Option-Implied Equity Premium Predictions via Entropic Tilting

Davide Pettenuzzo, Economics Department, Brandeis University

Konstantinos Metaxglou, Economics Department, Carleton University

Aaron Smith, Department of Agricultural & Resource Economics, University of California, Davis

Working Paper Series

2016 | 99R

Option-Implied Equity Premium Predictions via Entropic Tilting

Konstantinos Metaxoglou Davide Pettenuzzo Aaron Smith *

August 5, 2016

Abstract

We propose a new method to improve density forecasts of the equity premium using information from options markets. We obtain predictive densities from a state-of-the-art stochastic volatility (SV) model, which we then tilt towards the second moment of the risk-neutral distribution implied by options prices, while imposing a non-negativity constraint on the equity premium. By combining the *backward-looking* information contained in the SV model with the *forward-looking* information from options prices, our procedure delivers sharper predictive densities. Using density forecasts of the U.S. equity premium from January 1990 to December 2014, we find that tilting leads to more accurate predictions, both in terms of statistical and economic criteria.

JEL classification: C11, C22, G11, G12.

Keywords: entropic tilting, density forecasts, variance risk premium, equity premium, options.

---

*Metaxoglou: Department of Economics, Carleton University, konstantinos.metaxoglou@carleton.ca. Smith: Department of Agricultural and Resource Economics, University of California, Davis, adsmith@ucdavis.edu. Pettenuzzo: Brandeis University, dpettenu@brandeis.edu. The comments and suggestions of the editor Todd Clark, an anonymous associate editor, and two anonymous referees, significantly improved the paper. The usual disclaimer applies.

1 Introduction

Empirical asset pricing usually employs forecasting models that are *backward looking*—they use past observations on a set of variables to project future asset returns. The set of variables is often motivated by economic theory—for example, macroeconomic and financial variables, such as the dividend yield or the term spread. On the other hand, derivative prices convey information about the conditional density of future outcomes and, hence, are inherently *forward looking*. They contain information about market expectations, and thus should be useful for improving return forecasts.

In this paper, we provide a simple procedure to blend backward- and forward-looking information to sharpen the predictive densities of the equity premium obtained from a baseline econometric model. Our approach entails taking a given predictive density for excess returns and tilting it towards moments implied by options prices. Specifically, we proceed by extracting the variance of the risk-neutral distribution of returns from options prices, and subtracting from it a regression-based estimates of the variance risk premium to obtain a forward-looking variance estimate. In the spirit of Robertson et al. (2005), we then rely on entropic tilting to twist the original predictive distribution towards this forward-looking variance, while at the same time imposing a non-negativity constraint on the first moment. The latter constraint has been shown to substantially improve the out-of-sample (OOS) predictability of excess returns; see Campbell and Thompson (2008) and Pettenuzzo et al. (2014), among others. Our procedure is simple and has a low computational cost; using a few lines of code, we modify the original predictive density such that its moments conform with the additional restrictions we wish to impose.

To illustrate our method, we apply it to S&P 500 returns and to a set of industry-sorted portfolio returns. In both cases, we use a state-of-the-art stochastic-volatility (SV) model from the literature to form the baseline predictive density (Johannes et al., 2014). We find that tilting the baseline density using our procedure significantly improves the OOS predictability of stock returns both in terms of statistical and economic measures of forecasting accuracy.

Our paper contributes to a rapidly growing literature that looks at the role of option-implied information in improving forecasts. In particular, several papers show that option-implied volatility can predict future realized volatility as well as the equity premium; see for example, Szakmary et al.

(2003) and Bollerslev et al. (2009). We make two contributions to this literature. First, we provide a highly flexible non-parametric method for incorporating option-implied moments into baseline forecasts. Second, we work with density forecasts, whereas the bulk of the existing literature incorporates option-implied moments among the predictors in a point forecasting regression; see Altigan et al. (2015) for a recent example. Finally, it is worth noting that our method can be easily extended to higher moments, such as skewness and kurtosis, which have received increased attention recently in empirical asset pricing (Young Chang et al., 2013).

The remainder of the paper is organized as follows. In Section 2, we describe the entropic tilting procedure, along with our approach to constructing the model-based predictive densities for the equity premium. Our approach for removing the variance risk premium from the variance of the risk-neutral distribution implied by option prices follows. Section 3 presents our main results, and Section 4 focuses on OOS statistical and economic performance. Finally Section 5 provides some concluding remarks.

2 Entropic Tilting for Equity Premium Forecasting

Entropic tilting is a highly flexible non-parametric method to change the shape of a distribution to incorporate additional information about a random variable of interest. Such additional information may come in the form of moments and this is the approach we follow here. In what follows, we start from the predictive density implied by a state-of-the-art stochastic-volatility (SV) model from the asset pricing literature. We then use entropic tilting to alter this baseline distribution to incorporate moment restrictions derived from options prices and economic theory.

We begin by first outlining the general entropic tilting method and our approach to incorporating the moment-based information from the options markets into a baseline predictive density. Next, we describe the econometric model we use to produce the baseline density forecasts. We conclude this section by describing our approach to removing the variance risk premium from the risk-neutral variance we derive from option prices.

2.1 General Method

Let $p\left(r_{t+1}|\mathcal{D}^t\right)$ denote the baseline predictive density for the equity premium r_{t+1} with $\mathcal{D}^t$ being the information set available at time t, and $t = 1, ..., T-1$. The econometrician is assumed to have additional information about a function $g(r_{t+1})$, which was not used to generate the baseline predictive density. This additional information takes the form of moments of $g(r_{t+1})$ such that

$$E[g(r_{t+1})|\mathcal{D}^t] = \overline{g}_t. \tag{1}$$

For example, $g(r_{t+1})$ may represent quantities such as the mean, $g(r_{t+1}) = r_{t+1}$, the variance, $g(r_{t+1}) = (r_{t+1} - E[r_{t+1}])^2$, or higher moments of the predictive distribution; see Robertson et al. (2005) for a very informative exposition. The information could be in the form of moment restrictions implied by economic theory, such as Euler conditions in Giacomini and Ragusa (2014), or could be coming from survey forecasts and model-based nowcasts as in Altavilla et al. (2014) and Krüger et al. (2015).

Generally, the expected value of $g(r_{t+1})$ under the baseline distribution will not equal $\overline{g}_t$

$$\int g(r_{t+1})p\left(r_{t+1}|\mathcal{D}^t\right) dr_{t+1} \neq \overline{g}_t. \tag{2}$$

Thus, by transforming $p\left(r_{t+1}|\mathcal{D}^t\right)$ so that (1) holds, we sharpen the baseline predictive density. To implement the method, consider N random draws from the baseline predictive distribution $p\left(r_{t+1}|\mathcal{D}^t\right)$. We denote these draws with $\left\{r_{t+1}^i\right\}_{i=1}^N$, where each draw is associated with a weight $\pi_i = 1/N$. We construct a new set of weights $\{\pi_{it}^*\}_{i=1}^N$ that represent a new predictive density that is as close as possible to the baseline and also satisfies the moment restriction implied by (1). Following a standard approach in the literature, we use the empirical Kullback-Leibler Information Criterion (KLIC) to measure the distance between the baseline and the new predictive density[1]

$$KLIC(\pi_t^*; \pi) = \sum_{i=1}^N \pi_{it}^* \ln\left(\frac{\pi_{it}^*}{\pi_i}\right). \tag{3}$$

[1]Other measures of divergence are also available. As Giacomini and Ragusa (2014) note, the KLIC provides a convenient analytical expression for the tilted weights and, unlike other measures of distance, it has a direct counterpart in the logarithmic scoring rule, which is a common and well-studied measure for evaluating density forecasts (Amisano and Giacomini, 2007).

The objective is to find new weights that minimize (3) subject to the constraints

$$\pi_{it}^{*} \geq 0, \qquad \sum_{i=1}^{N} \pi_{it}^{*} = 1, \qquad \sum_{i=1}^{N} \pi_{it}^{*} g(r_{t+1}^{i}) = \overline{g}_{t}, \tag{4}$$

where the last constraint may be viewed as the Monte-Carlo approximation to the moment restriction in (1) using the language in Cogley et al. (2005).[2] The implied first-order conditions are given by

$$1 + \ln\left(\frac{\pi_i}{\pi_{it}^{*}}\right) - \mu_t - \gamma_t' g(r_{t+1}^{i}) = 0, \quad i = 1, \dots, N \tag{5}$$

with μ_t and γ_t being the Lagrange multipliers associated with the adding-up and moment constraints. The new weights are then given by

$$\pi_{it}^{*} = \frac{\pi_i \exp(\gamma_t^{*\prime} g(r_{t+1}^{i}))}{\sum_{i=1}^{N} \pi_i \exp(\gamma_t^{*\prime} g(r_{t+1}^{i}))}. \tag{6}$$

As a result, the baseline weights are *tilted* in an exponential fashion via (6) to generate the new weights. The tilting parameter γ_t^{*} can be found by solving the minimization problem

$$\gamma_t^{*} = \arg\min_{\gamma_t} \sum_{i=1}^{N} exp(\gamma_t' [g(r_{t+1}^{i}) - \overline{g}_t]). \tag{7}$$

In our case, we use the variance of the risk-neutral distribution for the equity premium, as implied by the option markets, to distort the baseline predictive distribution $p\left(r_{t+1}|\mathcal{D}^t\right)$ so that its dispersion, as captured by $Var\left(r_{t+1}|\mathcal{D}^t\right)$, resembles that of the option implied risk-neutral distribution. It is the forward-looking aspect of the options market that serves as the source of new information and is also the novelty in our approach. In addition, we follow the recent literature on stock return predictability (e.g., Campbell and Thompson, 2008; Pettenuzzo et al., 2014) and further impose a non-negativity constraint on the first moment of the tilted predictive density. In the spirit of Robertson et al., we incorporate restrictions that could in principle be built directly into the forecasting model in a manner that is less demanding computationally.

[2] See Robertson et al. and the references therein. The 2012 Econometric Reviews Special Issue on Entropy and the 2002 Journal of Econometrics Issue on Information and Entropy Econometrics offer more detailed treatments on entropy and the use of alternative divergence measures.

2.2 Baseline Predictive Densities

There is ample evidence pointing to time variation in both the conditional mean and volatility of the return distribution; see, for example, Rapach and Zhou (2013) and Andersen et al. (2006). Accordingly, we rely on the following model with time-varying first and second moments to produce the baseline predictive density $p\left(r_{t+1}|\mathcal{D}^t\right)$ of the monthly equity premium

$$r_{\tau+1} = \mu + \beta' x_\tau + \exp\left(h_{\tau+1}\right) u_{\tau+1}, \quad \tau = 1, ..., t-1, \tag{8}$$

where $h_{\tau+1}$ represents the log-volatility at time $\tau+1$, x_τ denotes a (vector of) lagged predictor(s), and $u_{\tau+1} \sim \mathcal{N}(0,1)$. We further assume that the log-volatility $h_{\tau+1}$ follow a stationary and mean-reverting process and depends on lagged intra-month information in the form of realized volatility

$$h_{\tau+1} = \lambda_0 + \lambda_1 h_\tau + \lambda_2 RV_\tau + \xi_{\tau+1}, \quad \xi_{\tau+1} \sim \mathcal{N}\left(0, \sigma_\xi^2\right), \tag{9}$$

where RV_τ denotes the realized volatility at time τ, computed by summing the squared daily returns within month τ, $|\lambda_1| < 1$. Note also that u_τ and ξ_s are mutually independent for all τ and s.

We estimate the parameters in (8) using Bayesian methods. Following standard practice in the Bayesian literature (Koop, 2003), the priors for μ and β in (8) are assumed to be normal

$$\begin{bmatrix} \mu \\ \beta \end{bmatrix} \sim N\left(\underline{b}, \underline{V}\right). \tag{10}$$

For the hyperparameters $\underline{b}$ and $\underline{V}$, we set aside an initial training sample of t_0 observations to calibrate them (e.g., Primiceri, 2005; Clark, 2011) and proceed as follows

$$\underline{b} = \begin{bmatrix} \overline{r}_{t_0} \\ 0 \end{bmatrix}, \quad \underline{V} = \underline{\psi}^2 \left[s_{r,t_0}^2 \left(\sum_{\tau=1}^{t_0-1} x_\tau x_\tau' \right)^{-1} \right], \tag{11}$$

where

$$\overline{r}_{t_0} = \frac{1}{t_0-1} \sum_{\tau=1}^{t_0-1} r_{\tau+1}, \quad s_{r,t_0}^2 = \frac{1}{t_0-2} \sum_{\tau=1}^{t_0-1} \left(r_{\tau+1} - \overline{r}_{t_0}\right)^2.$$

Our choice of $\underline{b}$ in (11) reflects the prior belief that the best predictor of stock returns is the average of past returns. Therefore, we center the prior intercept on the historical average of the excess returns, while we set the prior mean on the slope coefficient(s) to zero. Furthermore, the scalar $\underline{\psi}$ in (11) controls the tightness of the prior ($\underline{\psi} \rightarrow \infty$ corresponds to a diffuse prior on μ and β). We specify rather uninformative priors and set $\underline{\psi} = 1.0e^6$.

We also require priors on the sequence of volatilities, $h^t = \{h_1, ..., h_t\}$, and the SV parameters λ_0, λ_1, λ_2, and σ_ξ^2. Decomposing the joint probability of these parameters and using (9), we have

$$\begin{aligned} p\left(h^t, \lambda_0, \lambda_1, \lambda_2, \sigma_\xi^{-2}\right) &= p\left(h^t \middle| \lambda_0, \lambda_1, \lambda_2, \sigma_\xi^{-2}\right) p\left(\lambda_0, \lambda_1, \lambda_2\right) p\left(\sigma_\xi^{-2}\right) \\ &= \prod_{\tau=1}^{t-1} p\left(h_{\tau+1} \middle| \lambda_0, \lambda_1, \lambda_2, h_\tau, \sigma_\xi^{-2}\right) p\left(h_1\right), \end{aligned} \tag{12}$$

where

$$h_{\tau+1} | \lambda_0, \lambda_1, \lambda_2, h_\tau, \sigma_\xi^{-2} \sim \mathcal{N}\left(\lambda_0 + \lambda_1 h_\tau + \lambda_2 RV_\tau, \sigma_\xi^2\right). \tag{13}$$

To complete the prior elicitation for $p\left(h^t, \lambda_0, \lambda_1, \lambda_2, \sigma_\xi^{-2}\right)$, we choose priors for λ_0, λ_1, λ_2, the initial log volatility h_1, and σ_ξ^{-2}, from the normal-gamma family

$$h_1 \sim \mathcal{N}\left(\ln\left(s_{r,t_0}\right), \underline{k}_h\right), \tag{14}$$

$$\begin{bmatrix} \lambda_0 \\ \lambda_1 \\ \lambda_2 \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \underline{m}_{\lambda_0} \\ \underline{m}_{\lambda_1} \\ \underline{m}_{\lambda_2} \end{bmatrix}, \begin{bmatrix} \underline{V}_{\lambda_0} & 0 & 0 \\ 0 & \underline{V}_{\lambda_1} & 0 \\ 0 & 0 & \underline{V}_{\lambda_2} \end{bmatrix} \right), \quad \lambda_1 \in (-1, 1), \tag{15}$$

and

$$\sigma_\xi^{-2} \sim \mathcal{G}\left(1/\underline{k}_\xi, \underline{v}_\xi\left(t_0 - 1\right)\right). \tag{16}$$

We set $\underline{k}_\xi = 0.1$, $\underline{v}_\xi = 10$, and $\underline{k}_h = 10$. These choices restrict changes to the log-volatility to be roughly equal to 0.3, on average, and place a relatively diffuse prior on the initial log-volatility state.

Following Clark and Ravazzolo (2015), the hyperparameters are as follows: $\underline{m}_{\lambda_0} = \underline{m}_{\lambda_3} = 0$,

$\underline{m}_{\lambda_1} = 0.9$, $\underline{V}_{\lambda_0} = \underline{V}_{\lambda_3} = 0.25$, and $\underline{V}_{\lambda_0} = 1.0e^{-4}$. This corresponds to setting the prior means and standard deviations for the intercept and RV coefficient to 0 and 0.5, respectively. As for the AR(1) coefficient, these choices imply a prior mean of 0.9 with a standard deviation of 0.01. Overall, these are informative priors that match the persistent dynamics in the log volatility process.

We estimate the model in (8)-(9) using a Gibbs sampler that lets us compute posterior draws for μ, β, h^t, σ_ξ^{-2}, λ_0, λ_1, and λ_2. These draws are used to compute density forecasts for r_{t+1}

$$\begin{aligned} p\left(r_{t+1} \middle| \mathcal{D}^t\right) &= \int p\left(r_{t+1} \middle| h_{t+1}, \Theta, h^t, \mathcal{D}^t\right) \\ &\quad \times p\left(h_{t+1} \middle| \Theta, h^t, \mathcal{D}^t\right) p\left(\Theta, h^t \middle| \mathcal{D}^t\right) d\Theta dh^{t+1}. \end{aligned} \tag{17}$$

where $\Theta = \left(\mu, \beta, \sigma_\xi^{-2}, \lambda_0, \lambda_1, \lambda_2\right)$ contains the time-invariant parameters. The Online Appendix of the paper contains details on the Gibbs sampler and the computation of the integral in (17).

2.3 Removing the Variance Risk Premium

We capitalize on the literature that has demonstrated the predictive power of implied volatility for future realized volatility; see Jorion (1995) and, more recently, Szakmary et al. (2003), among others. The basic argument is that implied volatility—inferred from options data as in our case—can be perceived as the market's expectation of future volatility and, hence, it is a market-based volatility forecast (Poon and Granger, 2003). The feature of the implied volatility that is particularly appealing for a forecasting exercise like the one undertaken here is that it is inherently forward-looking.[3]

In the presence of a variance risk premium, the implied or risk-neutral variance is a biased estimate of the variance of the physical predictive density. Economic agents dislike the uncertainty of future variance and, in equilibrium, command a premium for accepting this risk, which gives rise to the variance risk premium. Bollerslev et al. (2009) provides strong evidence of variance risk premia in financial assets. Thus, we first remove the variance risk premium from the risk-neutral variance

[3] Implied volatility reflects options traders' judgment about short-term volatility, due in part to information such as forthcoming announcements (e.g., an upcoming election, macroeconomic data releases) known to market participants but not to the econometrician. It resembles the judgmental component of Blue Chip, the Survey of Professional Forecasters, and the Greenbook surveys in forecasting inflation (among other macroeconomic series) as in Faust and Wright (2013).

before tilting the baseline predictive density $p\left(r_{t+1}|\mathcal{D}^t\right)$ towards it.

Let $\hat{u}_{\mathbb{P},t+1}$ denote the forecast error from the baseline physical predictive distribution at time $t+1$ obtained following the approach in Section 2.2

$$\widehat{u}_{\mathbb{P},t+1} = r_{t+1} - E\left(r_{t+1}|\mathcal{D}^t\right), \tag{18}$$

where $E\left(r_{t+1}|\mathcal{D}^t\right)$ is the posterior mean of $p\left(r_{t+1}|\mathcal{D}^t\right)$. The posterior variance of the predictive distribution is $\sigma^2_{\mathbb{P},t+1} \equiv Var\left(r_{t+1}|\mathcal{D}^t\right)$. From options prices, we can compute the variance of the risk-neutral distribution, $\sigma^2_{\mathbb{Q},t}$, which differs from $\sigma^2_{\mathbb{P},t+1}$ by the variance risk premium VRP_{t+1}

$$\sigma^2_{\mathbb{P},t+1} = \sigma^2_{\mathbb{Q},t} - VRP_{t+1}. \tag{19}$$

We assume that the variance risk premium is such that the following holds

$$\log\left(\sigma^2_{\mathbb{P},t+1}\right) = \alpha + \beta \log\left(\sigma^2_{\mathbb{Q},t}\right). \tag{20}$$

Because the log squared forecast error is a noisy measure of $\log \sigma^2_{\mathbb{P},t+1}$, we can estimate α and β using a regression of $\log\left(\widehat{u}^2_{\mathbb{P},\tau+1}\right)$ on $\log\left(\sigma^2_{\mathbb{Q},\tau}\right)$, where $\tau = 1, ..., t-1$. Thus, we tilt the predictive distribution towards a variance given by

$$\hat{\sigma}^2_{\mathbb{P},t+1} = \exp\left(\hat{\alpha} + \hat{\beta}\log\left(\sigma^2_{\mathbb{Q},t}\right)\right), \tag{21}$$

which implies that the variance risk premium is

$$\begin{aligned} VRP_{t+1} &= \sigma^2_{\mathbb{Q},t+1} - \exp\left(\alpha + \beta\log\left(\sigma^2_{\mathbb{Q},t}\right)\right) \\ \widehat{VRP}_{t+1} &= \sigma^2_{\mathbb{Q},t+1} - \exp\left(\hat{\alpha} + \hat{\beta}\log\left(\sigma^2_{\mathbb{Q},t}\right)\right). \end{aligned} \tag{22}$$

An alternative and more computationally demanding approach to incorporate forward-looking information into return forecasts would be to adapt a GARCH-type model such as the MEM of Engle and Gallo (2006), the HEAVY of Shephard and Sheppard (2010), or the realized GARCH of Hansen et al. (2012). This adaptation would entail replacing realized volatility with a measure of implied volatility and developing an approach for handling the variance risk premium.[4]

3 Empirical Results

We obtain the data necessary to generate the density forecasts in (17) from Goyal and Welch (2008) and Rapach et al. (2016). End-of-month stock returns are computed from the S&P500 index and include dividends. A short T-bill rate is subtracted from stock returns to obtain the monthly excess returns. Overall, we consider a set of 14 popular predictor variables from Goyal and Welch, augmented with the short interest index (SII) introduced by Rapach et al.[5] Our sample starts in January 1973 ($t = 1$) and extends to December 2014 ($t = T$), as in Rapach et al.

We begin by computing the baseline predictive densities for the equity premium using (17) and, one by one, all 15 of the predictors considered. To explicitly denote the dependence of the predictive density in (17) on predictor i, we write $p\left(r_{t+1}|\mathcal{M}_i, \mathcal{D}^t\right)$, where $i = 1, ..., K$ and K=15. Next, to generate the predictive densities, we start in January 1990 (this is the first month when the VIX series becomes available) and proceed in a recursive fashion using an expanding-window approach until the last observation in the sample.[6] This process yields 15 time series of one-step-ahead density forecasts—one for each predictor—between January 1990 and December 2014.

3.1 Moments

To assess the degree of time variation in the excess return volatility implied by our econometric model, the top panel of Figure 1 plots the monthly excess return volatility implied by the SV model in (8)-(9) over the full estimation sample, January 1973 to December 2014. We plot the volatility series for a single predictor, the earnings-price ratio (EP), noting that the series provided here are very similar across the 15 predictors. The blue solid line depicts the posterior mean of

[4]See Table 1 in Hansen et al. for a succinct comparison of the three types of models. We thank an anonymous referee for suggesting this alternative approach.

[5]The data on the monthly market returns, risk-free rate, and the Goyal and Welch predictors, are available from Amit Goyal's website, updated and extended to December 2014, at `http://www.hec.unil.ch/agoyal/`. The SII data are available at `http://sites.slu.edu/rapachde/home/research`. For a detailed discussion of the predictors considered, see Goyal and Welch and Rapach et al.

[6]Accordingly, we set aside the data from January 1973 to December 1989 to train the priors in (10), (14), and (16). Hence, we set $t_0 = 204$.

$\exp(h_t)$, $t = 1, ..., T$, while the dashed red lines show the 5th and 95th percentiles of its posterior distribution. The volatility hovers around 5% per month, but it spikes in 1975, after October 1987, and during the recent financial crisis.

As a comparison, the bottom panel of Figure 1 provides a time series plot of the end-of-month values of the Chicago Board Options Exchange (CBOE) Volatility Index (VIX). We use VIX to summarize the risk-neutral volatility of the S&P 500 returns, that is, $\sqrt{\sigma^2_{\mathbb{Q},t+1}}$. In 1993, the CBOE introduced VIX, originally designed to measure the market's expectation of 30-day volatility implied by ATM S&P 100 Index (OEX) option prices. In 2003, CBOE together with Goldman Sachs updated the methodology and formula for VIX. The new VIX is based on the S&P 500 Index (SPX) and estimates expected volatility by averaging the weighted prices of SPX puts and calls over a wide range of strikes and its values are available back to January 1990. Setting aside the very prominent spike in October 2008 around the peak of the most recent financial crisis, the risk-neutral variance is highest during 1997–2003, a period of well-documented turmoil in financial markets (Bloom, 2009). Events during this period include the Asian crisis (Fall 1997), the Russian Financial Crisis (Fall 1998) September 11 (Fall 2001), the Enron and WorldCom scandals (Summer/Fall 2002), and Gulf War II (Spring 2003).

3.2 Entropic Tilting

We use the entropic tilting approach described in Section 2.1 to modify each of the baseline predictive densities, such that their variances match the corresponding option implied risk-neutral variance—adjusted using (21) to remove the variance risk premium—and their means are non-negative. Setting aside the period January 1990 to December 1994 to estimate the first variance risk premium, our final OOS period is January 1995 to December 2014.[7]

In Figure 2, we show the first two moments for the tilted and the baseline distributions over the OOS period for the model in which EP is the predictor. Starting with the top panel, the two mean

[7]We obtain our estimates for the $i-$th variance risk premium by regressing the log squared forecast error implied by the $i-$th baseline predictive density, $p(r_{\tau+1}|\mathcal{M}_i, \mathcal{D}^\tau)$, on the the log squared VIX using (20) and an expanding-window approach, where $\tau = t_0, ..., t-1$. Thus, the estimated variance risk premium for each forecast month comes from a regression using data from January 1990 through the previous month. The slope parameter $\hat{\beta}$ in (20) averages about 1.25 in these regressions. We refer the reader to the Online Appendix for additional details.

series track each other very closely, except for the early part of 2009 when we see a big dip in the mean of the baseline return distribution but not in its tilted counterpart, an immediate consequence of the fact that we impose the non-negativity restriction on the mean of the tilted distribution. The bottom panel shows that the volatility of the baseline distribution generally exceeds that of the tilted distribution, except for 2000–2003 and 2008–2011. For example, in December of 2008, the volatility of the tilted distribution is close to 0.15 while its baseline counterpart is around 0.07.

Overall, the shape of the tilted predictive densities exhibits much more variation over time compared to its baseline counterpart. For example, if we focus on the far left tail of the distributions, the 1% quantile for the baseline density forecasts is between -0.176 and -0.112, while that for the tilted density forecasts is between -0.365 and -0.032. In the case of the far right tail of the distributions, the 99% quantile for the baseline density forecasts is 0.123–0.171, while that for the tilted density forecasts is 0.042–0.406. Similar conclusions are drawn by looking at the shoulders of the two distributions.

The empirical KLIC defined in (3) gauges how much the baseline density is altered by the tilting procedure. That is, small values of the empirical KLIC signify agreement between the baseline predictive model and outside information, while large values signify disagreement.[8] As a practical matter, large discrepancies also serve as warnings about the accuracy of statistics computed from the tilted densities. In fact, a large KLIC value implies that the distribution of the weights is highly skewed, with many draws from the baseline density being ignored and a few draws becoming highly influential. The average KLIC for the entire OOS period is 0.17, which is within the range of KLIC values reported in Cogley et al. (2005) and Robertson et al. (2002), 0.12–0.68 and 0.06–0.66, respectively.[9] One of the three examples in Robertson et al. uses an intertemporal consumption-CAPM to add moment restrictions on a VAR forecasting real consumption growth and interest rates. This is the example that gives rise to the largest KLIC value reported in their paper (0.66), and—according to Cogley et al.—these values can serve as benchmark for aggressive twisting given that the consumption-CAPM is known to fit the data poorly. In our case, although the KLIC achieves some of its largest value in 1995–1996, 2004–2007, and also in 2009, its annual averages

[8]Note that when $KLIC(\pi_t^*; \pi)$ is zero, it means that the baseline and tilted densities coincide.

[9]The ranges reported here are based on Table 2 in Cogley et al., and on the KLIC statistics reported in Tables 1b, 2b, and 3b in Robertson et al..

never exceed 0.6 in any of these years. In summary, it appears that for the largest part of out sample, the twisting of the baseline distributions is not excessively aggressive.

4 Out-of-Sample Performance

In this section, we examine whether the approach introduced in Section 2 leads to more accurate equity premium forecasts, both in terms of statistical and economic criteria. As with previous studies, such as Goyal and Welch (2008) and Campbell and Thompson (2008), we measure the predictive accuracy relative to the historical average (HA) model, which assumes constant expected excess returns. However, since all the models we consider in this study allow for time-varying volatility, we augment the HA model to also include this feature and label it HA-SV. The reader should note that the HA-SV benchmark corresponds to the model in (8)–(9) when $\beta = 0$.[10]

4.1 Statistical Forecasting Performance

We consider several evaluation metrics for both point and density forecasts. Starting with point-forecast accuracy, we follow Campbell and Thompson and summarize the predictive ability of the various models over the whole evaluation period by reporting the OOS R-squared for the forecasting model associated with each predictor k

$$R^2_{OOS,kd} = 1 - \frac{\sum_{\tau=m+1}^{T} e^2_{kd,\tau}}{\sum_{\tau=m+1}^{T} e^2_{bcmk,\tau}}, \tag{23}$$

where $m+1$ denotes the beginning of the forecast evaluation period (January 1995) and $bcmk$ refers to the HA-SV benchmark. The additional subscript $d \in \{baseline, tilted\}$ allows us to distinguish between the baseline and the tilted densities associated with predictor k. Furthermore, $e_{kd,\tau}$ and $e_{bcmk,\tau}$ denote the time τ forecast error for the baseline or tilted, and the HA-SV benchmark densities, respectively. We obtain point forecasts to compute the forecast errors in (23) by averaging over the draws from the corresponding predictive densities. A positive $R^2_{OOS,kd}$ indicates that the

[10]For consistency, the HA-SV model is estimated using priors analogous to those we used with the various predictors. In particular, we slightly alter the prior on (μ, β) to impose a dogmatic "no predictability" prior on $\beta = 0$, while using the same priors for $h^t, \lambda_0, \lambda_1, \lambda_2, \sigma_\xi^{-2}$.

point forecast associated with the baseline or tilted densities for predictor k is more accurate than the HA-SV benchmark forecast.

To quantify the accuracy of density forecasts, we follow Amisano and Giacomini (2007) and report the average log score difference

$$ALSD_{kd} = \frac{1}{T-m} \sum_{\tau=m+1}^{T} LS_{kd,\tau} - LS_{bcmk,\tau}, \tag{24}$$

where $LS_{kd,\tau}$ and $LS_{bcmk,\tau}$ denote the time-τ log predictive scores of the baseline or tilted densities, and the HA-SV predictive density, respectively. The logarithmic score gives a high value to a predictive density that assigns a high probability to the event that actually occurred. Hence, a positive $ALSD_{kd}$ value indicates that, on average, the SV model with predictor k is more accurate than the HA-SV benchmark in predicting the outcome of interest.

To test the statistical significance of differences in point and density forecasts, we consider Diebold and Mariano (1995) (DM) tests of equal predictive accuracy using mean squared forecast errors (MSFEs) and average log scores (ALSs), respectively. We perform two DM tests. First, we test whether the improvements in the MSFEs or the ALSs for the baseline densities relative to their HA-SV benchmark counterparts are statistically significant. Second, we test whether the improvements in the MSFEs or the ALSs for the tilted densities relative to their baseline counterparts are statistically significant. In both cases, we use standard normal critical values and incorporate the finite sample correction due to Harvey et al. (1997).[11]

The top panel of Table 1 pertains to point forecasts. Columns (1) and (2) of the table report the R^2_{OOS} associated with the baseline and tilted density forecasts for all 15 predictors considered over the full OOS period, January 1995–December 2014; the remaining columns report the R^2_{OOS} values for the earlier (January 1995–December 2006) and later (January 2007–December 2014) parts of the OOS period. For example, the dividend-price ratio (DP) produces an R^2_{OOS} of -1.287 in the case of the baseline forecasts and an R^2_{OOS} of -1.103 in the case of the tilted forecasts for the full OOS period. The bold entry in column (2) indicates that the tilted forecasts perform better than the

[11] Citing Monte Carlo evidence in Clark and McCracken (2011), with nested models, Clark and Ravazzolo (2015) argue that the DM test with normal critical values is a somewhat conservative test—has sizes that tend to fall below the nominal—for equal accuracy in finite samples.

baseline forecasts in terms of MSFEs generating a higher R^2_{OOS}. Furthermore, the two asterisks (**) next to the same entry indicate that the tilted forecasts are significantly better than their baseline counterparts at the 5% using a DM test for their MSFE differences. The lack of an asterisk next to the entry in column (1) for the same predictor indicates that the baseline forecasts fail to be significantly better than the HA-SV benchmark forecasts. Analogous notational conventions hold for the other combinations of predictors and OOS periods in the table.

In the case of baseline point forecasts for the full OOS period, we observe negative R^2_{OOS} values for all predictors except the short interest index (SII), whose superior forecast accuracy relative to the HA-SV benchmark is significant at the 10% level. These results are consistent with the findings of Rapach et al. despite the fact that our OOS period is slightly shorter. Although tilting leads to R^2_{OOS} improvements for 9 out of 15 predictors (3 of them are significant at 5%), it fails to produce positive R^2_{OOS} values for all predictors except for SII. However, it appears that tilting has a mild negative impact on the overall point-forecast predictability for this particular predictor; note the lower R^2_{OOS} value for the tilted forecasts.

The bottom panel of Table 1 reports the ALSDs for the baseline and tilted density forecasts for each of the 15 predictors we considered. Over the full OOS period, SII is the only predictor for which the ALSD associated with the baseline density forecasts is positive. However, this single positive value is not statistically significant at conventional levels. The remaining ALSDs lie between -0.009 for DP and -0.001 for the three-month T-bill rate (TBL). Our tilting procedure delivers a substantial improvement in the ALSDs for all predictors. The resulting improvement is statistically significant at the 1% level on the basis of the DM test on the ALSs. We now see ALSDs ranging between 0.128 for net equity expansion (NTIS) and 0.151 for SII.

To see how point-forecast performance changes over time, we compute the cumulative sum of squared forecast error difference (CSSED)

$$CSSED_{kd,t} = \sum_{\tau=m+1}^{t} \left(e^2_{bcmk,\tau} - e^2_{kd,\tau}\right), \ t = m+1, ..., T. \tag{25}$$

A positive $CSSED_{kd,t}$ indicates that the point forecasts associated with the baseline or tilted predictive densities for predictor k are more accurate than their benchmark HA-SV counterparts up to time t. We also examine how the density-forecast performance changes over time using the cumulative log-score difference (CLSD)

$$CLSD_{kd,t} = \sum_{\tau=m+1}^{t} LS_{kd,\tau} - LS_{bcmk,\tau},\ t = m+1, ..., T. \tag{26}$$

If the baseline or tilted density forecast for predictor k is more accurate than the HA-SV benchmark during the entire OOS period, then the corresponding CLSD line is monotonically increasing. Conversely, episodes with the density forecasts of predictor k being less accurate than the HA-SV benchmark generate dips in the CLSD line.

The top panel of Figure 3 plots the CSSED associated with the baseline and tilted density forecasts for EP and illustrates the role of the non-negativity constraint in our tilting procedure. For EP, as the top panel of Figure 2 shows, the point forecasts turn negative for a short period of time around the latest financial crisis. As a result, the baseline and tilted CSSEDs are essentially identical up until the end of 2008. Right around that time, the onset of the latest financial crisis leads to a large positive shock in predictability for both the baseline and tilted models. On the one hand, in the case of the baseline model, this positive shock is very short-lived, and by the mid-2009 the baseline CSSED turns negative. On the other hand, thanks to the non-negativity constraint we impose on the first moment of the tilted density, the tilted CSSED line remains significantly above its baseline counterpart for the remainder of the OOS period.

The bottom panel of Figure 3 plots the CLSDs for the EP baseline and tilted densities. The tilted CLSD line lies above the baseline one throughout the OOS period with a clear upward trend right at the beginning of the OOS period, and then again during 2003–2007. The baseline CLSD line remains very close to zero exhibiting a very slight downward trend.

We conclude this section by investigating the stability over time of the point and density forecasts. In particular, we perform two separate analysis. First, we separately report the R^2_{OOS} and ALSD statistics for two different parts of the OOS period. The first part, January 1995–December 2006 (OOS I), predates the global financial crisis. The second part, January 2007–December 2014 (OOS II), surrounds the recent crisis. We also report the results of the Giacomini and Rossi (2010) fluctuation test for the baseline and tilted density forecasts in terms of mean squared forecast error differences (MSFEDs) and ALSDs.

We begin with the results in columns (4)–(7) of Table 1. Similar to the full OOS results, tilting leads to improvements in R^2_{OOS} for both OOS periods I and II. The gains are generally larger for OOS period I. The tilted R^2_{OOS} values are higher than their benchmark counterparts for 9 out of the 15 predictors; the improvements in 5 cases are statistically significant. Turning to period OOS II, we notice that the baseline point forecasts imply a positive R^2_{OOS} value in several instances, ranging between 0.116 for the long-term rate of return (LTR) and 4.198 for SII. In this case, the improvement due to tilting the first moment of the baseline densities is limited to 5 predictors only, with the most notable one being the long-term yield (LTY), where we see an increase from 0.198 to 0.280 that is statistically significant at the 5% level. Moving to the bottom panel of Table 1, we find that the effect of tilting on the density forecast accuracy for periods OOS I and OOS II is highly comparable to the full OOS period. In every instance, the ALSD values for the tilted densities are higher than their baseline counterparts, and in all instances the differences are statistically significant at the 1% level.

Next, Figure 4 plots the results of the Giacomini-Rossi (GR) fluctuation test for the baseline and tilted densities, both in terms of MSFEDs and ALSDs. For the baseline densities, we test the null hypothesis that the baseline and the HA-SV densities have equal predictive performance over every 5-year centered window in our OOS period. The alternative hypothesis is that the baseline densities perform better. For the tilted densities, we test the null hypothesis that the tilted and baseline densities have equal performance. The alternative hypothesis is that the tilted densities perform better. We test these hypotheses for each of the 15 predictors. As a result, the maximum number of rejections reported is 15.

If the forecasting performance is stable over time, we expect the rejection rate to be relatively constant over time. Starting with the two left panels, which use the MSFED metric, we see that for the baseline predictive densities (top left panel) we reject the null hypothesis only for a few predictors, mostly during 2006–2012. For the same metric, we see mostly between 1 and 5 rejections for the tilted densities (bottom left panel), clustering around 1998–2006 and 2008–2012.

Using the ALSD metric, we almost never reject the null hypothesis for the baseline densities (top right panel), except for a single predictor for a short period around 2010. For the tilted densities (bottom right panel), we consistently reject the null for almost all predictors during 2002–2009 and 2011–2012. We do not see rejections prior to 2002 and for a good part of 2010. This finding mirrors the bottom panel of Figure 3, which shows that the CLSD is quite flat until 2002. In sum, the improvement in density forecasts is consistently strong in the final two thirds of the OOS period and is not driven by a few isolated events.

4.2 Economic Performance

Up to this point, we have focused on the statistical performance of the baseline and tilted predictive densities. In this section, we turn to their economic performance. We posit a representative investor using these predictions to make optimal portfolio decisions, taking parameter uncertainty into consideration.[12]

In particular, our interest lies in the optimal asset allocation of a representative investor facing a utility function $U(\omega_{t-1}, r_t)$ with ω_{t-1} denoting the share of her portfolio allocated into risky assets, and r_t being time t equity premium.[13] The representative agent solves the optimal asset allocation problem

$$\omega_{t-1}^{*} = \arg\max_{\omega_{t-1}} E\left[U\left(\omega_{t-1}, r_t\right) \middle| \mathcal{D}^{t-1} \right], \tag{27}$$

with $t = m+1, ..., T$. She is assumed to have power utility of the form

$$U\left(\omega_{t-1}, r_t\right) = \frac{\left[\left(1-\omega_{t-1}\right) \exp\left(r_{f,t-1}\right) + \omega_{t-1} \exp\left(r_{f,t-1} + r_t\right)\right]^{1-A}}{1-A}, \tag{28}$$

where $r_{f,t-1}$ is the continuously compounded T-bill rate available at time $t-1$, and A is the coefficient of relative risk aversion. The subscript $t-1$ on the portfolio implies that the investor solves the optimization problem using information available only at time $t-1$. The power utility

[12] Our discussion follows closely Kandel and Stambaugh (1996) and Barberis (2000). Parameter uncertainty is accounted for in the Bayesian framework because the parameter posterior distribution is integrated out of the predictive density of returns (see equation (17)).

[13] Given the availability of density forecasts as opposed to just point forecasts, we are not restricted to rely on a mean-variance utility function, and we can focus on functions with better properties such as the power utility. The power utility avoids the major limitation of the mean-variance utility, namely, that investors care only about the first two moments of returns. Furthermore, it is well known that mean-variance portfolio optimization is consistent with expected utility maximization only under special circumstances. Sufficient conditions include quadratic utility or elliptical return distributions. See, for example, Back (2010).

function exhibits the useful property of constant relative risk aversion (CRRA). Moreover, the optimal portfolio weights do not depend on initial wealth.

Taking expectations with respect to the predictive density of r_t, we can rewrite (27) as follows

$$\omega_{t-1}^{*} = \arg\max_{\omega_{t-1}} \int U\left(\omega_{t-1}, r_t\right) p\left(r_t | \mathcal{D}^{t-1}\right) dr_t. \tag{29}$$

The integral in (29) can be approximated using draws from the competing predictive densities. Specifically, using the HA-SV predictive density, we can approximate the solution to (29) using a large number (J) of draws, $\left\{r_{bcmk,t}^{j}\right\}_{j=1}^{J}$, and the following expression

$$\widehat{\omega}_{bcmk,t-1} = \arg\max_{\omega_{t-1}} \frac{1}{J} \sum_{j=1}^{J} \left\{ \frac{\left[\left(1-\omega_{t-1}\right)\exp\left(r_{f,t-1}\right) + \omega_{t-1}\exp\left(r_{f,t-1} + r_{bcmk,t}^{j}\right)\right]^{1-A}}{1-A} \right\}. \tag{30}$$

Similarly, using kd with $d \in \{baseline, tilted\}$ to denote either the baseline or the tilted density forecasts for predictor k, we can approximate (29) via

$$\widehat{\omega}_{kd,t-1} = \arg\max_{\omega_{t-1}} \frac{1}{J} \sum_{j=1}^{J} \left\{ \frac{\left[\left(1-\omega_{t-1}\right)\exp\left(r_{f,t-1}\right) + \omega_{t-1}\exp\left(r_{f,t-1} + r_{kd,t}^{j}\right)\right]^{1-A}}{1-A} \right\}. \tag{31}$$

The sequence of portfolio weights $\{\widehat{\omega}_{bcmk,t-1}\}_{t=m+1}^{T}$ and $\{\widehat{\omega}_{kd,t-1}\}_{t=m+1}^{T}$ are next used to compute the realized utilities under the HA-SV, baseline, and tilted densities. Let $\widehat{W}_{bcmk,t}$ and $\widehat{W}_{kd,t}$ be the corresponding realized wealth at time t, where $\widehat{W}_{bcmk,t}$ and $\widehat{W}_{kd,t}$ are functions of time t realized excess return, r_t, as well as the optimal allocations to the risky and risk-free assets computed in (30) and (31)

$$\begin{aligned} \widehat{W}_{bcmk,t} &= \left(1-\widehat{\omega}_{bcmk,t-1}\right)\exp\left(r_{f,t-1}\right) + \widehat{\omega}_{bcmk,t-1}\exp\left(r_{f,t-1} + r_t\right) \\ \widehat{W}_{kd,t} &= \left(1-\widehat{\omega}_{kd,t-1}\right)\exp\left(r_{f,t-1}\right) + \widehat{\omega}_{kd,t-1}\exp\left(r_{f,t-1} + r_t\right). \end{aligned} \tag{32}$$

Following Cenesizoglu and Timmermann (2012), we assess the performance of the predictive densities by calculating the implied annualized certainty equivalent return (CER) values for the OOS period as follows

$$
\begin{aligned}
CER_{bcmk} &= \left((1-A)(T-m)^{-1} \sum_{\tau=m+1}^{T} \widehat{U}_{bcmk,\tau} \right)^{12/(1-A)} - 1 \\
CER_{kd} &= \left((1-A)(T-m)^{-1} \sum_{\tau=m+1}^{T} \widehat{U}_{kd,\tau} \right)^{12/(1-A)} - 1,
\end{aligned}
\tag{33}
$$

where $\widehat{U}_{bcmk,\tau} = \widehat{W}^{1-A}_{bcmk,\tau}/(1-A)$ and $\widehat{U}_{kd,\tau} = \widehat{W}^{1-A}_{kd,\tau}/(1-A)$ denote the time-τ realized utility associated with the HA-SV and the baseline or tilted predictive density of predictor k, respectively. Finally, we compute the certainty equivalent return difference for predictor k using

$$
CERD_{kd} = CER_{kd} - CER_{bcmk} \tag{34}
$$

Table 2 reports the annualized CERD estimates associated with the baseline and tilted density forecasts for all predictors considered. For the remainder of our discussion here, we will refer to the former as baseline CERDs and we will refer to the latter as tilted CERDs. As in Table 1, we separately report results for the entire OOS period as well as for the two shorter periods, January 1995–December 2006 (OOS I), and January 2007–December 2014 (OOS II). We also examine the sensitivity of the CERDs to different risk preferences by considering risk-aversion coefficients of 3 (top panel) and 5 (bottom panel).[14]

Starting with $A = 3$ and the full OOS period, the tilted CERDs exceed the corresponding baseline CERDs for all 15 predictors considered. Across the 15 predictors, the average baseline CERDs is -0.69%, and the individual CERDs range between -3.439% for the default spread yield (DFY) and 3.256% for SII. On the other hand, the average tilted CERDs is 1.21%, ranging between -1.197% for DFY and 4.384% for SII. This corresponds to an average annual CER increase of 190 basis points (bps). For 1995–2006, the tilted CERDs exceed the baseline ones for all but two predictors, with an average improvement over the baseline CERDs of 118 bps. During this period, we find positive tilted CERDs for 10 of the 15 predictors, with the largest gains, 3.075%, associated with

[14] We compute the optimal portfolio weights for the CRRA investor using the approximation in Equation 2.4 of Campbell and Viceira (2002). Additionally, we restrict the portfolio weights to lie between -0.5 and 1.5 as in Rapach et al. (2016). We have also experimented with tighter bounds on the portfolio weights, ruling out short-selling and leverage (that is, $\omega_t \in [0, 1)$), as well as fully unconstrained portfolio weights. The results from these experiments are qualitatively very similar to the main results we report in Table 2.

the default return spread (DFR).

For the OOS II period, we find that once again the CERDs exceed the baseline CERDs for all 15 predictors. The average CER improvement from the baseline to the tilted densities is close to 300 bps, with the tilted density forecasts giving rise to CER gains as high as 8.579% in the case of SII. Similar results and improvements occur in the case of $A = 5$, where the average CER gains one would obtain by switching from the baseline to the tilted densities become 178 bps (1995–2014), 120 bps (OOS I period), and 262 bps (OOS II period).

The top two panels of Figure 5 plot the time series of equity weights for the monthly portfolios based on the EP and SII baseline and tilted densities, along with the equity weights implied by the HA-SV benchmark densities. While the HA-SV equity weights are relatively stable, ranging between 0.5 and 1, the baseline and tilted equity weights exhibit more variation. This is especially true during the late 1990s and right after the financial crisis. The tilted weights are generally larger than the baseline and benchmark ones, which means that the tilted densities tend to imply larger equity positions. This is particularly the case during the late 1990s and the mid 2000s when we tilt towards lower volatility. The bottom panels of the same figure show the corresponding log cumulative wealth for the three portfolios, computed using (32). By and large, the wealth generated by the tilted density forecasts lies above the baseline and benchmark counterparts, a pattern that is consistent throughout the whole OOS period, and in line with CER gains reported in Table 2.

4.3 Industry-Sorted Portfolios

We also investigate the performance of our entropic tilting method when applied to industry-sorted portfolios focusing on the following five industries: Consumers, Manufacturing, High-Tech, Health, and Other.[15] Although there is a vast literature examining the OOS predictability of U.S. aggregate returns, the analysis of OOS return predictability for industry portfolios is relatively rare; exceptions include Huang et al. (2015) and Pettenuzzo and Ravazzolo (2016). In terms of the predictors, we use market-wide measures of bond yields and inflation, and we construct industry-specific dividend

[15] For our analysis of industry portfolios, we use monthly CRSP data on the universe of NYSE firms, along with industry portfolio returns and industry classifications available at `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

price ratios, dividend yields, and realized volatilities.

We start by modifying the baseline specification in (8)–(9) as follows:

$$r^j_{\tau+1} = \mu^j + \beta^j x^j_\tau + \exp\left(h^j_{\tau+1}\right) u^j_{\tau+1}, \quad u^j_{\tau+1} \sim \mathcal{N}(0,1), \tag{35}$$

where $\tau = 1, ..., t-1$, $j = 1, ..., 5$, $r^j_{\tau+1}$ is time $\tau+1$ monthly excess return for industry j, x^j_τ is one of the industry predictors, and $h^j_{\tau+1}$ denotes the log-volatility of returns at time $\tau+1$. Similar to (9), we assume that the log-volatility $h^j_{\tau+1}$ follows a stationary and mean-reverting process and further depends on lagged intra-month information in the form of realized volatility estimates:

$$h^j_{\tau+1} = \lambda^j_0 + \lambda^j_1 h^j_\tau + \lambda^j_2 RV^j_\tau + \xi^j_{\tau+1}, \quad \xi^j_{\tau+1} \sim \mathcal{N}\left(0, \sigma^{2,j}_\xi\right), \tag{36}$$

where RV^j_τ denotes the realized volatility of returns for industry j at time τ, computed by summing the squared industry daily returns within month τ, $|\lambda^j_1| < 1$, and u^j_τ and ξ^j_s are mutually independent for all τ and s.[16]

To generate the volatility statistics that we tilt towards, we continue to use the VIX. This reveals a useful feature of our approach, namely that the options market from which the implied volatility is extracted need not match the asset that is being predicted. The projection in (20) allows us to map VIX to the industry volatility.[17]

The top panel of Table 3 shows the baseline and tilted ALSDs, relative to the corresponding industry's HA-SV benchmark, by industry, for each of the 10 predictors considered. For the Consumer, Health, and Others industries, the tilted ALSD is positive and significantly better than the baseline at the 1% level. For the Manufacturing and High-Tech industries, although the tilted ALSDs are positive, they fail to be significant at conventional levels in most instances. The baseline ALSDs have both positive and negative signs and generally fail to be significant at conventional levels for all combinations of industries and predictors considered.

The bottom panel of Table 3 shows that for the case of $A = 3$ the tilted densities deliver superior CER gains for the majority of combinations of industries and predictors. The tilted CER gains lie

[16] Daily industry returns going back to July 1926 are available from Kenneth French's Data Library.

[17] We thank an anonymous referee for pointing out to us this aspect of our approach.

between 0.339% for the Consumer industry using LTR as a predictor to 6.120% for the Manufacturing industry using DFR as a predictor. Tilting, however, seems to be quite detrimental for the High-Tech industry overall, and especially when using DP and DY as predictors. This result arises because volatility in the technology sector increased far above its typical value relative to the VIX during the dot-com frenzy. As a result, tilting the High-Tech distribution using the VIX works very poorly during this period. Overall, although our method performs well, when there is a mismatch between the options market used to obtain the implied volatilities and the asset being predicted, there are exceptions to our method's performance.

5 Conclusions

The paper introduces a novel approach to sharpen density forecasts for the equity premium using information from the derivatives markets in a time-series setting. We tilt predictive densities from a state-of-the-art stochastic volatility model in the empirical asset-pricing literature towards the second moment of the distribution implied by option prices, while imposing a non-negativity constraint on the mean of the resulting density. Tilting augments the *backward-looking* information in the baseline models with *forward-looking* information from the options in a straightforward manner that is not computationally intensive. Using monthly density forecasts for the S&P 500 and a number of industry-sorted portfolios between 1990 and 2014, we show that tilting significantly improves both the statistical and economic predictability of stock returns. Although improvements in forecasting the equity premium using information from the derivative markets have been previously documented, they have been limited to point forecasts, incorporating option-implied moments among predictors in forecasting regressions. Extending our method to higher moments, such as skewness and kurtosis, which have been receiving increased attention in empirical asset pricing, as well as longer investment horizons is a research agenda worth pursuing, especially as more options data become available.

References

Altavilla, C., R. Giacomini, and R. Constantini. 2014. "Bond Returns and Market Expectations." *Journal of Financial Econometrics*, 12(4): 708–729.

Altigan, Y., T.G. Bali, and K.O Demitras. 2015. "Implied Volatility Spreads and Expected Market Returns." *Journal of Business & Economic Statistics*, 33(1): 87–101.

Amisano, G., and R. Giacomini. 2007. "Comparing Density Forecasts via Weighted Likelihood Ratio Tests." *Journal of Business & Economic Statistics*, 25(2): 177–190.

Andersen, T.G., T. Bollerslev, P.F. Christoffersen, and F.X. Diebold. 2006. "Volatility and Correlation Forecasting." In *Handbook of Economic Forecasting, volume 1.* eds. by G. Elliott, C.W.J. Granger, and A. Timmermann: North Holland, 777–878.

Back, K. 2010. *Asset Pricing and Portfolio Choice Theory.*: Oxford University Press.

Barberis, N. 2000. "Investing for the Long Run when Returns are Predictable." *Journal of Finance*, LV 225–264.

Bloom, N. 2009. "The Impact of Uncertainty Shocks." *Econometrica*, 77(3): 623–685.

Bollerslev, T., G. Tauchen, and H. Zhou. 2009. "Expected Stock Returns and Variance Risk Premia." *Review of Financial Studies*, 22(11): 4463–4492.

Campbell, J.Y., and S. Thompson. 2008. "Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?" *Review of Financial Studies*, 21(4): 1509–1531.

Campbell, J.Y., and L.M. Viceira. 2002. *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors.*: Oxford University Press.

Cenesizoglu, T., and A. Timmermann. 2012. "Do return prediction models add economic value?" *Journal of Banking and Finance*, 36(11): 2974–2987.

Clark, T.E. 2011. "Real-Time Density Forecasts From Bayesian Vector Autoregressions With Stochastic Volatility." *Journal of Business & Economic Statistics*, 29(3): 327–341.

Clark, T.E., and M. McCracken. 2011. "Nested Forecast Model Comparisons: a New Approach to Testing Equal Accuracy." *Federal Reserve Bank of St. Louis Working Paper.*

Clark, T.E., and F. Ravazzolo. 2015. "Macroeconomic Forecasting Performance Under Alternative Specifications of Time-Varying Volatility." *Journal of Applied Econometrics*, 30(1): 551–575.

Cogley, T., S. Morozov, and T.J. Sargent. 2005. "Bayesian Fan Charts for U.K. inflation: Forecasting and sources of uncertainty in an evolving monetary system." *Journal of Economic Dynamics & Control*, 29(11): 1893–1925.

Diebold, F.X., and R.S. Mariano. 1995. "Comparing Predictive Accuracy." *Journal of Business & Economic Statistics*, 13(3): 253–263.

Engle, R.F., and G. Gallo. 2006. "A multiple indicators model for volatility using intra-daily data." *Journal of Econometrics*, 131(1): 3–27.

Faust, J., and J.H. Wright. 2013. "Forecasting Inflation." In *Handbook of Economic Forecasting, volume 2.* eds. by G. Elliott, and A. Timmermann: North Holland, 2–56.

Giacomini, R., and G. Ragusa. 2014. "Theory-coherent forecasting." *Journal of Econometrics*, 182(1): 145–155.

Giacomini, R., and B. Rossi. 2010. "Forecast comparisons in unstable environments." *Journal of Applied Econometrics*, 25(4): 595–620.

Goyal, A., and I. Welch. 2008. "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction." *Review of Financial Studies*, 21(4): 1455–1508.

Hansen, P.R., Z. Huang, and H.H. Shek. 2012. “Realized GARCH: A Joint Model for Returns and Realized Measures of Volatility.” *Journal of Applied Econometrics*, 27(6): 877–906.

Harvey, D., S. Leybourne, and P. Newbold. 1997. “Testing the equality of prediction mean squared errors.” *International Journal of Forecasting*, 13(2): 281–291.

Huang, D., F. Jiang, J. Tu, and G. Zhou. 2015. “Investor Sentiment Aligned: A Powerful Predictor of Stock Returns.” *Review of Financial Studies*, 28(3): 791–837.

Johannes, M.S., A. Korteweg, and N. Polson. 2014. “Sequential Learning, Predictive Regressions, and Optimal Portfolio Returns.” *Journal of Finance*, 69(2): 611–644.

Jorion, P. 1995. “Predicting Volatility in the Foreign Exchange Market.” *Journal of Finance*, 50(2): 507–528.

Kandel, S., and R. Stambaugh. 1996. “On the Predictability of Stock Returns: An Asset-Allocation Perspective.” *Journal of Finance*, 51(2): 385–424.

Koop, G. 2003. *Bayesian Econometrics.*: John Wiley & Sons, Ltd.

Krüger, F., T. Clark, and F. Ravazzolo. 2015. “Using Entropic Tilting to Combine BVAR Forecasts with External Newcasts.” *Journal of Business & Economic Statistics*, Fortchoming.

Pettenuzzo, D., and F. Ravazzolo. 2016. “Optimal Portfolio Choice Under Decision-Based Model Combinations.” *Journal of Applied Econometrics*, p. forthcoming.

Pettenuzzo, D., A. Timmermann, and R. Valkanov. 2014. “Forecasting Stock Returns under Economic Constraints.” *Journal of Financial Economics*, 114(3): 517–553.

Poon, S-H., and C.W.J Granger. 2003. “Forecasting Volatility in Financial Markets: A Review.” *Journal of Economic Literature*, 41(2): 478–539.

Primiceri, G.E. 2005. “Time Varying Structural Vector Autoregressions and Monetary Policy.” *The Review of Economic Studies*, 72(3): 821–852.

Rapach, D.E., M.C. Ringgenberg, and G. Zhou. 2016. "Short interest and aggregate stock returns." *Journal of Financial Economics*, p. Forthcoming.

Rapach, D.E., and G. Zhou. 2013. "Forecasting stock returns." In *Handbook of Economic Forecasting, volume 2*. eds. by G. Elliot, and A. Timmermann: North Holland, 329–383.

Robertson, J.C., E.W. Tallman, and C.H. Whiteman. 2002. "Forecasting Using Relative Entropy." *Federal Reserve Bank of Atlanta WP 2002-22*.

Robertson, J.C., E.W. Tallman, and C.H. Whiteman. 2005. "Forecasting Using Relative Entropy." *Journal of Monery, Credit and Banking*, 37(3): 383–401.

Shephard, N., and K. Sheppard. 2010. "Realising the Future: Forecasting with High-Frequency-Based Volatility (HEAVY) Models." *Journal of Applied Econometrics*, 25(2): 197–231.

Szakmary, A., E. Ors, J.K Kim, and W.N. Davidson. 2003. "The predictive power of implied volatility: Evidence from 35 futures markets." *Journal of Banking and Finance*, 27(3): 2151–2175.

Young Chang, B., P. Christoffersen, and K. Jacobs. 2013. "Market skewness risk and the cross section of stock returns." *Journal of Financial Economics*, 107(1): 46–68.

Table 1: Out-of-sample statistical predictability

Panel A: Out-of-sample R^2 (vs. HA-SV)						
	1995:01-2014:12		*1995:01-2006:12*		*2007:01-2014:12*	
	(1)	(2)	(3)	(4)	(5)	(6)
Predictor	*Baseline*	*Tilted*	*Baseline*	*Tilted*	*Baseline*	*Tilted*
DP	-1.287	**-1.103****	-2.310	**-2.002****	-0.045	**-0.011**
DY	-1.122	**-1.002****	-2.171	**-1.946****	0.152	0.144
EP	-0.481	**-0.005**	-0.345	-0.352	-0.647	**0.416**
DE	-1.414	**-0.562**	-1.137	**-1.060****	-1.751	**0.042**
RVOL	-0.368	-0.419	-1.614	-1.671	1.147	1.103
BM	-0.443	**-0.418**	-0.737	**-0.687**	-0.087	-0.092
NTIS	-2.186	**-2.153**	-1.158	**-1.131**	-3.436	**-3.395**
TBL	-0.432	-0.492	-0.685	-0.759	-0.125	-0.168
LTY	-0.213	**-0.193**	-0.551	-0.582	0.198	**0.280****
LTR	-0.415	-0.576	-0.852	**-0.596***	0.116	-0.552
TMS	-0.902	-0.958	-1.424	-1.455	-0.267	-0.354
DFY	-2.312	**-1.911****	-2.663	**-1.897****	-1.885	-1.928
DFR	-1.066	-1.743	-2.121	**-1.074**	0.217	-2.556
INFL	-0.852	**-0.784**	-0.214	**-0.071**	-1.627	-1.651
SII	1.902*	1.725	0.013	-0.036	4.198*	3.864
Panel B: Average log score differences (vs. HA-SV)						
	1995:01-2014:12		*1995:01-2006:12*		*2007:01-2014:12*	
	(1)	(2)	(3)	(4)	(5)	(6)
Predictor	*Baseline*	*Tilted*	*Baseline*	*Tilted*	*Baseline*	*Tilted*
DP	-0.009	**0.130*****	-0.013	**0.118*****	-0.003	**0.148*****
DY	-0.009	**0.130*****	-0.014	**0.116*****	-0.001	**0.151*****
EP	-0.004	**0.141*****	-0.004	**0.136*****	-0.005	**0.150*****
DE	-0.008	**0.132*****	-0.009	**0.123*****	-0.007	**0.146*****
RVOL	-0.005	**0.131*****	-0.010	**0.117*****	0.003	**0.152*****
BM	-0.003	**0.139*****	-0.005	**0.131*****	-0.001	**0.151*****
NTIS	-0.008	**0.128*****	-0.006	**0.133*****	-0.012	**0.119*****
TBL	-0.001	**0.141*****	-0.003	**0.136*****	0.002	**0.148*****
LTY	-0.004	**0.139*****	-0.006	**0.133*****	-0.002	**0.149*****
LTR	-0.002	**0.145*****	-0.003	**0.146*****	0.000	**0.145*****
TMS	-0.006	**0.133*****	-0.009	**0.121*****	-0.001	**0.151*****
DFY	-0.012	**0.131*****	-0.014	**0.121*****	-0.008	**0.146*****
DFR	-0.006	**0.133*****	-0.009	**0.128*****	-0.002	**0.139*****
INFL	-0.002	**0.149*****	-0.002	**0.143*****	-0.002	**0.158*****
SII	0.004	**0.151*****	-0.003	**0.141*****	0.014	**0.165*****

Note: The table reports the out-of-sample (OOS) R^2 in (23) and the average log score differences (ALSDs) in (24) for the 15 predictors considered, over the entire OOS period, 1995:01–2014:12, as well as for 1995:01–2006:12 and 2007:01–2012:14. All forecasts are OOS using recursive estimates for 1995:01–2014:12. Bold numbers indicate all instances where the tilted forecasts improve upon the corresponding baseline forecasts. The asterisks indicate statistical significance at 10%(*), 5%(**), and 1%(***) levels, using the Diebold and Mariano (1995) tests discussed in Section 4.1. The predictor nomenclature is as follows: (1) DP: Log dividend price-ratio; (2) DY: Log dividend yield; (3) EP: Log earning-price ratio; (4) DE: Log dividend-payout ratio; (5) RVOL: Excess stock return volatility; (6) BM: Book-to-market ratio; (7) NTIS: Net equity expansion; (8) TBL: Treasury bill rate; (9) LTY: Long-term yield; (10) LTR: Long-term return; (11) TMS: Term spread; (12) DFY: Default yield spread; (13) DFR: Default return spread; (14) INFL: Inflation; (15) SII: Short interest index. We use HA-SV to refer to the historical-average model augmented with stochastic volatility.

Table 2: Out-of-sample economic predictability

	Panel A: CER gains, $A = 3$ (vs. HA-SV)					
	1995:01-2014:12		*1995:01-2006:12*		*2007:01-2014:12*	
	(1)	(2)	(3)	(4)	(5)	(6)
Predictor	*Baseline*	*Tilted*	*Baseline*	*Tilted*	*Baseline*	*Tilted*
DP	-2.226	**-0.529**	-3.571	**-2.888**	-0.257	**2.975**
DY	-2.041	**-0.540**	-3.420	**-2.815**	-0.021	**2.834**
EP	-0.472	**2.585**	-0.622	**1.836**	-0.255	**3.673**
DE	-1.262	**1.070**	-1.739	**-0.041**	-0.571	**2.691**
RVOL	-0.691	**0.277**	-2.119	-2.834	1.402	**4.945**
BM	-0.531	**1.312**	-0.793	**0.368**	-0.153	**2.687**
NTIS	-1.735	**1.281**	-0.609	**0.496**	-3.328	**2.422**
TBL	0.490	**1.607**	0.200	**1.053**	0.908	**2.410**
LTY	0.200	**1.601**	-0.467	**0.728**	1.167	**2.873**
LTR	-0.707	**0.276**	-0.710	**0.355**	-0.702	**0.163**
TMS	-0.928	**1.569**	-1.369	**0.997**	-0.291	**2.396**
DFY	-3.439	**-1.197**	-3.902	**-3.490**	-2.767	**2.204**
DFR	0.294	**3.358**	-0.390	**3.075**	1.287	**3.765**
INFL	-0.544	**1.101**	0.255	0.026	-1.680	**2.670**
SII	3.256	**4.384**	0.010	**1.576**	8.135	**8.579**
	Panel B: CER gains, $A = 5$ (vs. HA-SV)					
	1995:01-2014:12		*1995:01-2006:12*		*2007:01-2014:12*	
	(1)	(2)	(3)	(4)	(5)	(6)
Predictor	*Baseline*	*Tilted*	*Baseline*	*Tilted*	*Baseline*	*Tilted*
DP	-1.324	**-0.297**	-2.151	-2.365	-0.124	**2.766**
DY	-1.205	**-0.247**	-2.028	-2.227	-0.010	**2.682**
EP	-0.280	**2.456**	-0.393	**1.778**	-0.116	**3.437**
DE	-0.879	**0.593**	-1.063	**-0.598**	-0.613	**2.333**
RVOL	-0.403	**0.084**	-1.285	-2.319	0.878	**3.662**
BM	-0.296	**1.840**	-0.481	**1.510**	-0.030	**2.316**
NTIS	-1.339	**0.591**	-0.340	**1.235**	-2.746	**-0.324**
TBL	0.324	**1.962**	0.092	**1.661**	0.657	**2.395**
LTY	0.138	**2.227**	-0.281	**1.739**	0.744	**2.931**
LTR	-0.396	**0.787**	-0.401	**1.065**	-0.388	**0.389**
TMS	-0.564	**1.066**	-0.833	**0.261**	-0.176	**2.236**
DFY	-2.036	**-0.729**	-2.324	**-2.008**	-1.622	**1.142**
DFR	-0.641	**2.986**	-1.284	**2.052**	0.290	**4.345**
INFL	-0.311	**2.298**	0.153	**1.775**	-0.971	**3.054**
SII	2.476	**4.323**	-0.003	**1.874**	6.173	**7.973**

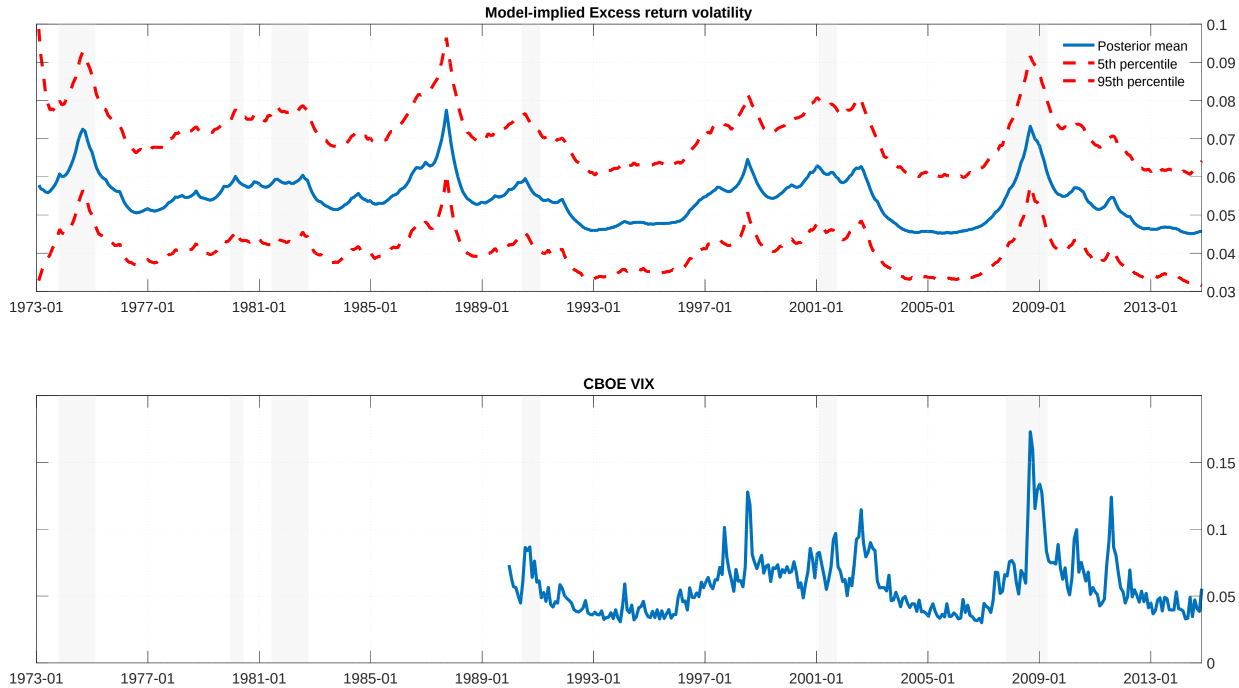
Note: The table reports the annualized certainty equivalent return differences (CERDs) in (34) for portfolio decisions based on recursive out-of-sample (OOS) forecasts of excess returns. Each period, an investor with power utility and coefficient of relative risk aversion $A = 3$ (top panel) or $A = 5$ (bottom panel) selects stocks and T-bills based on a predictive density differing both by the predictor considered and the model entertained (baseline or tilted). See the notes of Table 1 for the predictor nomenclature. The equity weights are constrained to lie in the $[-0.5, 1.5]$ interval. All forecasts are OOS using recursive estimates of the models for 1995:01–2014:12. Bold numbers indicate all instances where CER gains for the tilted densities exceed the CER gains for the baseline densities. We use HA-SV to refer to the historical-average model augmented with stochastic volatility.

Table 3: Out-of-sample predictability: industry-sorted portfolios

	Panel A: Average log score differences (vs. HA-SV), 1995:01–2014:12									
	Consumer		*Manufacturing*		*High-Tech*		*Health*		*Others*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Predictor	*Baseline*	*Tilted*	*Baseline*	*Tilted*	*Baseline*	*Tilted*	*Baseline*	*Tilted*	*Baseline*	*Tilted*
DP	-0.003	**0.149*** **	-0.017	**0.055* **	-0.006	**0.006**	-0.012	**0.103*** **	-0.006	**0.127*** **
DY	-0.001	**0.150*** **	-0.015	**0.058* **	-0.006	**0.007**	-0.011	**0.099*** **	-0.009	**0.125*** **
RVOL	0.000	**0.149*** **	-0.003	**0.070* **	-0.002	**0.027**	-0.002	**0.103*** **	-0.004	**0.128*** **
LTR	-0.002	**0.156*** **	-0.004	**0.071** **	0.000	**0.022**	0.000	**0.107*** **	-0.004	**0.142*** **
TBL	0.000	**0.153*** **	0.000	**0.054**	0.001	**0.020**	-0.003	**0.097*** **	-0.002	**0.128*** **
LTY	-0.004	**0.149*** **	0.000	**0.066**	0.001	**0.020**	-0.003	**0.101*** **	-0.004	**0.131*** **
TMS	-0.002	**0.155*** **	-0.004	**0.049**	-0.005	**0.021**	-0.002	**0.098*** **	-0.006	**0.126*** **
DFY	-0.012	**0.142*** **	-0.006	**0.061* **	-0.008	**0.026**	-0.006	**0.100*** **	-0.010	**0.117*** **
DFR	0.000	**0.133*** **	0.002	**0.065**	-0.001	**0.030**	-0.006	**0.099*** **	0.000	**0.124*** **
INFL	-0.002	**0.157*** **	0.000	**0.080** **	0.002	**0.024**	-0.002	**0.098*** **	-0.002	**0.134*** **
	Panel B: CER gains, $A = 3$ (vs. HA-SV), 1995:01–2012:14									
	Consumer		*Manufacturing*		*High-Tech*		*Health*		*Others*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Predictor	*Baseline*	*Tilted*	*Baseline*	*Tilted*	*Baseline*	*Tilted*	*Baseline*	*Tilted*	*Baseline*	*Tilted*
DP	-0.773	**2.139**	-4.272	**-0.148**	-3.787	-7.483	-3.755	**0.987**	-2.233	**-0.712**
DY	-0.785	**1.687**	-3.885	**-0.130**	-3.781	-7.570	-3.707	**1.068**	-2.463	**-0.699**
RVOL	-0.273	**2.699**	-0.381	**2.335**	0.684	**2.670**	-0.774	**1.005**	-0.828	**2.506**
LTR	-0.465	**0.339**	-1.180	**0.740**	0.170	-1.832	0.393	**1.314**	-1.269	**0.427**
TBL	0.641	**2.733**	0.861	**2.359**	1.085	-2.147	-0.034	**1.526**	-0.338	**2.234**
LTY	-0.704	**2.213**	0.978	**2.306**	0.503	-1.965	-0.468	**0.957**	-0.243	**2.070**
TMS	-0.329	**2.413**	-0.892	**2.170**	-0.984	-2.177	-0.288	**1.573**	-1.047	**2.074**
DFY	-4.341	**0.493**	-2.040	**0.906**	-2.878	**-2.770**	-1.988	**-0.320**	-3.142	**-1.376**
DFR	1.227	**3.947**	2.366	**6.120**	-0.224	-1.755	-0.924	**1.105**	0.901	**3.749**
INFL	-0.401	**2.063**	0.480	**2.859**	1.061	-0.159	0.097	**2.078**	-0.936	**1.629**

Note: The table reports the average log score differences (ALSDs) in (24) and the annualized certainty equivalent return differences (CERDs) in (34) for 5 industry-sorted portfolios following the industry definitions at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/datalibrary.html. All forecasts are out-of-sample using recursive estimates for 1995:01–2014:12. See the notes of Table 1 for the predictor nomenclature. Bold numbers indicate all instances where the ALSDs and the CER gains for the tilted densities exceed the ALSDs and CER gains for the baseline densities. For the ALSDs, the asterisks indicate statistical significance at 10%(*), 5%(**), and 1%(***) levels, using the Diebold and Mariano (1995) tests discussed in Section 4.1. We use HA-SV to refer to the historical-average model augmented with stochastic volatility.

Figure 1: Model-implied excess return volatility and VIX

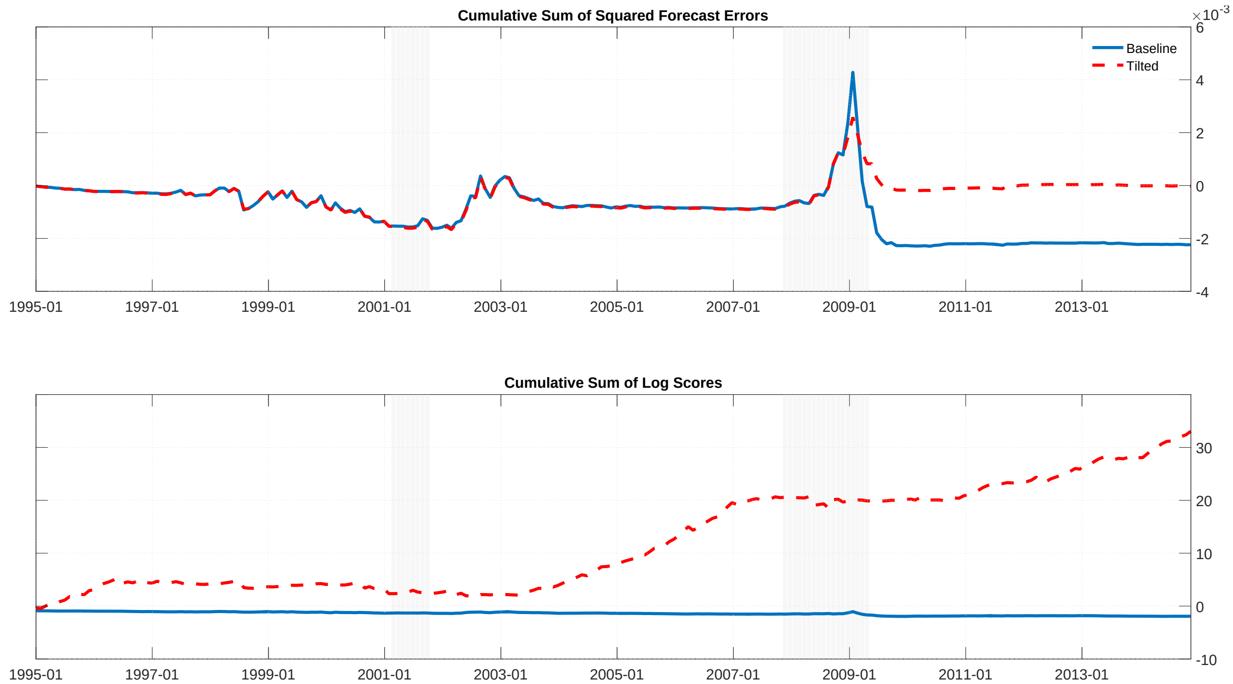


Note: The top panel of the figure shows the monthly excess return volatility implied by the stochastic volatility model in (8)-(9) for 1973:01–2014:12 using the earnings-price ratio (EP) as predictor. The solid line shows the posterior mean of $\exp(h_t)$, $t = 1, ..., T$, while the dashed lines show the 5th and 95th percentiles of its posterior distribution. The bottom panel of the figure shows the end-of-month values of the Chicago Board Options Exchange (CBOE) Volatility Index (VIX) for 1990:01–2014:12. The grey shading identifies NBER recessions.

Figure 2: Moments of the baseline and tilted predictive densities

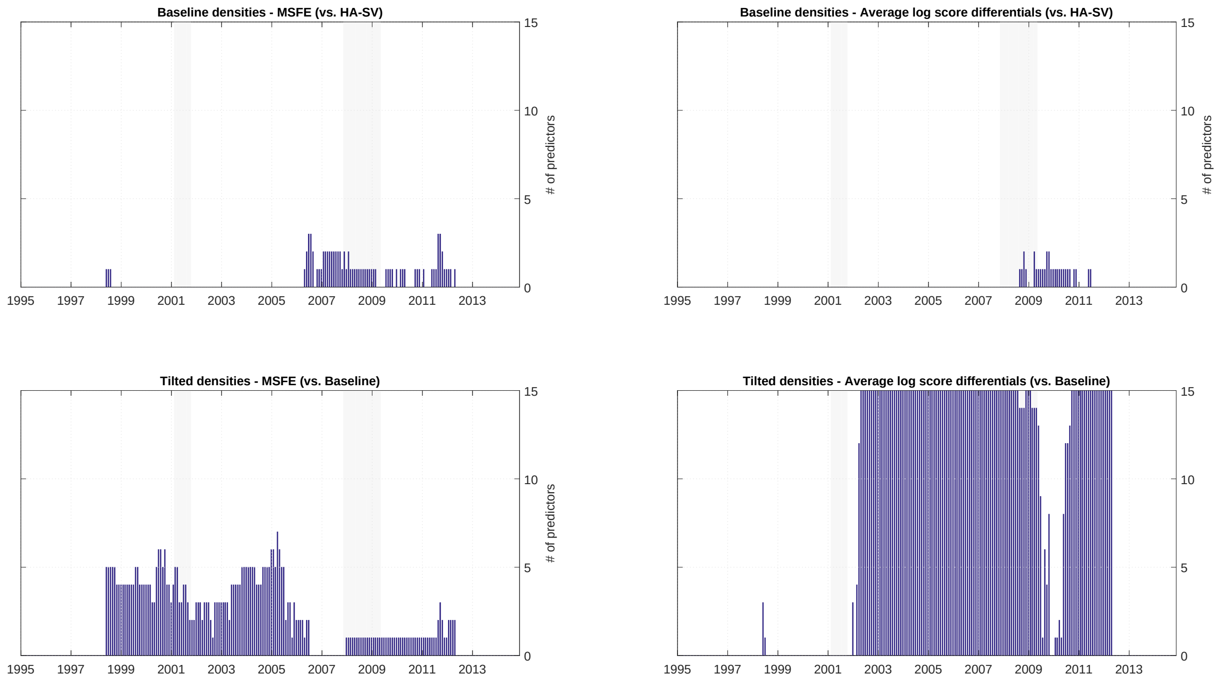
Note: The figure shows the first two moments of excess returns for the baseline and tilted predictive densities of the stochastic volatility model in (8)-(9) using the earning price ratio (EP) as predictor. Both predictive densities are out-of-sample using recursive estimates for 1995:01–2014:12. The top panel shows the posterior mean of the baseline and tilted predictive densities. The bottom panel shows the posterior volatility of the baseline and tilted predictive densities. The grey shading identifies NBER recessions.

Figure 3: Cumulative sums of squared forecast error and log score differences



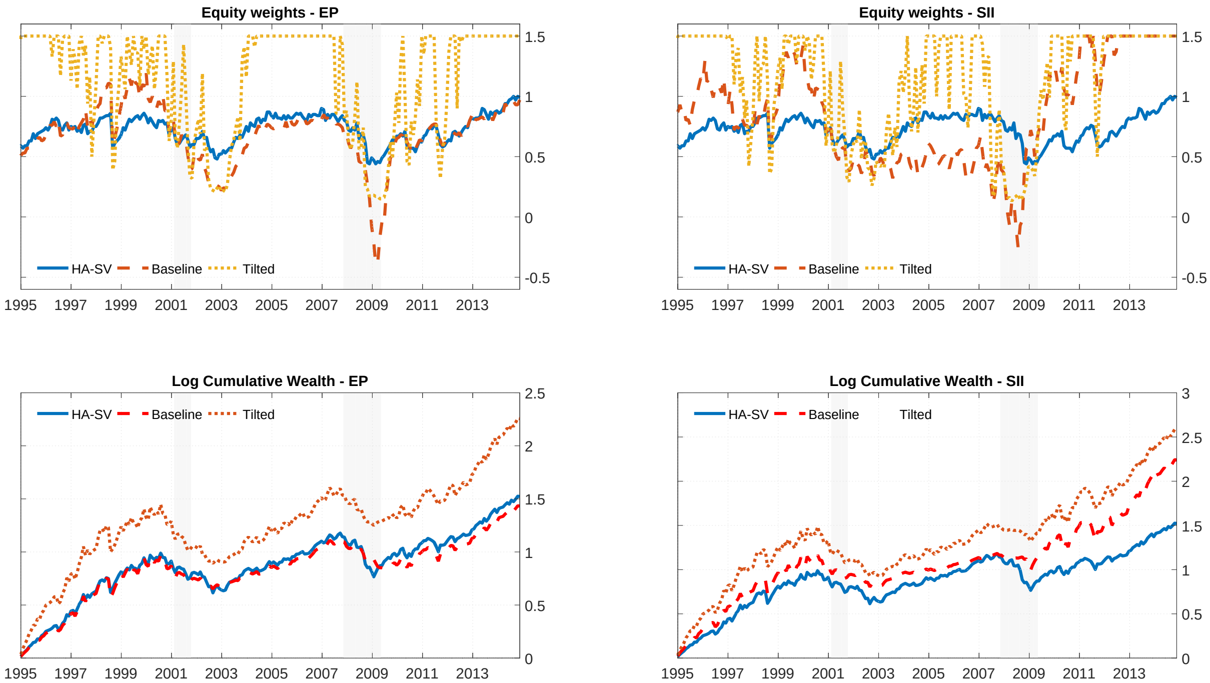
Note: The top panel of the figure shows the cumulative sum of squared error differences (CSSEDs) in (25) using the earnings-price ratio (EP) as a predictor. CSSEDs above zero indicate that the baseline and/or tilted densities generate better predictions than the historical-average with stochastic volatility (HA-SV) benchmark densities. Negative CSSED values suggest the opposite. The bottom panel shows the cumulative log score differences (CLSDs) in (26) using the same predictor. CLSDs above zero indicate that the baseline and/or tilted models generate better performance than the HA-SV benchmark. Negative values suggest the opposite. All forecasts are out-of-sample using recursive estimates for 1995:01–2014:12. The gray shading identifies NBER recessions.

Figure 4: Fluctuation test across predictors



Note: The figure shows the number of rejections in the one-sided Giacomini and Rossi (2010) fluctuation test for the baseline and tilted predictive densities of all 15 predictors considered over centered rolling windows of 60 observations as discussed in Section 4.1. The critical values have been adjusted to reflect our recursive forecasts. The two left panels focus on the mean squared forecast error differences (MSFEDs), while the two right panels focus on the average log score differences (ALSDs) in (24). All forecasts are out-of-sample using recursive estimates for 1995:01–2014:12. The grey shading identifies NBER recessions.

Figure 5: Asset allocation weights and cumulative wealth



The top two panels of the figure show the time series of equity weights in (31) of the monthly portfolios for the earnings-price (EP) and the short-interest index (SII) baseline and tilted densities, along with the equity weights from the historical-average with stochastic volatility (HA-SV) benchmark. We compute the optimal allocation to stocks and T-bills based on the predictive density of excess returns. The investor is assumed to have power utility with a coefficient of relative risk aversion A=3 in (28), while the equity weights are constrained to lie in the $[-0.5, 1.5]$ interval. The bottom two panels of the figure show the corresponding log cumulative wealth. All forecasts are out-of-sample using recursive estimates for 1995:01—2014:12.