Table 11: Forecast performance in terms of out-of-sample R^2 for 20 Dow Jones constituents (sample: 2004 - 2015) for various forecasting models

Stock	AA	AAPL	AIG	AXP	BA	CAT	KO	DD	GE	HD
AR1	-0.10	**0.09**	-0.07	**0.33**	**0.06**	-0.18	-0.36	-0.28	-0.11	-0.17
VAR-Full	-1.16	-1.57	-1.78	-1.25	-0.57	-1.08	-1.86	-0.50	-0.33	-0.71
VAR-Minnesota	**0.74**	**0.67**	**0.88**	**0.39**	**0.65**	**0.11**	**0.47**	-0.09	**0.33**	**0.22**
TVPVAR-DMA	**0.85**	**0.67**	-0.03	**0.39**	**0.91**	**0.30**	**0.04**	**0.55**	**0.27**	**0.39**
TVPVAR-DMS	**0.62**	**0.37**	-0.09	**0.15**	**0.85**	**0.00**	-0.20	**0.43**	**0.06**	**0.24**
TVPVAR-DMAm	**0.89**	**0.70**	-0.02	**0.41**	**0.94**	**0.35**	**0.09**	**0.60**	**0.31**	**0.41**
TVPVAR-DMAm/v	**1.91*****	**0.80**	**0.02**	**0.72**	**1.55****	**0.88**	**1.04***	**1.15***	**0.75**	**0.98**
TVPVAR-DMSm	**0.63**	**0.38**	-0.07	**0.20**	**0.87**	**0.02**	-0.19	**0.47**	**0.08**	**0.28**
TVPVAR-DMSm/v	**1.66****	**1.30****	**0.31**	**1.05***	**1.77****	**0.79**	**0.53**	**0.83**	**0.34**	**0.55**
Bayesian lasso	**0.85**	**0.74**	**0.00**	**0.46**	**0.95**	**0.34**	**0.05**	**0.56**	**0.28**	**0.44**

Stock	INTC	IBM	JNJ	MCD	MRK	MSFT	PG	UTX	WMT	DIS
AR1	-0.36	-0.41	-0.64	-0.62	-0.32	-0.32	-0.13	-0.09	-0.17	**0.27**
VAR-Full	-1.41	-2.36	-2.06	-1.24	-0.90	-2.02	-1.96	-1.37	-0.68	**0.10**
VAR-Minnesota	**0.47**	**0.26**	-0.43	-0.30	-0.19	**0.35**	**0.44**	**0.08**	-0.02	**0.75**
TVPVAR-DMA	**0.39**	**0.50**	-0.29	**0.06**	**0.70**	-0.24	**0.48**	**0.34**	**0.15**	**0.55**
TVPVAR-DMS	**0.15**	**0.39**	-0.31	-0.11	**0.43**	-0.30	**0.35**	**0.12**	**0.14**	**0.27**
TVPVAR-DMAm	**0.44**	**0.53**	-0.28	**0.08**	**0.74**	-0.21	**0.52**	**0.39**	**0.20**	**0.58**
TVPVAR-DMAm/v	**1.17***	**0.52**	**0.59**	**0.22**	**1.23****	**0.04**	**0.89**	**1.07***	**0.34**	**0.86**
TVPVAR-DMSm	**0.18**	**0.41**	-0.27	-0.08	**0.45**	-0.26	**0.40**	**0.14**	**0.16**	**0.30**
TVPVAR-DMSm/v	**0.47**	**0.92**	-0.28	**0.89**	**1.28****	-0.24	**0.51**	**0.53**	**0.56**	**0.67**
Bayesian lasso	**0.41**	**0.59**	-0.21	**0.16**	**0.75**	-0.22	**0.50**	**0.40**	**0.23**	**0.59**

Note: The table provides forecast performance results in terms of mean out-of-sample R^2 for 20 Dow Jones constituents (sample: 2004 - 2015) with a one month forecast horizon. The benchmark model is a simple mean model. For each asset, we estimate various Bayesian VAR systems described sections 3 and 4. Values above zero indicate that a given predictor has better forecast performance than the benchmark model, while negative values suggest the opposite. All values are multiplied by 100. We test statistical significance in the average loss between the each model and a simple mean model using the Diebold and Mariano (1995) test. One/two/three asterisks denote rejection of the null hypothesis of equal predictive ability at the ten/five/one percent test level.

Table 12: Forecast performance in terms of average log predictive score differentials for 20 Dow Jones constituents (sample: 2004 - 2015) for various forecasting models

Stock	AA	AAPL	AIG	AXP	BA	CAT	KO	DD	GE	HD
AR1	**0.01**	-0.01	-0.46	**0.33**	**0.12**	-0.09	-0.06	-0.01	**0.05**	-0.05
VAR-Full	-0.91	-1.29	-2.29	**0.01**	-1.31	-1.24	-0.93	-1.78	-0.74	-0.41
VAR-Minnesota	**0.10**	**0.32**	**0.25**	**0.77**	**0.42**	**0.08**	**0.29**	**0.75**	**0.15**	**0.32**
TVPVAR-DMA	**0.21**	**0.74**	-0.11	**0.75**	**0.28**	**0.73**	**0.56**	**0.72**	**0.85**	**0.01**
TVPVAR-DMS	-0.06	**0.57**	-0.28	**0.49**	**0.27**	**0.47**	**0.44**	**0.71**	**0.63**	-0.03
TVPVAR-DMAm	**0.21**	**0.76**	-0.07	**0.76**	**0.32**	**0.77**	**0.60**	**0.77**	**0.87**	**0.03**
TVPVAR-DMAm/v	**0.80**	**1.83*****	**0.72**	**1.83*****	**0.54**	**1.31****	**0.62**	**1.56****	**1.52****	**0.96**
TVPVAR-DMSm	-0.02	**0.60**	-0.27	**0.53**	**0.30**	**0.51**	**0.48**	**0.73**	**0.66**	-0.01
TVPVAR-DMSm/v	**0.88**	**1.55****	**0.11**	**0.87**	**0.48**	**1.05***	**0.82**	**1.56****	**0.90**	**0.55**
Bayesian lasso	**0.23**	**0.84**	-0.04	**0.85**	**0.32**	**0.83**	**0.56**	**0.79**	**0.93**	**0.04**

Stock	INTC	IBM	JNJ	MCD	MRK	MSFT	PG	UTX	WMT	DIS
AR1	-0.09	-0.07	-0.08	-0.05	-0.09	-0.08	-0.04	-0.01	-0.01	**0.18**
VAR-Full	-1.35	-1.32	-0.73	-1.66	-2.09	-2.04	-0.29	-0.48	-0.06	-1.03
VAR-Minnesota	**0.60**	-0.03	**0.48**	**0.73**	**0.13**	**0.03**	**0.51**	**0.72**	**0.48**	**0.90**
TVPVAR-DMA	**0.86**	**0.43**	**0.68**	**0.69**	**0.74**	**0.08**	**0.42**	**0.61**	**0.92**	**1.02***
TVPVAR-DMS	**0.82**	**0.25**	**0.60**	**0.60**	**0.62**	-0.04	**0.30**	**0.42**	**0.87**	**0.96**
TVPVAR-DMAm	**0.89**	**0.45**	**0.72**	**0.73**	**0.77**	**0.09**	**0.45**	**0.62**	**0.95**	**1.04***
TVPVAR-DMAm/v	**1.95*****	**1.45****	**1.13***	**0.69**	**1.33****	**0.31**	**0.66**	**0.97**	**1.03***	**1.84*****
TVPVAR-DMSm	**0.85**	**0.29**	**0.64**	**0.60**	**0.66**	-0.04	**0.32**	**0.46**	**0.91**	**0.98**
TVPVAR-DMSm/v	**1.01***	**1.18***	**0.83**	**1.14***	**1.64****	**0.40**	**0.46**	**0.80**	**1.28****	**1.32****
Bayesian lasso	**0.96**	**0.51**	**0.76**	**0.75**	**0.82**	**0.11**	**0.44**	**0.64**	**0.98**	**1.10***

Note: The table provides forecast performance results in terms of average log predictive score differentials between the benchmark mean model and a single regressor model for 20 Dow Jones constituents (sample: 2004 - 2015) with a one month forecast horizon. For each asset, we estimate various Bayesian VAR systems described sections 3 and 4. Values above zero indicate that a given predictor has better forecast performance than the benchmark model, while negative values suggest the opposite. All values are multiplied by 100. We test statistical significance in the average loss between the each model and a simple mean model using the Diebold and Mariano (1995) test. One/two/three asterisks denote rejection of the null hypothesis of equal predictive ability at the ten/five/one percent test level.