Does realized volatility help bond yield density prediction?

Minchul Shin
University of Illinois

Molin Zhong*
Federal Reserve Board

This version: August 24, 2016

Abstract

We suggest using "realized volatility" as a volatility proxy to aid in model-based multivariate bond yield density forecasting. To do so, we develop a general estimation approach to incorporate volatility proxy information into dynamic factor models with stochastic volatility. The resulting model parameter estimates are highly efficient, which one hopes would translate into superior predictive performance. We explore this conjecture in the context of density prediction of U.S. bond yields by incorporating realized volatility into a dynamic Nelson-Siegel (DNS) model with stochastic volatility. The results clearly indicate that using realized volatility improves density forecasts relative to popular specifications in the DNS literature that neglect realized volatility.

Key words: Dynamic factor model, forecasting, stochastic volatility, term structure of interest rates, dynamic Nelson-Siegel model

JEL codes: C5, G1, E4

*Correspondence: Minchul Shin: 214 David Kinley Hall, 1407 W. Gregory Dr., Urbana, Illinois 61801. E-mail: mincshin@illinois.edu. Molin Zhong: 20th Street and Constitution Avenue N.W., Washington, D.C. 20551. E-mail: Molin.Zhong@frb.gov. We are grateful for the advice of Frank Diebold, Jesus Fernandez-Villaverde, and Frank Schorfheide. We also thank Manabu Asai, Luigi Bocola, Todd Clark, Xu Cheng, Frank DiTraglia, Nikolaus Hautsch, Kyu Ho Kang, Michael McCracken, Andrew Patton, Neil Shephard, Dongho Song, Allan Timmermann, Jonathan Wright, and seminar participants at the University of Pennsylvania, International Symposium on Forecasting 2013, and OMI-SoFiE Financial Econometrics Summer School 2013 for their comments. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of any other person associated with the Federal Reserve System.

1 Introduction

Time-varying volatility exists in U.S. government bond yields[1]. In this paper, we introduce volatility proxy data in the hopes of better capturing this time-varying volatility for predictive purposes. To do so, we develop a general estimation approach to incorporate volatility proxy information into dynamic factor models with stochastic volatility. We apply it to the dynamic Nelson-Siegel (DNS) model of bond yields. We find that the higher frequency movements of the yields in the realized volatility data contain valuable information for the stochastic volatility and lead to significantly better density predictions, especially in the short term.

Our approach can be applied to the existing classes of dynamic factor models with stochastic volatility. Specifically, we can account for stochastic volatility on the latent factors or stochastic volatility on the measurement errors. We derive a measurement equation to link realized volatility to the model-implied conditional volatility of the original observables. Incorporating realized volatility improves estimation of the stochastic volatility by injecting precise volatility information into the model.

The DNS model is a dynamic factor model that uses latent level, slope, and curvature factors to drive the intertemporal movements of the yield curve. This reduces the high-dimensional yields to be driven by just three factors. The level of the yield curve has traditionally been linked to inflation expectations while the slope to the real economy. Our preferred specification introduces stochastic volatility on these latent factors. This leads to a nice interpretation of the stochastic volatility as capturing the uncertainty surrounding well-understood aspects of the yield curve. It also reduces the dimension of modeling the time-varying volatility of the yield curve.

[1]Early empirical work was done by Engle et al. (1990). In the years that followed, there have been many attempts to empirically validate and analyze time-varying volatility in US government bond yields/prices (See, for example, papers cited in Diebold and Rudebusch, 2012).

We then compare this specification to several others in the DNS framework, including random walk dynamics for the factors and stochastic volatilities, as well as stochastic volatility on the yield measurement equation. In a forecasting horserace on U.S. bond yields, our preferred specification features slight improvements in the point forecast performance and significant gains in the density forecast performance. The realized volatility data injects accurate conditional second moment information into the model, which aids in both the extraction of the current volatility state and also the estimation of the volatility process parameters. Both are important for density prediction, with the accurate volatility state estimation effect dominating for short horizon forecasts and the accurate parameter estimation effect dominating at longer horizons. We also find that allowing for time-varying volatility is important for density prediction, especially in the short run. Unlike conditional mean dynamics, modeling volatility as first-order autoregressive processes rather than random walks leads to better predictive performance. Furthermore, having stochastic volatility on the factor equation better captures the time-varying volatility in the bond yield data when compared to stochastic volatility on the measurement equation.

Our paper relates to the literature in three main areas. First, our paper relates to work started by Barndorff-Nielsen and Shephard (2002) in incorporating realized volatility in models with time-varying volatility. Takahashi et al. (2009) use daily stock return data in combination with high-frequency realized volatility to more accurately estimate the stochastic volatility. Maheu and McCurdy (2011) show that adding realized volatility directly into a model of stock returns can improve density forecasts over a model that only uses level data, such as the EGARCH. Jin and Maheu (2013) propose a model of stock returns and realized covariance based on time-varying Wishart distributions and find that their model provides superior density forecasts for returns. There also exists work adding realized volatility in observation-driven volatility models (Shephard and Sheppard, 2010; Hansen et al., 2012). As opposed to the other papers, we consider a dynamic factor model with stochastic volatility on the factor equation and use the realized volatility to help in the extraction of this

stochastic volatility. In this sense, we bring the factor structure in the conditional mean to the conditional volatility as well. Cieslak and Povala (2015) have a similar framework in a no-arbitrage term structure model. Furthermore, we are the first paper to investigate the implications of realized volatility for bond yield density predictability.

Second, we contribute to a large literature on bond yield forecasting. Most of the work has been done on point prediction (see for example, Diebold and Rudebusch, 2012; Duffee, 2012, for excellent surveys). There has been, however, a growing interest in density forecasting. Egorov et al. (2006) were the first to evaluate the joint density prediction performance of yield curve models. They overturn the point forecasting result of the superiority in random walk forecasts and find that affine term structure models perform better when forecasting the entire density, especially on the conditional variance and kurtosis. However, they do not consider time-varying conditional volatility dynamics in the bond yield predictive distribution. Hautsch and Ou (2012) and Hautsch and Yang (2012) add stochastic volatility to the DNS model by considering an independent AR(1) specification for the log volatilities of the latent factors. They do not do formal density prediction evaluation of the model, but give suggestive results of the possible improvements in allowing for time-varying volatility. Carriero et al. (2013) find that using priors from a Gaussian no-arbitrage model in the context of a VAR with stochastic volatility improves short-run density forecasting performance. Building on this previous work, we introduce potentially highly accurate volatility information into the model in the form of realized volatility and evaluate bond yield density predictions to see whether this extra information about the bond yield volatility can improve the quality of the predictive distribution.

Another related class of bond yield prediction papers in the literature uses external information to improve the quality of prediction. Altavilla et al. (2013) exploit information contained in survey expectations data and use it to restrict model-implied forecasts via a flexible informational projection method. van Dijk et al. (2014) use various sources of external information, including survey expectations, to capture a shift in the endpoint of the

yield curve. These papers attempt to improve the point prediction by incorporating external information. On the other hand, our paper exploits external information to improve the density prediction by accurately estimating the latent volatility states and their related parameters.

Finally, we also add to a growing literature on including realized volatility information in bond yield models. Andersen and Benzoni (2010) and Christensen et al. (2014) view realized volatility as a benchmark on which to compare the fits of affine term structure models. Cieslak and Povala (2015) are interested in using realized covariance to better extract stochastic volatility and linking the stochastic volatility to macroeconomic and liquidity factors. These papers focus on in-sample investigations of incorporating realized volatility in bond yield models. Another stream of research exploits information in high-frequency movements of bond prices to achieve better point prediction performance. For example, Wright and Zhou (2009) report that the realized jump mean measure constructed from Treasury bond futures improves excess bond return point prediction by 40%. Our paper, in contrast to these others, considers the improvement from using realized volatility in out-of-sample bond yield density prediction.

In section 2, we introduce our methodology for incorporating volatility proxies into dynamic factor models in the context of the DNS model and other competitor specifications. We discuss the data in section 3. We present our estimation and forecast evaluation methodology in section 4. In section 5, we present in-sample and out-of-sample results. We conclude in section 6.

2 Model

We introduce the dynamic Nelson-Siegel model with stochastic volatility (DNS-SV) proposed by Bianchi et al. (2009), Hautsch and Ou (2012), and Hautsch and Yang (2012). Then, we

discuss the incorporation of realized volatility information into this framework. Finally, we consider alternatives to our main approach.

2.1 The Dynamic Nelson-Siegel model and time-varying bond yield volatility

Denote $y_t(\tau)$ as the continuously compounded yield to maturity on a zero coupon bond with maturity of τ periods at time t. Following Diebold and Li (2006), we consider a factor model for the yield curve,

$$y_t(\tau) = f_{l,t} + f_{s,t}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right) + f_{c,t}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right) + \epsilon_t(\tau), \quad \epsilon_t \sim N(0,\ Q) \qquad (1)$$

where $f_{l,t}, f_{s,t}$ and $f_{c,t}$ serve as latent factors and ϵ_t is a vector that collects the idiosyncratic component $\epsilon_t(\tau)$ for all maturities. As is well documented in the literature, the first factor mimics the level of the yield curve, the second the slope, and the third the curvature. We assume that the Q matrix is diagonal. This leads to the natural interpretation of a few common factors driving the comovements in a large number of yields. All of the other movements in the yields are considered idiosyncratic. We model the dynamic factors as independent autoregressive processes, given by,

$$f_{i,t} = (1-\Phi_{i,f})\mu_{i,f} + \Phi_{i,f} f_{i,t-1} + \eta_{i,t}, \quad \eta_{i,t} \sim N(0,\ e^{h_{i,t}}) \qquad (2)$$

for $i = l, s, c$, where $\eta_{i,t}$ is the corresponding innovation to each factor, with potentially time-varying volatility $h_{i,t}$. We also assume that idiosyncratic shocks ϵ_t and factor shocks η_t are independent. Following Bianchi et al. (2009), Hautsch and Ou (2012), and Hautsch and Yang (2012), we model the logarithm of the variance of the shocks to the factor equation as

AR(1) processes,

$$h_{i,t} = \mu_{i,h}(1 - \phi_{i,h}) + \phi_{i,h} h_{i,t-1} + \nu_{i,t}, \quad \nu_{i,t} \sim N(0,\ \sigma^2_{i,h}) \tag{3}$$

for $i = l, s, c$, where $e^{h_{i,t}}$ corresponds to the volatility of the innovation to factor $f_{i,t}$[2]. In addition, shocks to the stochastic volatilities of the factor innovations are assumed to be independent. We call this specification the DNS-SV model (dynamic Nelson-Siegel with stochastic volatility).

2.2 DNS-RV

We claim that by using high-frequency data to construct realized volatilities of the yields, it is possible to aid in the extraction of the stochastic volatilities governing the level, slope, and curvature of the DNS-SV model. Using realized volatility to augment our algorithm should make estimation of the stochastic volatility parameters more accurate and produce a superior predictive distribution. Crucially, we need to find an appropriate linkage between our volatility proxy - realized volatility - and the stochastic volatility in the model. Given the definition of the model-implied conditional volatility, we propose[3]

$$RV_t \approx Var_{t-1}(y_t) = diag(\Lambda_f H_t \Lambda_f' + Q) \tag{4}$$

[2] This formulation implies that shocks to the factors, $\{\eta_{l,t}, \eta_{s,t}, \eta_{c,t}\}$ are independent of each other. We have a discussion regarding this assumption later in the paper. We can relax this assumption by decomposing the covariance matrix η_t as in Cogley and Sargent (2005) and Primiceri (2005) to obtain,

$$cov(\eta_t) = C H_t C' = \begin{pmatrix} 1 & 0 & 0 \\ c_{ls} & 1 & 0 \\ c_{lc} & c_{sc} & 1 \end{pmatrix} \begin{pmatrix} exp(h_{l,t}) & 0 & 0 \\ 0 & exp(h_{s,t}) & 0 \\ 0 & 0 & exp(h_{c,t}) \end{pmatrix} \begin{pmatrix} 1 & c_{ls} & c_{lc} \\ 0 & 1 & c_{sc} \\ 0 & 0 & 1 \end{pmatrix},$$

where c_{ls}, c_{lc}, and c_{sc} are real numbers. Our main idea goes through with this formulation by redefining λ_f in equation 4 as $\widetilde{\lambda}_f = \lambda_f C$.

[3] This strategy of linking an observed volatility measure to the model is also used in other papers (Maheu and McCurdy (2011) in a univariate model and Cieslak and Povala (2015) in a multivariate context).

where RV_t is the realized volatility of bond yields, which has the same dimension as the bond yield vector y_t, Λ_f is the factor loading matrix given by equation 1, and H_t is a diagonal variance covariance matrix with diagonal elements given by $e^{h_{i,t}}$. Insofar as realized volatility provides an accurate approximation to the true underlying conditional time-varying volatility, equation 4 is the one that links this information to the model.

Upon adding measurement error, one can view equation 4 as a nonlinear measurement equation. In principle, we have several tools to handle this nonlinearity, including the particle filter. To keep estimation computationally feasible, when estimating h_t, we choose to take a first order Taylor approximation of the logarithm of this equation around a 3×1 vector $\mu_h = [\mu_{l,h}, \mu_{s,h}, \mu_{c,h}]'$ with respect to h_t. This leads to a set of linear measurement equations that links the realized volatility of the bond yields and the underlying factor volatility,

$$\log(RV_t) = \beta + \Lambda_h \widetilde{h}_t + \zeta_t, \quad \zeta_t \sim N(0, S), \tag{5}$$

where we write the logarithm of volatility in deviation form $\widetilde{h}_{i,t} = h_{i,t} - \mu_{i,h}$ for $i = l, s, c$. We assume that RV_t follows a log-normal distribution conditional on past histories of bond yields and bond yield realized volatilities. This assumption leads to a normally distributed measurement error ζ_t[4]. We call this new model the dynamic Nelson-Siegel with realized volatility (DNS-RV) model. The difference between this model and DNS-SV comes from augmenting equation 1 with a new measurement equation 5. This equation has a constant β and a factor loading Λ_h. The parameter β comes from the linearization[5] while we can interpret Λ_h as a loading for the factor volatility used to reduce the dimension of the volatility data, $\log(RV_t)$. This very naturally extends the dynamic factor model, which transforms high-dimensional data (y_t) into a few number of factors (f_t) via the factor loading matrix Λ_f. The volatility factor loading (Λ_h) is a function of other model parameters (Λ_f, Q, μ_h)

[4] A detailed derivation for equation 5 can be found in the Appendix.

[5] We estimate β as a separate parameter as in Takahashi et al. (2009). See the Appendix for more detail.

with the functional form given by the linearized version of equation 4[6]. We can view this as a model-consistent restriction on the linkage between the conditional volatility of observed data, approximated by $\log(RV_t)$, and the factor volatility h_t. For the same reasons as in the baseline DNS-SV model, we set the S matrix to be diagonal. These are interpreted as idiosyncratic errors, and we therefore do not model them to be contemporaneously or serially correlated.

In summary, the DNS-RV model introduces a new measurement equation into the state space of the DNS-SV model.

(Measurement equation)

$$\begin{aligned} y_t &= \Lambda_f f_t + \epsilon_t, \quad \epsilon_t \sim N(0,\ Q) \\ \log(RV_t) &= \beta + \Lambda_h \widetilde{h}_t + \zeta_t, \quad \zeta_t \sim N\left(0, S\right), \end{aligned} \tag{6}$$

(Transition equation)

$$\begin{aligned} f_{i,t} &= (1-\Phi_{i,f})\mu_{i,f} + \Phi_{i,f} f_{i,t-1} + \eta_{i,t}, \quad \eta_{i,t} \sim N(0,\ e^{h_{i,t}}) \\ h_{i,t} &= \mu_{i,h}(1-\phi_{i,h}) + \phi_{i,h} h_{i,t-1} + \nu_{i,t}, \quad \nu_{i,t} \sim N(0,\ \sigma^2_{i,h}) \end{aligned} \tag{7}$$

for $i = l, s, c$. In our application, both observed bond yields (y_t) and realized volatilities ($\log(RV)_t$) are 17×1 vectors. Moreover, both sets of variables have a factor structure with dynamics following the transition equations.

2.3 Alternative specifications

We have four classes of alternative specifications to compare forecasts to our baseline model. We briefly introduce them in this section and list all specifications considered in the paper in Table 1.

[6]Detailed formulas for the volatility factor loading matrix can be found in the Appendix.

Table 1 Model Specifications

Label	Factors (level, slope, curvature)	Conditional variance	Realized volatility
RW-C	Random walk	Constant	Not used
RW-SV	Random walk	log AR(1) in each factor	Not used
RW-RV	Random walk	log AR(1) in each factor	Used
DNS-C	Diagonal Φ_f	Constant	Not used
DNS-SV	Diagonal Φ_f	log AR(1) in each factor	Not used
DNS-RV	Diagonal Φ_f	log AR(1) in each factor	Used
RW-SV-RW	Random walk	Random walk	Not used
DNS-SV-RW	Diagonal Φ_f	Random walk	Not used
DNS-ME-SV	Diagonal Φ_f	log AR(1) in measurement equation	Not used
DNS-ME-RV	Diagonal Φ_f	log AR(1) in measurement equation	Used

Note: We list the specifications for the DNS model considered in this paper.

2.3.1 Dynamic Nelson-Siegel (DNS-C)

The first model is the standard Diebold-Li DNS model discussed at the beginning of the paper. It does not allow for stochastic volatility. This model has been shown to forecast the level of bond yields quite well, at times beating the random walk model of yields.

2.3.2 Dynamic Nelson-Siegel-Stochastic Volatility (DNS-SV)

The second is the DNS-SV model that adds stochastic volatility to the transition equation. It is summarized at the beginning of this section. By allowing for stochastic volatility, this model should improve upon the standard DNS model, especially in the second moments, as it can capture the time-varying volatility present in the bond yield data.

2.3.3 Dynamic Nelson-Siegel-Random Walk (RW)

The bond yield forecasting literature (e.g. van Dijk et al. (2014) and references therein) has shown that random walk specifications of the yield curve generally perform quite well. Oftentimes, the no-change forecast from a current period does best among a large group of forecasting models. It is in this sense that bond yield forecasting is difficult. Given these

results, we also augment the DNS-C, DNS-SV, and DNS-RV model classes with random walk parameterizations of the factor processes.

The empirical macroeconomic literature (Cogley and Sargent, 2005; Justiniano and Primiceri, 2008; Clark, 2011) often specifies stochastic volatility as following a random walk. Doing so reduces the number of parameters estimated while also providing a simple no-change forecast benchmark for time-varying volatility. As long-horizon bond yield volatility links with macroeconomic volatility, we also have random walk specifications for the stochastic volatilities.

2.3.4 Dynamic Nelson-Siegel-Measurement Error Stochastic Volatility (+ Realized Volatility) (DNS-ME)

Another way of modeling time-varying volatility in the bond yield data is to put the time-varying conditional volatility on the measurement errors, as opposed to the shocks in the factor equation. More specifically, we model independent AR(1) specifications for the measurement error stochastic volatilities[7]. While following this strategy greatly increases the number of parameters estimated, it could improve forecasting as each yield has its own stochastic volatility process. Therefore, as opposed to a time-varying H_t and constant Q matrix in the DNS-SV and DNS-RV setups, now H is constant while Q_t has stochastic volatility.

$$Q_t = diag(e^{\mathbf{q_t}}) \tag{8}$$

$$q_{i,t} - \mu_{q,i} = \phi_{q,i}(q_{i,t-1} - \mu_{q,i}) + \nu_{i,t}, \quad \nu_{i,t} \sim N(0,\ \overline{q}_i), \tag{9}$$

[7]Koopman et al. (2010) argue that putting the time-varying conditional volatility on the measurement errors provides an improvement for the in-sample fit of the DNS class of models. However, there is an important difference between our DNS-ME model and their model. Koopman et al. (2010) assume that there is a common factor of time-varying volatility in interest rates whereas our DNS-ME model treats the conditional volatility of each measurement error as following separate volatility processes.

for $i = 1, ..., N$ where N is the number of bond yields in the observation equation. $\mathbf{q_t}$ is a vector that collects all stochastic volatilities in the measurement errors. Q_t remains a diagonal matrix, as equation 8 shows. We again model the logarithm of the variances as independent first order autoregressive processes.

We also consider incorporating realized volatility information into this model. Doing so leads to the following relationship

$$RV_t \approx Var_{t-1}(y_t) = diag(\Lambda_f H \Lambda_f' + Q_t). \tag{10}$$

As before, we do a first order Taylor approximation of the logarithm of this equation around the 17×1 vector $\mu_{\mathbf{q}}$. We also add in measurement error for estimation. However, in contrast to the DNS-RV model, we link each element of $\log(RV_t)$ to its corresponding element in $\mathbf{q_t}$.

2.4 Discussion

Note that our DNS-SV and DNS-RV model specifications have three factors governing time-varying conditional volatility that are independent of the yield curve factors. Hence, the total number of latent factors explaining the joint distribution of bond yields is larger than that of standard no-arbitrage affine term structure models considered in the literature (Duffie and Kan, 1996). In addition, our volatility factors are unspanned volatility in the sense that the cross-sectional properties of bond yields (i.e., the yield curve) is determined by f_t in equation 6. As we are not imposing no-arbitrage restrictions, the volatility factors (h_t) do not face the dual role of fitting the cross-section and the conditional volatility dynamics of bond yields, as pointed out by Collin-Dufresne et al. (2009). Relative to the standard no-arbitrage affine term structure model, the DNS-SV and DNS-RV models provide a more flexible way to achieve our main goal of forecasting the joint distribution of bond yields. They do so by using three factors (f_t) to fit yield conditional mean dynamics and three other factors (h_t)

to fit yield volatility dynamics[8].

One more point worthy of discussion is our choice of independent $AR(1)$ dynamics for the level, slope, and curvature factors in the Dynamic Nelson Siegel model, as opposed to the more flexible $VAR(1)$ specification. We believe that such a benchmark is a particularly important model in the bond yield forecasting literature, which is why we have chosen to build the stochastic volatility off of that model.

The original Diebold and Li (2006) paper considered the question of $AR(1)$ versus $VAR(1)$ dynamics for the factors, opting for the more parsimonious $AR(1)$. The justification relied on the empirical finding that for forecasting purposes, a more parsimonious model often outperforms a more richly specified model. Other papers, including Pooter (2007), Christensen et al. (2011), Exterkate et al. (2013), have subsequently reconfirmed this finding when forecasting with the Dynamic Nelson Siegel class of models. Of particular note is the work of Christensen et al. (2011). In this paper, the authors compare $AR(1)$ to $VAR(1)$ dynamics for both the Dynamic Nelson Siegel model and its no-arbitrage counterpart. While the results in-sample suggest the dominance of the $VAR(1)$ specifications, upon moving to out-of-sample evaluation, the $AR(1)$ models outperforms their more richly specified competitors for most maturities and forecast horizons. This holds true for both the standard model and the no-arbitrage version, leading the authors to focus on the $AR(1)$ versions of the model as benchmarks.

Such results have led to many subsequent papers in the literature to adopt the $AR(1)$ assumption, including Diebold et al. (2008), van Dijk et al. (2014), and Kang (2015), as well as Diebold and Rudebusch (2012) to recommend using the $AR(1)$ model. In the stochastic volatility context, several previous papers have also chosen to embrace the $AR(1)$ model, including Hautsch and Ou (2012) and Christensen et al. (2014). The Hautsch and Ou (2012) paper is especially relevant for our purposes, as it includes stochastic volatility in the

[8]It seems that at least more than one volatility factor is essential to explain the bond yield distribution in the affine term structure model framework (See for example, Creal and Wu, 2014; Cieslak and Povala, 2015).

Dynamic Nelson Siegel model without imposing arbitrage. The authors compare the $AR(1)$ and $VAR(1)$ specifications, but choose to use the $AR(1)$ model as they find that many of the off-diagonal terms in the $VAR(1)$ specification are not significant.

While we consider the $AR(1)$ specification for the factor dynamics due both to its good forecasting performance and popularity in the literature, we must make one caveat. The forecasting results discussed above that have led the bond yield forecasting literature to adopt the more parsimonious model come from point forecast evaluations. For density forecasts, the question of $AR(1)$ versus $VAR(1)$ factor dynamics in a model with stochastic volatility is still an interesting open question. Although the paper does not directly consider this issue, the in-sample results of Hautsch and Ou (2012) suggest that perhaps the more parsimonious $AR(1)$ model with stochastic volatility will continue to do well. We focus on the more parsimonious case given the guidance of the previous literature. One important point to note, however, as shown in footnote 2 and discussed in more detail in a later section of the paper, is that in theory our DNS-RV model is fully consistent with VAR factor dynamics.

3 Data

We use a panel of unsmoothed Fama and Bliss (1987) U.S. government bond yields at the monthly frequency with maturities of 3, 6, 9, 12, 15, 18, 21 months and 2, 2.5, 3, 4, 5, 6, 7, 8, 9, 10 years from January 1981 to December 2009. This dataset is provided by Jungbacker et al. (2013)[9]. To construct the monthly realized volatility series, we use daily U.S. government bond yield data with the same maturities from January 2, 1981 to December 30, 2009 taken from the Federal Reserve Board of Governors with the methodology of Gürkaynak et al. (2007) [10] [11].

[9] http://qed.econ.queensu.ca/jae/datasets/jungbacker001/

[10] http://www.federalreserve.gov/econresdata/researchdata/feds200628_1.html

[11] Because most papers in the literature estimate the DNS model with unsmoothed Fama and Bliss data, we generate and evaluate predictions for the unsmoothed Fama and Bliss data. Unfortunately, this data is only available at the monthly frequency. Even though the two datasets use different methodologies, monthly

We construct the realized variance of each month's yields using daily bond yield data. The formula for realized variance at time t is

$$RV_t = \sum_{d=1}^{D} \left(\Delta y_{t+\frac{d}{D}} \right)^2 .$$

where D is the number of daily data in one time period t. This formula converges in probability to the true conditional variance as the sampling frequency goes to infinity under assumptions laid out in Andersen et al. (2003). Usually, there are around 21 days in each month, with less depending upon the number of holidays in a month that fall on normal trading days.

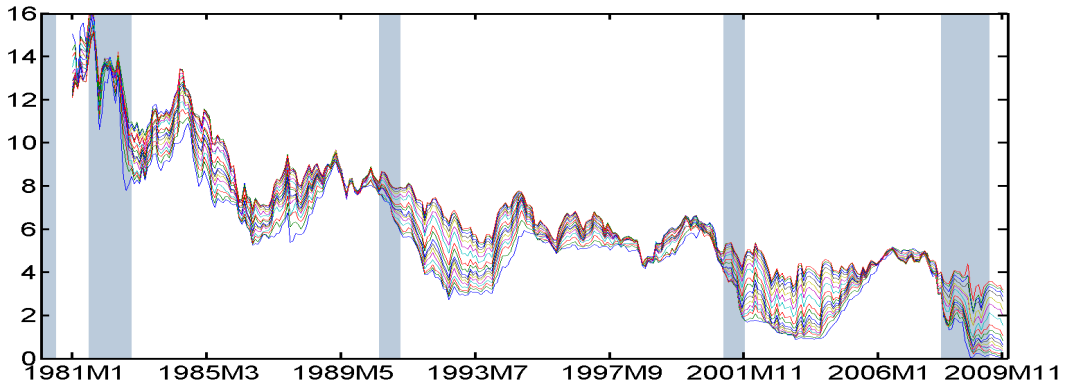
We use daily data to construct our realized volatilities for a few reasons. First, we want to use realized volatility information starting in 1981 to use a sample period similar to other bond yield forecasting studies. The availability of higher-frequency intraday data begins much later. For instance, Cieslak and Povala (2015) start their estimation in 1992 for specifically this reason. Second, the month-to-month volatility movements we want the volatility proxy to capture do not necessitate using ultra-high frequency data. Finally, while results may improve with higher frequency data, we show that positive effects are present even with lower frequency realized volatility.

Figure 1 plots the time series of monthly U.S. government bond yields and logarithm of realized volatilities. All yields exhibit a general downward trend from the start of the sample. For around the first 25 months, the realized volatility seems quite high and exhibits large time variation. After around 1983, yield volatility dies down and largely exhibits only temporary spikes in volatility. For a period of 2 years starting in 2008, the realized volatility picks up across all yields. We attribute this to the financial crises. Another interesting feature of log realized volatility is that it shows large autocorrelation. Its first-order autocorrelation coefficients range from 0.59 to 0.69 and the 12th-order autocorrelation coefficients range from

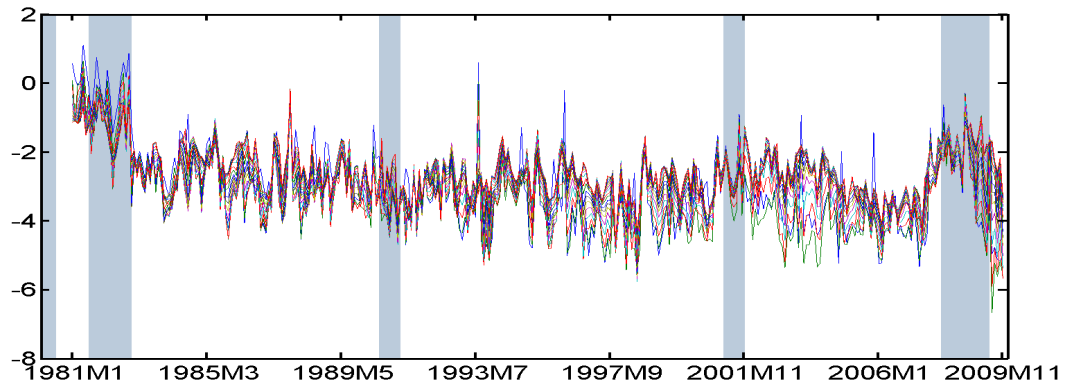
yield data based on the daily bond yield from Gürkaynak et al. (2007) is very close to the one based on the unsmoothed Fama and Bliss method.

Figure 1 U.S. Treasury Yields

(a) Yields (monthly, annualized %)



(b) Yield Realized Volatilities (Monthly, Log)



Notes: We present monthly U.S. Treasury yields with maturities of 3, 6, 9, 12, 15, 18, 21 months and 2, 2.5, 3, 4, 5, 6, 7, 8, 9, 10 years over the period January 1981 – November 2009. Monthly yields are constructed using the unsmoothed Fama-Bliss method. Monthly yield realized volatilities are constructed based on daily yields using Wright's dataset. Blue shaded bars represent NBER recession dates.

0.20 to 0.31[12]. This means that realized volatility data could help even for the long-horizon forecasts that we consider in this paper.

Both the yields and realized volatilities do seem to exhibit a factor structure, meaning that each set of series co-move over time. In fact, a principal components analysis shows that the first three factors for yields explain 99.95% of the variation in the U.S. yield curve. The first three factors for realized volatilities explain 98.53% of the variation (Table 2). Although

[12]We provide tables in the supporting material for the descriptive statistics of monthly realized volatility of bond yields.

Table 2 Variance explained by the first five principal components (cumulative, %)

	Yield	log(RV)
pc 1	98.16	84.30
pc 2	99.84	94.62
pc 3	99.95	98.53
pc 4	99.98	99.46
pc 5	99.98	99.85

Notes: Numbers in the table are the cumulative percentage of total variance explained by the first five principal components for U.S. Yield data and log(RV). log(RV) is the logarithm of the monthly realized volatility constructed based on the daily U.S. yield data.

the fact that U.S. bond yields can be explained by the first few principal components is well documented, it is interesting that the same feature carries over to the realized volatility of U.S. bond yields.

4 Estimation/Evaluation Methodology

4.1 Estimation

We perform a Gibbs sampling Markov Chain Monte Carlo algorithm for $15,000$ draws. We keep every 5th draw and burn in for the first $5,000$ draws. Due to our linearization approximation in introducing realized volatility, all specifications that we consider can be sampled by using the method developed in Kim et al. (1998) for the stochastic volatility state space model. Details of the state space representation and the estimation procedure can be found in the appendix. Details on the prior can be found in the appendix as well, although we comment that our choice of prior is loose and is not expected to impact the estimation results radically.

We highlight the difference in estimation procedure due to the additional measurement equation for the realized volatility. Roughly speaking, there are two sources of information for the latent volatility factor h_t. The first source is from the latent factor f_t. To see this,

one can transform the transition equation for f_t as follows

$$\log\left((f_{i,t} - \mu_{i,f}(1-\phi_{i,f}) - \phi_{i,f} f_{i,t})^2\right) = h_{i,t} + \log(x_{i,t}^2), \quad x_{i,t} \sim N(0,1) \tag{11}$$

for $i = l, s, c$. This is common to both DNS-SV and DNS-RV. The second source is from equation 6, which relates $\log(RV_t)$ with h_t and is unique to DNS-RV. For the estimation of the DNS-RV model, we augment the realized volatility measurement equation to the Kim et al. (1998)'s state space model representation defined by equation 3 and equation 11. Then, conditional on the other parameters and data, extraction of h_t amounts to running a simulation smoother in conjunction with the Kalman filter with and without equation 6.

4.2 Forecast evaluation

We consider model performance along both the point and density forecasting dimensions. The appendix contains further details on the Bayesian simulation algorithm we use to generate the forecasts. We begin forecasting on February 1994 and reestimate the model in an expanding window and forecast moving forward two months at a time. For every forecast run at a given time t, we forecast for all yields in our dataset and for horizons ranging from 1 month to 12 months ahead. This leads to a total of 94 repetitions.

Point forecast

To evaluate the point prediction, we use the Root Mean Square Forecast Error (RMSE) statistic,

$$RMSE^M_{\tau,ho} = \sqrt{\frac{1}{F}\sum \left(\hat{y}^M_{t+ho}(\tau) - y(\tau)_{t+ho}\right)^2}. \tag{12}$$

Call the yield τ forecast at horizon ho made by model M as $\hat{y}^M_{t+ho}(\tau)$ and $y(\tau)_{t+ho}$ the realized value of the yield at time $t + ho$. F is the number of forecasts made. Then, equation 12 provides the formula for the RMSE. To gauge whether there are significant differences in the

RMSE, we use the Diebold and Mariano (1995) t-test of equal predictive accuracy[13].

Density forecast

The log predictive score (Geweke and Amisano, 2010) gives an indication of how well a model performs in density forecasting,

$$LPS^{M}_{\tau,ho} = \frac{1}{F}\sum \log p(y_{t+ho}(\tau)|y^t, M). \tag{13}$$

where $p(y_{t+ho}(\tau)|y^t, M)$ denotes the ho-step ahead predictive distribution of yield τ generated by model M given time t information. Following Carriero et al. (2013), we estimate the log predictive density by a kernel density estimator using MCMC draws for parameters and latent states and compute the p-value for the Amisano and Giacomini (2007) t-test of equal means to gauge whether there exist significant differences in the log predictive score[14].

5 Results

We first present in-sample results of the model, focusing on time-varying volatility. Then, we move to point and density forecasting results.

5.1 In-sample

We first present the full sample estimation from January 1981 - November 2009. We focus on how adding realized volatility information alters the model estimates. Adding in second-moment information does not significantly change conditional mean dynamics, so we relegate

[13]When models are nested (as is sometimes the case in our paper (e.g. DNS-C versus DNS-SV)), Diebold and Mariano (1995)'s testing procedure with Gaussian critical value turns out to be conservative. However, it still provides a rough idea about the size of difference in RMSEs across models. Discussion and application of the Diebold-Mariano test in a nested model comparison environment can be found in Carriero et al. (2015) and Clark and Ravazzolo (2015).

[14]Strictly speaking, p-values from this test are valid when the forecasting exercise is done with rolling window estimation/prediction as opposed to recursive estimation/prediction. However, we think that it still provides a rough idea about the size of difference in LPSs across models (For example, Carriero et al., 2015, follow the same strategy.).

Table 3 Posterior Estimates of Parameters on h_t Equation

	DNS-SV			DNS-RV		
	5%	50%	95%	5%	50%	95%
$\mu_{l,h}$	-4.40	-2.56	-1.48	-4.19	-3.92	-2.21
$\mu_{s,h}$	-3.41	-2.30	-1.67	-3.22	-2.85	-1.94
$\mu_{c,h}$	-1.48	-0.96	-0.47	-2.18	-1.88	-1.39
$\phi_{l,h}$	0.93	0.98	0.999	0.58	0.66	0.73
$\phi_{s,h}$	0.92	0.97	0.995	0.57	0.64	0.72
$\phi_{c,h}$	0.81	0.92	0.98	0.39	0.49	0.60
$\sigma_{l,h}$	0.01	0.03	0.08	0.41	0.75	0.92
$\sigma_{s,h}$	0.01	0.04	0.10	1.27	1.72	2.64
$\sigma_{c,h}$	0.02	0.09	0.27	1.30	1.78	4.30

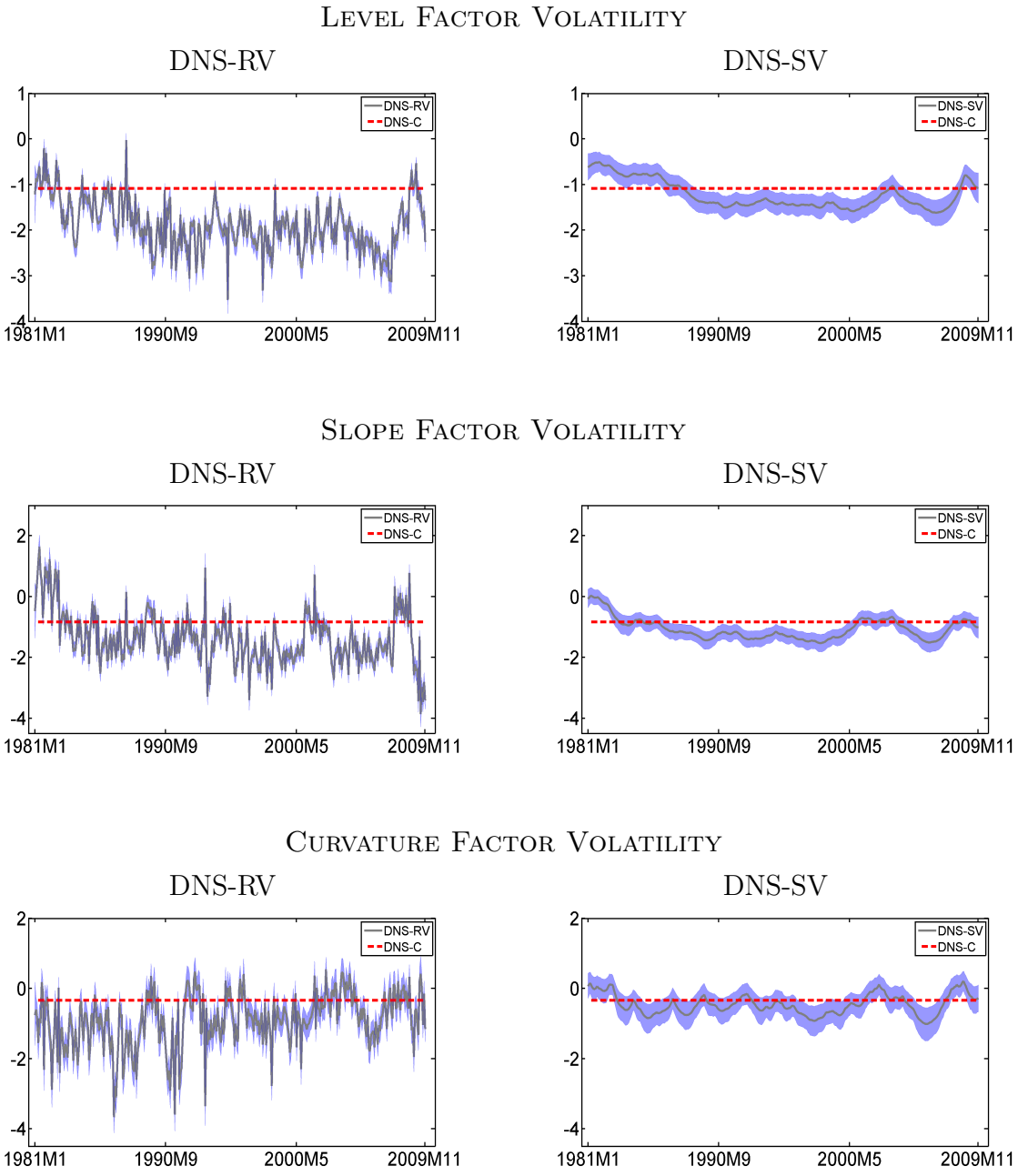
Notes: Posterior moments are based on estimation sample from January 1981 to November 2009.

our discussion of the extracted factors to the appendix[15]. Our extracted factors are similar to those found in Diebold and Li (2006).

The stochastic volatility dynamics deserve some more precise discussion. Figure 2 shows the volatility estimates from the DNS-C, DNS-SV, and DNS-RV models. The fluctuations of the extracted stochastic volatilities in both the DNS-SV and DNS-RV models show that there exists conditional time-varying volatility in the data. Relative to the DNS-SV specification, adding in realized volatility data makes the extracted stochastic volatility much less persistent and more variable. This leads to a lower autoregressive parameter and higher innovation standard deviation estimates for all of the stochastic volatility processes (Table 3). The DNS-RV model also delivers lower stochastic volatility mean estimates. These differences lead to differences in forecasting. For example, the lower autoregressive parameter in the DNS-RV model means that it predicts faster mean reversion of the stochastic volatilities relative to the DNS-SV model and the lower long-run mean estimate implies that the DNS-RV model produces a tighter density prediction in the long run. The smoothed stochastic volatilities from the DNS-SV model, however, generally captures the low-frequency volatility movements from the DNS-RV model.

[15]All other in-sample estimation results are in the appendix.

Figure 2 Stochastic volatility for bond yield factors



Notes: Posterior median of log stochastic volatility (h_t) for bond yield factors from DNS-RV (left column) and DNS-SV (right column). Red dotted line is volatility level estimated from DNS-C. Blue band is 80% credible interval. Estimation sample is from January 1981 to November 2009.

Table 4 RMSE COMPARISON

Maturity	RW-C	RW-SV	RW-RV	DNS-C	DNS-SV	DNS-RV	RW-SV-RW	DNS-SV-RW	DNS-ME-SV	DNS-ME-RV
1-step-ahead prediction										
3	0.267	0.992	1.011	1.004	1.034**	1.007	**0.991**	1.038**	1.071	0.953**
12	**0.229**	1.005	1.002	1.076**	1.055**	1.003	1.002	1.060**	1.070**	1.072**
36	0.274	1.001	0.996	1.015	1.020	**0.990**	1.001	1.025	1.010	1.012
60	0.274	1.000	0.995	1.003	1.010	**0.987**	1.000	1.013	1.002	1.004
120	0.277	1.001	0.996	0.988	0.989	0.999	1.000	0.988	**0.986**	0.997
3-step-ahead prediction										
3	**0.506**	1.002	1.011	1.036	1.066**	1.018	1.002	1.069**	1.055	1.024
12	0.537	1.002	**0.999**	1.072	1.073**	**0.999**	1.002	1.077**	1.064	1.063
36	0.580	1.000	0.998	1.022	1.036	**0.990**	0.999	1.040	1.018	1.018
60	0.551	1.000	0.998	1.004	1.017	**0.986**	0.999	1.021	1.005	1.004
120	0.489	1.000	0.998	**0.982**	0.987	0.995	1.000	0.988	0.987	0.987
6-step-ahead prediction										
3	0.932	1.000	1.002	1.003	1.047	1.007	**0.999**	1.048	1.010	1.000
12	0.915	1.001	1.001	1.040	1.065	**0.999**	1.000	1.068	1.036	1.034
36	0.881	1.001	1.001	1.017	1.041	**0.987**	1.000	1.046	1.011	1.011
60	0.819	1.000	0.998	0.998	1.016	**0.978**	1.000	1.020	0.997	0.997
120	0.665	0.999	0.992*	**0.977**	0.982	0.980	0.998	0.986	0.981	0.980
12-step-ahead prediction										
3	1.592	1.002*	0.999	**0.936****	1.013	0.991	1.000	1.015	0.943**	0.940*
12	1.476	1.001	0.999	**0.975**	1.034	0.991	1.000	1.036	0.976	**0.975**
36	1.242	1.001	1.001	0.997	1.043	**0.979****	1.001	1.046	0.993	0.993
60	1.069	1.002	1.000	0.997	1.031	**0.967***	1.001	1.033	0.994	0.994
120	0.833	1.002	0.998	0.976	0.994	**0.966**	1.001	0.994	0.980	0.976

Notes: The first column shows the RMSE based on the RW-C. Other columns show the relative RMSE compared to the first column. The RMSE from the best model for each variable and forecast horizon is in bold letter. Units are in percentage points. Divergences in accuracy that are statistically different from zero are given by * (10%), ** (5%), *** (1%). We construct the p-values based on the Diebold and Mariano (1995) t-statistics with a variance estimator robust to serial correlation using a rectangular kernel of $h-1$ lags and the small-sample correction proposed by Harvey et al. (1997).

We argue that this difference in the stochastic volatilities matters for density forecasting. The high-frequency data used to construct the realized volatilities brings information that the low-frequency monthly yield data misses. By having more accurate estimates of the current level of time-varying volatility and volatility process parameters, the DNS-RV model both starts off forecasting at a more accurate point and better captures the dynamics of the data moving forward. We provide a careful decomposition of these two effects when analyzing our density prediction results in Section 5.4.

5.2 Point prediction

Table 4 shows the RMSE of selected maturities for 1, 3, 6, and 12-step ahead predictions. The second column has the calculated RMSE values for the RW-C model. All other values reported are ratios relative to the RW-C RMSE. Values below 1 indicate superior performance relative to the random walk benchmark. Stars in the table indicate significant gains relative to the RW-C model. As expected, the RMSE increases as the forecasting horizon lengthens. The models with random walk dynamics in the factors do well for short-horizon forecasts but deteriorate when compared to the stationary models as the prediction horizon lengthens. In general, all RMSE values have numbers close to 1, reproducing the well-known result in the bond yield forecasting literature on the difficulty in beating the no-change forecast. As alluded to in the previous section, adding in time-varying second moments does not largely impact point predictions, although the DNS-RV model forecasts middle maturities well across all horizons. For 12-month horizon forecasts, the DNS-C and DNS-ME-RV models also do well for short maturity yields.

5.3 Density prediction

Table 5 shows the density evaluation result in terms of the log predictive score. Similar to the RMSE table, the RW-C column gives the value of the log predictive score for the random walk case while the numbers for the other models are differences relative to that column. A higher value indicates larger log predictive score and better density forecasting results. As opposed to the point prediction results, three interesting findings emerge when we consider the log predictive score.

First, for the short-run horizon, having realized volatility gives significant gains in density prediction. This is on top of a large improvement in log predictive score from adding stochastic volatility, which Carriero et al. (2013) find. We conduct pairwise tests for differences in the log predictive score for the DNS-RV forecasts and the forecasts from the other models

Table 5 Log predictive score comparison

Maturity	RW-C	RW-SV	RW-RV	DNS-C	DNS-SV	DNS-RV	RW-SV-RW	DNS-SV-RW	DNS-ME-SV	DNS-ME-RV
1-step-ahead prediction										
3	-0.538***	0.205***	0.343***	0.008***	0.198***	**0.375**	0.181***	0.173***	0.048***	0.102***
12	-0.357***	0.211***	0.304	-0.009***	0.195***	**0.316**	0.189***	0.172***	0.012***	0.013***
36	-0.318***	0.129	0.158**	0.000***	0.118	**0.176**	0.111	0.104	0.022**	0.017**
60	-0.265*	**0.109**	0.077**	-0.001	0.098	0.105	0.098	0.088	0.019	0.015
120	-0.248*	**0.110**	0.111	0.011	**0.121**	0.110	0.106	0.108	0.048	0.037
3-step-ahead prediction										
3	-1.061***	0.183***	0.302	0.016***	0.165***	**0.322**	0.151***	0.133***	0.036***	0.061***
12	-0.985***	0.107**	0.195	-0.012***	0.064**	**0.196**	0.089**	0.045***	0.007***	0.010***
36	-0.951	0.027	0.053	-0.002	-0.002	**0.067**	0.013	-0.018	0.009	0.007
60	-0.882	0.013	0.016	0.003	-0.011	**0.035**	0.007	-0.015	0.011	0.010
120	-0.779	0.033	0.020	0.014	0.033	0.031	**0.036**	**0.038**	0.027	0.029
6-step-ahead prediction										
3	-1.485	0.048**	0.157	0.028	0.023**	**0.159**	0.033**	0.005**	0.038	0.053
12	-1.417	0.010	0.065	0.000	-0.058	**0.066**	0.000	-0.082	0.008	0.012
36	-1.342	-0.017	0.021	0.001	-0.077	**0.030**	-0.039*	-0.095	0.010	0.012
60	-1.263	-0.031	0.008	0.010	-0.072	**0.018**	-0.045	-0.083	0.016	0.016
120	-1.100	0.031	0.049	0.023	0.038	**0.061**	0.029	0.033	0.036	0.034
12-step-ahead prediction										
3	-1.940	-0.096	0.031	**0.091**	-0.098	0.041	-0.096*	-0.119	0.086	0.094
12	-1.850	-0.071	-0.012	**0.048**	-0.147	-0.004	-0.089	-0.168	0.048	0.051
36	-1.688	-0.017**	-0.001	0.023	-0.122	0.019	-0.049***	-0.137	0.032	**0.033**
60	-1.560	0.005***	0.036	0.024	-0.068	**0.066**	-0.011***	-0.073	0.034	0.036
120	-1.387**	0.099***	0.151	0.030	0.101	**0.193**	0.098***	0.092	0.047*	0.045*

Notes: The first column shows the log predictive score based on the RW-C. Other columns show the difference of log predictive score from the first column. Log predictive score differences represent percentage point differences. Therefore, a difference of 0.1 corresponds to a 10% more accurate density forecast. The log predictive score from the best model is in bold letter for each variable and forecast horizon. We construct the p-values from Amisano and Giacomini (2007) tests comparing the hypothesis of equal log predictive score of the DNS-RV with alternative models. Divergences in accuracy that are statistically different from zero are given by * (10%), ** (5%), *** (1%). Test statistics are computed with a variance estimator robust to serial correlation using a rectangular kernel of $h-1$ lags and the small-sample correction proposed by Harvey et al. (1997).

(the stars in Table 5.3 show the results of these tests). They indicate that the DNS-RV model has significantly higher log predictive score values for one- and three-month ahead predictions for short term maturities when compared to most competitors[16]. By producing improved estimates of the current state of volatility, we would expect that short horizon forecasts have the largest gain. The improved density forecasting performance for the DNS-RV and RW-RV models continues even up to a 6-month forecasting horizon. At one year ahead, most models with realized volatility have their volatility processes returning close to the unconditional mean, so the gain diminishes.

[16]We also compute the model confidence set of Hansen et al. (2011) and find a similar result. See the supporting material for the model confidence set results.

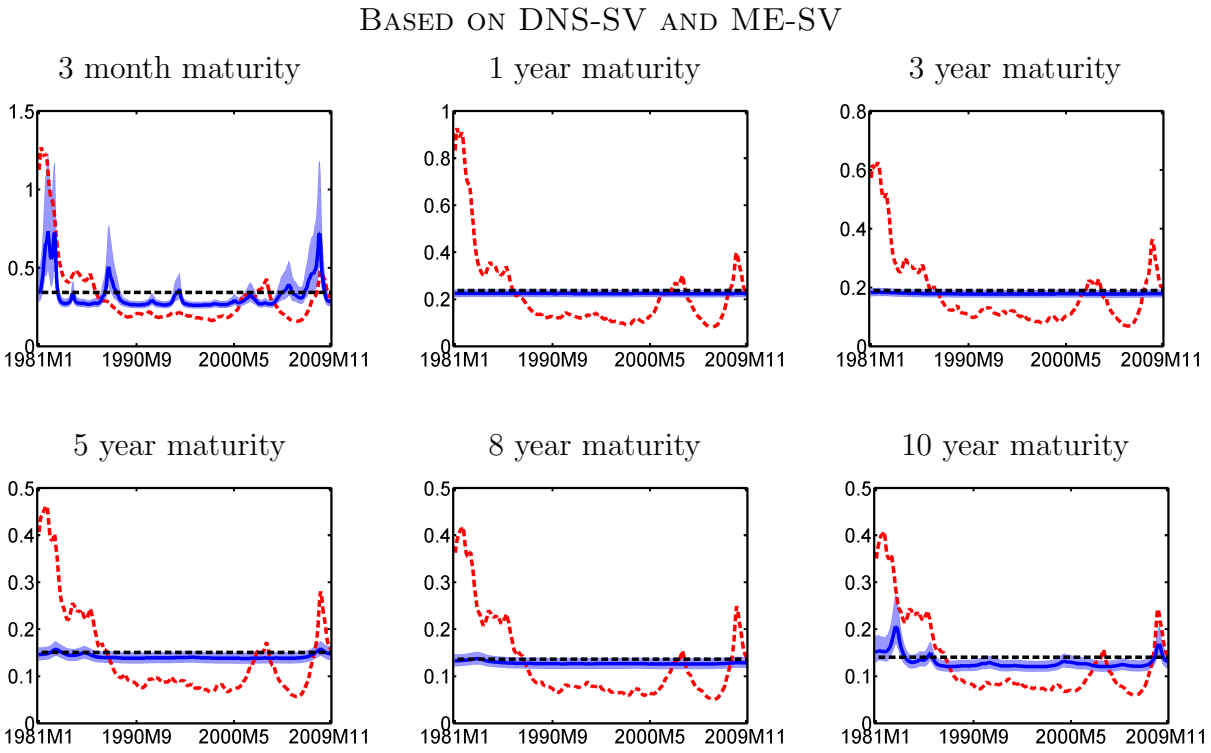
Second, comparing RW-SV to RW-SV-RW, RW-RV to RW-RV-RW, and DNS-SV to DNS-SV-RW shows that given a fixed conditional mean specification, a random walk specification on conditional volatility dynamics in general leads to poorer results. This illustrates the fact that even though conditional mean dynamics of bond yields approximate a random walk in our sample, conditional volatility dynamics exhibit mean reversion. Bond yields therefore do have forecastability, although simply looking at the conditional mean dynamics do not reveal this fact strongly.

Third, an alternative specification for introducing stochastic volatility into the model by putting it on the measurement equation does not forecast as well as the specification with stochastic volatility on the transition equation. Comparing DNS-SV to DNS-ME-SV shows that for short horizon forecasts, DNS-SV performs better whereas for longer horizon forecasts, DNS-ME-SV does better. A similar story holds when looking at DNS-RV and DNS-ME-RV, although DNS-RV does better even up to 6-month horizon forecasts with mixed 12-month horizon results. The measurement error specifications give consistent improvements in the log predictive score over and above the constant volatility models, although the gains are small.

One fact holding back the performance of the measurement error specifications is that the measurement error variance explains a small portion of total bond yield variance. For example, the ratio of the standard deviations of smoothed measurement errors to the standard deviations of smoothed factors in the DNS-ME-SV model is often below 3 percent and never above 8 percent[17]. In Figure 3, we see that the model with stochastic volatility on the measurement errors does not generate movements in the conditional time-varying volatility of various middle maturity yields. In fact, the conditional variances of the 1-year, 3-year, 5-year, and 8-year maturities are nearly on top of the black dotted line, which is the variance of the yields implied by the DNS-C model. Therefore, putting time-varying volatility in the

[17]This fact is similar across different model specifications. The values from all models are presented in the supporting materials.

Figure 3 Stochastic variance of individual yield



Notes: Red dotted line: Conditional variance from DNS-SV. Blue solid line: Conditional variance from DNS-ME-SV with 80% credible interval. Black dotted horizontal line: Conditional variance from DNS-C. Estimation sample is from January 1981 to November 2009.

measurement errors does not drastically change the model-implied predictive distributions. This explains why the density forecasting performance for this class of models mimics that of the DNS-C model.

In contrast, the Figure 3 shows that putting stochastic volatility in the shocks to the bond yield factor can better capture time-varying volatility. The DNS-SV model-implied time-varying volatility consistently fits the narrative evidence of the Great Moderation from the mid-1980's until the mid-2000's. It also picks up the increases in volatility from the early 2000's recession and recent financial crises.

5.4 How does RV information help density prediction?

In the previous section, we provide evidence that RV information improves the quality of density prediction. In this section, we investigate how RV information leads to better density prediction. Broadly speaking, there are two potential major sources of prediction gains by incorporating RV information. The first is from accurate parameter estimation. The second is from realistic volatility state estimation (in other words, the volatility estimates are closer to the true latent volatility). The latter effect is mostly coming from having more realistic volatility state estimates at the time of the prediction.

To decompose the forecast gains from the incorporation of RV information, we generate two additional forecasts based on posterior draws from the DNS-SV model and the DNS-RV model. Note that the construction of the predictive distribution can be done by simulating future paths of the bond yields based on the model at hand. This requires two main ingredients. The first is a set of model parameters (Θ) and the second is the state estimate at the current period (the end point of the estimation sample, h_T and f_T). The latter plays a role as a starting point of the predictive path simulation.

We construct two additional predictive distributions by propagating the model using the posterior draws, $\{\Theta^{(s)}, f_{1:T}^{(s)}\}_{s=1}^{S}$ and $\{h_{1:T}^{(s)}\}_{s=1}^{S}$ from the DNS-SV model and the DNS-RV model:

1. **DNS-SV-SVend:** We construct the predictive distribution using $\{\Theta^{(s)}, f_{1:T}^{(s)}\}_{s=1}^{S}$ from the DNS-SV model estimation and fixing the current level of the volatility state (h_T) at the posterior mean based on the DNS-SV model.
2. **DNS-RV-SVend:** We construct the predictive distribution using $\{\Theta^{(s)}, f_{1:T}^{(s)}\}_{s=1}^{S}$ from the DNS-RV model estimation and fixing the current level of the volatility state (h_T) at the posterior mean based on the DNS-SV model.

We decompose the impact of RV information on density prediction performance (log predictive score) through pairwise comparisons of the different predictive distributions: the

difference between DNS-SV and DNS-SV-SVend captures the effect of fixing volatility states at the posterior mean; the difference between DNS-SV-SVend and DNS-RV-SVend captures the effect of accurate parameter estimation; the difference between DNS-RV-SVend and DNS-RV captures the effect of realistic volatility estimation[18].

The Table 6 presents the density evaluation results in terms of the log predictive score. For consistency with the previous table, the RW-C column shows the value of the log predictive score for the random walk case while the numbers for the other models are differences relative to that column.

Several findings emerge from this table. First, there is not much difference between the DNS-SV model and the DNS-SV-SVend model, indicating that fixing the volatility state at the posterior mean does not have a large impact on density prediction performance. Second, for h=1, most of the RV information gain comes from the realistic estimation of the volatility state. This can be seen from the big jump in LPS from the DNS-RV-SVend model to the DNS-RV model. For example, 80% $\sim$ 98% of LPS gain from the DNS-SV model to the DNS-RV model comes from realistic volatility estimation for short maturities (3 month, 12 month, 36 month). Third, as we move from short-run prediction (h=1) to longer-run prediction (h=3, 6, and 12), the gain from realistic volatility estimation drops and the gain from accurate parameter estimation increases. For example, for h=3, the gain from realistic volatility estimation ranges from 24% to 81% and for h=6 they range from 1% to 37% for short to medium maturities (3, 12, 36 month maturities).

Based on our findings, we can conclude that RV information plays a different role in density prediction depending on the forecasting horizon. For short-run prediction, the gain mostly comes from the realistic volatility state estimation while the gain from accurate parameter estimation helps for medium- to long-run predictions (h=3, 6, 12). This is quite

[18]Note that we are not taking account the joint (interactive) effect. Specifically, good volatility estimates could lead to good parameter estimates or good parameter estimates could lead to good volatility estimates. We view this exercise as just a rough calibration of the two different effects.

Table 6 Decomposition of gain from incorporating RV information

Maturity	RW-C	DNS-SV	DNS-SV-SVend	DNS-RV-SVend	DNS-RV
1-step-ahead prediction					
3	-0.538	0.198	0.206	0.223	**0.375**
12	-0.357	0.195	0.209	0.211	**0.316**
36	-0.318	0.118	0.127	0.138	**0.176**
60	-0.265	0.098	0.110	**0.116**	0.105
120	-0.248	**0.121**	0.116	0.105	0.110
3-step-ahead prediction					
3	-1.061	0.165	0.160	0.190	**0.322**
12	-0.985	0.064	0.062	0.127	**0.196**
36	-0.951	-0.002	-0.009	0.049	**0.067**
60	-0.882	-0.011	-0.017	0.041	**0.035**
120	-0.779	0.033	0.036	**0.056**	0.031
6-step-ahead prediction					
3	-1.485	0.023	0.006	0.103	**0.159**
12	-1.417	-0.058	-0.083	0.045	**0.066**
36	-1.342	-0.077	-0.099	0.029	**0.030**
60	-1.263	-0.072	-0.094	**0.029**	0.018
120	-1.100	0.038	0.036	**0.079**	0.061
12-step-ahead prediction					
3	-1.940	-0.098	-0.148	0.023	**0.041**
12	**-1.850**	-0.147	-0.205	-0.010	-0.004
36	-1.688	-0.122	-0.136	**0.020**	0.019
60	-1.560	-0.068	-0.073	**0.074**	0.066
120	-1.387	0.101	0.106	0.191	**0.193**

Notes: The first column shows the log predictive score based on the RW-C. Other columns show the difference of log predictive score from the first column. Log predictive score differences represent percentage point differences. Therefore, a difference of 0.1 corresponds to a 10% more accurate density forecast. The log predictive score from the best model is in bold letter for each variable and forecast horizon.

intuitive. The bond yield factor volatilities follow fairly persistent, but still stationary processes. Therefore, in predicting the second moment, a good short-run prediction requires a good estimate about the current volatility level. Because a prediction based on a stationary process tends to converge to its long-run mean, a good estimate about the long-run mean of the volatility process is necessary for superior long-run density prediction. This is exactly what we find from our decomposition, namely that realistic volatility state estimation is important for the short run, whereas accurate volatility process parameter estimates are important for the long run.

5.5 General second moment factor dynamics.

So far, we have only allowed each volatility factor to evolve independently. Alternatively, one could allow the factor volatilities to follow a VAR process with correlated shocks. We call this model DNS-RV-Full. We present the density prediction evaluation results (in terms of log predictive score) in Table 7 together with the results based on DNS-SV and DNS-RV. All numbers are differences between the LPS of the corresponding model and RW-C to make all numbers comparable to other tables in the paper.

From this table, several findings emerge. First, the improvement in density prediction from allowing for general volatility dynamics is not as drastic as the improvement from incorporating RV information into the DNS-SV model. Second, the DNS-RV model does better than the DNS-RV-Full model for short/medium maturity yields, while the DNS-RV-Full model tends to have higher LPS values for longer maturity yields. The reason why allowing for general volatility dynamics does not always improve the quality of the density prediction is quite straightforward. When forecasting, there is a tension between model complexity and simplicity. In this case, allowing for a more complex model seems to have mixed results when compared to the simpler one. Overall, however, the DNS-RV-Full model still outperforms DNS-SV.

5.6 Potential extensions

5.6.1 General first moment factor dynamics

As this paper focuses on the importance of realized volatility data in inferring conditional volatility dynamics for forecasting purposes, we use a popular benchmark $AR(1)$ model of the factors to model the conditional mean dynamics[19].

[19] We justify our choice in section 2.4.

Table 7 LOG PREDICTIVE SCORE COMPARISON: GENERAL DYNAMICS FOR h_t

Maturity	RW-C	DNS-SV	DNS-RV	DNS-RV-Full
1-step-ahead				
3	-0.538	0.198	**0.375**	0.335
12	-0.357	0.195	**0.316**	0.270
36	-0.318	0.118	**0.176**	0.161
60	-0.265	0.098	0.105	**0.112**
120	-0.248	0.121	0.110	**0.131**
12-step-ahead				
3	-1.940	-0.098	**0.041**	0.015
12	**-1.850**	-0.147	-0.004	-0.033
36	-1.688	-0.122	**0.019**	0.019
60	-1.560	-0.068	0.066	**0.086**
120	-1.387	0.101	0.193	**0.203**

Notes: The first column shows the log predictive score based on the RW-C. Other columns show the difference of log predictive score from the first column. Log predictive score differences represent percentage point differences. Therefore, a difference of 0.1 corresponds to a 10% more accurate density forecast. The log predictive score from the best model is in bold letter for each variable and forecast horizon.

It is possible, however, to extend the DNS-RV model to allow for VAR dynamics for the factors. Although we do not do so in the paper, we describe here the necessary ingredients for an extension to VAR factor dynamics. Equation 14 presents the modified model:

$$\begin{aligned} f_t &= (I_3 - \Phi_f)\mu_f + \Phi_f f_{t-1} + C\eta_t, \quad \eta_t \sim N(0,\ H_t = diag(e^{[h_{l,t}, h_{s,t}, h_{c,t}]})) \\ h_{i,t} &= \mu_{i,h}(1-\phi_{i,h}) + \phi_{i,h} h_{i,t-1} + \nu_{i,t}, \quad \nu_{i,t} \sim N(0,\ \sigma^2_{i,h}) \ \ i = l, s, c \end{aligned} \tag{14}$$

Φ_f is now allowed to be a full matrix and C is a lower triangular matrix with 1s on the main diagonal. Standard Bayesian Gibbs sampling techniques can handle the estimation of Φ_f. On the other hand, relative to previous methods (Cogley and Sargent (2005)), sampling C is now more complicated. The difficulty comes from the fact that C enters into the measurement equation for the realized volatilities. One can see this by examining equation 4 appropriately modified:

$$RV_t \approx Var_{t-1}(y_t) = diag\left(\Lambda_f C H_t C' \Lambda_f' + Q\right) \tag{15}$$

By redefining $\tilde{\Lambda}_f = \Lambda_f C$, it continues to be possible to apply the linearization technique and form the realized volatility measurement equation. One can use an importance sampler or the Metropolis-Hastings algorithm for the various elements of C, taking into account that the matrix now enters both the realized volatility measurement equation and the factor transition equation.

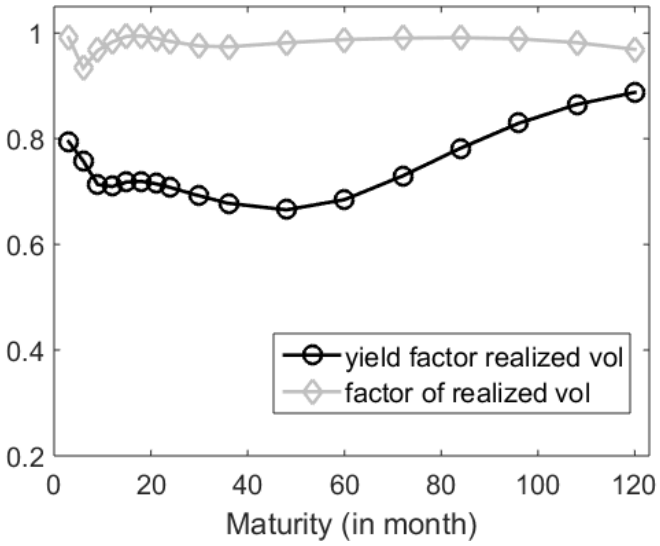
5.6.2 Volatility factor structure

Our benchmark DNS-RV model decomposes the fluctuations in yield curve volatility into the volatilities of the level, slope, and curvature factors. A natural question is how well this decomposition can explain the volatility data.

To answer this question, we conduct the following exercise. We use our daily bond yield data to construct the level, slope, and curvature factors implied by the Dynamic Nelson Siegel model day-by-day. This can easily be done by following the strategy in Diebold and Li (2006) of regressing the daily yields y_t^d on the factor loadings using the formulation found in equation 1 with ordinary least squares[20] . Doing so gives us a daily time series of factor estimates $\hat{f}_{l,t}, \hat{f}_{s,t}, \hat{f}_{c,t}$. We can then construct the monthly realized variance of these factors $RV_t^{f_i}$ for $i = l, s, c$. This gives us an empirical proxy for the volatility of the level, slope, and curvature factors. A regression of the realized volatility of the individual yields on the realized volatility of the factors can give us a sense of what fraction of yield realized volatility is explained by these volatility factors. The motivation for our regression is equation 4, which specifies a linear relationship between the yield volatilities and factor volatilities. Therefore, we run the following regression maturity-by-maturity:

[20]Consistent with Diebold and Li (2006), we fix $\lambda = 0.0609$.

Figure 4 R^2 COMPARISON OF PRINCIPAL COMPONENTS VERSUS DNS FACTOR DECOMPOSITION



Notes: Black line shows the R^2 by maturity of the yield factor realized volatility. Gray line shows the R^2 by maturity of the first three realized volatility principal components. Maturities range from 3 months to 120 months.

$$RV_t(\tau) = \alpha(\tau) + \sum_{i=l,s,c} \beta_i(\tau) RV_t^{f_i} + \epsilon_t(\tau) \tag{16}$$

Figure 4 shows the fraction of variation by maturity explained by the volatility factors compared with the fraction explained by the first three principal components of the realized volatility. The volatility factors explain the fluctuations best for longer maturity yields, topping out at around 89% of total variation. They still explain around 67% of total variation for short maturity yields (the average over maturities is around 74%). Comparing the results when using the three volatility factors to those using the first three principal components, we see that there still are non-negligible differences in the explanatory power. Overall, we conclude that the Dynamic Nelson Siegel framework with stochastic volatility has the potential to explain a majority of the bond yield volatility fluctuations, although there still may be further room for improvement. Our forecasting results show that the Dynamic Nelson Siegel model-implied decomposition of yield volatility does in fact lead to significant gains

in density prediction.

This result adds to the burgeoning literature on developing bond yield models with stochastic volatility that can simultaneously fit the yield curve and realized volatility properties of the data well (Collin-Dufresne et al., 2009; Andersen and Benzoni, 2010; Cieslak and Povala, 2015). As emphasized previously, an important feature of our model is that our volatility factors are not spanned, which allows us much more flexibility in fitting the conditional second moments. We look forward to the further development of bond yield models that can simultaneously fit the first and second moments of the yield curve.

6 Conclusion

We investigate the effects of introducing realized volatility information on U.S. bond yield density forecasting. To do so, we develop a general estimation approach to incorporate realized volatility into dynamic factor models with stochastic volatility. We compare the performance of our benchmark model DNS-RV with a variety of different models proposed in the literature and find that the DNS-RV model produces superior density forecasts, especially for the short-run. In addition to this, incorporating time-varying volatility in general improves density prediction, time-varying volatility is better modeled as a stationary process as opposed to a random walk, and time-varying volatility in the factor equation generates better density predictions when compared to time-varying volatility on the measurement equation.

References

Altavilla, C., R. Giacomini, and G. Ragusa (2013): "Anchoring the Yield Curve Using Survey Expectations," *Working Paper*.

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

Andersen, T. G. and L. Benzoni (2010): "Do Bonds Span Volatility Risk in the U.S. Treasury Market? A Specification Test for Affine Term Structure Models," *The Journal of Finance*, 65, 603–653.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003): "Modeling and Forecasting Realized Volatility," *Econometrica*, 71, 579–625.

Barndorff-Nielsen, O. E. and N. Shephard (2002): "Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64, 253–280.

Bianchi, F., H. Mumtaz, and P. Surico (2009): "The Great Moderation of the Term Structure of UK Interest Rates," *Journal of Monetary Economics*, 56, 856 – 871.

Carriero, A., T. E. Clark, and M. Marcellino (2013): "No Arbitrage Priors, Drifting Volatilities, and the Term Structure of Interest Rates," Tech. rep., Working paper.

——— (2015): "Bayesian VARs: Specification Choices and Forecast Accuracy," *Journal of Applied Econometrics*, 30, 46–73.

Carter, C. K. and R. Kohn (1994): "On Gibbs Sampling for State Space Models," *Biometrika*, 81, 541–553.

Christensen, J. H., F. X. Diebold, and G. D. Rudebusch (2011): "The affine arbitrage-free class of Nelson-Siegel term structure models," *Journal of Econometrics*, 164, 4–20.

Christensen, J. H., J. A. Lopez, and G. D. Rudebusch (2014): "Can Spanned Term Structure Factors Drive Stochastic Volatility?" *Federal Reserve Bank of San Francisco Working Paper Series*.

Cieslak, A. and P. Povala (2015): "Information in the Term Structure of Yield Curve Volatility," *Journal of Finance*, forthcoming.

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

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

Cogley, T. and T. J. Sargent (2005): "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII US," *Review of Economic Dynamics*, 8, 262 – 302.

Collin-Dufresne, P., R. Goldstein, and C. Jones (2009): "Can Interest Rate Volatility Be Extracted from the Cross Section of Bond Yields?" *Journal of Financial Economics*, 94, 47–66.

Creal, D. and J. Wu (2014): "Interest Rate Uncertainty and Economic Fluctuations," *Working Paper, The University of Chicago Booth School of Business*.

Del Negro, M. and F. Schorfheide (2013): "DSGE Model-Based Forecasting," in *Handbook of Economic Forecasting*, ed. by G. Elliott and A. Timmermann, Elsevier, vol. 2A, chap. 2, 57 – 140.

Diebold, F. X. and C. Li (2006): "Forecasting the Term Structure of Government Bond Yields," *Journal of Econometrics*, 130, 337 – 364.

Diebold, F. X., C. Li, and V. Z. Yue (2008): "Global yield curve dynamics and interactions: A dynamic Nelson-Siegel approach," *Journal of Econometrics*, 146, 351–363.

Diebold, F. X. and R. S. Mariano (1995): "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics*, 20, 134–144.

Diebold, F. X. and G. D. Rudebusch (2012): *Yield Curve Modeling and Forecasting*, Princeton University Press.

Duffee, G. R. (2012): "Forecasting Interest Rates," *Working papers, The Johns Hopkins University, Department of Economics*.

Duffie, D. and R. Kan (1996): "A Yield-Factor Model Of Interest Rates," *Mathematical Finance*, 6, 379–406.

Egorov, A. V., Y. Hong, and H. Li (2006): "Validating Forecasts of the Joint Probability Density of Bond Yields: Can Affine Models Beat Random Walk?" *Journal of Econometrics*, 135, 255 – 284.

Engle, R. F., V. K. Ng, and M. Rothschild (1990): "Asset Pricing with a Factor-ARCH Covariance Structure: Empirical Estimates for Treasury Bills," *Journal of Econometrics*, 45, 213–237.

Exterkate, P., D. V. Dijk, C. Heij, and P. J. F. Groenen (2013): "Forecasting the Yield Curve in a Data?Rich Environment Using the Factor?Augmented NelsonSiegel Model," *Journal of Forecasting*, 32, 193–214.

Fama, E. F. and R. R. Bliss (1987): "The Information in Long-Maturity Forward Rates," *The American Economic Review*, 77, pp. 680–692.

Geweke, J. and G. Amisano (2010): "Comparing and Evaluating Bayesian Predictive Distributions of Asset Returns," *International Journal of Forecasting*, 26, 216 – 230.

Gürkaynak, R. S., B. Sack, and J. H. Wright (2007): "The U.S. Treasury Yield Curve: 1961 to the Present," *Journal of Monetary Economics*, 54, 2291 – 2304.

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, 877–906.

Hansen, P. R., A. Lunde, and J. M. Nason (2011): "The Model Confidence Set," *Econometrica*, 79, 453–497.

Harvey, D., S. Leybourne, and P. Newbold (1997): "Testing the Equality of Prediction Mean Squared Errors," *International Journal of Forecasting*, 13, 281 – 291.

Hautsch, N. and Y. Ou (2012): "Analyzing Interest Rate Risk: Stochastic Volatility in the Term Structure of Government Bond Yields," *Journal of Banking & Finance*, 36, 2988 – 3007.

Hautsch, N. and F. Yang (2012): "Bayesian Inference in a Stochastic Volatility Nelson–Siegel Model," *Computational Statistics & Data Analysis*, 56, 3774 – 3792.

Jin, X. and J. M. Maheu (2013): "Modeling Realized Covariances and Returns," *Journal of Financial Econometrics*, 11, 335–369.

Jungbacker, B., S. J. Koopman, and M. van der Wel (2013): "Smooth Dynamic Factor Analysis with Application to the US Term Structure of Interest Rates," *Journal of Applied Econometrics*.

Justiniano, A. and G. E. Primiceri (2008): "The Time-Varying Volatility of Macroeconomic Fluctuations," *The American Economic Review*, 98, pp. 604–641.

Kang, K. H. (2015): "The predictive density simulation of the yield curve with a zero lower bound," *Journal of Empirical Finance*, 33, 51 – 66.

Kim, S., N. Shephard, and S. Chib (1998): "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *The Review of Economic Studies*, 65, 361–393.

Koopman, S. J., M. I. P. Mallee, and M. Van der Wel (2010): "Analyzing the Term Structure of Interest Rates Using the Dynamic Nelson–Siegel Model With Time–Varying Parameters," *Journal of Business & Economic Statistics*, 28, 329–343.

Maheu, J. M. and T. H. McCurdy (2011): "Do High-Frequency Measures of Volatility Improve Forecasts of Return Distributions?" *Journal of Econometrics*, 160, 69 – 76.

Pooter, M. D. (2007): "Examining the Nelson-Siegel Class of Term Structure Models," Tinbergen Institute Discussion Papers 07-043/4, Tinbergen Institute.

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

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

Takahashi, M., Y. Omori, and T. Watanabe (2009): "Estimating Stochastic Volatility Models Using Daily Returns and Realized Volatility Simultaneously," *Computational Statistics & Data Analysis*, 53, 2404 – 2426.

van Dijk, D., S. J. Koopman, M. van der Wel, and J. H. Wright (2014): "Forecasting Interest Rates with Shifting Endpoints," *Journal of Applied Econometrics*, 29, 693–712.

Wright, J. and H. Zhou (2009): "Bond Risk Premia and Realized Jump Risk," *Journal of Banking & Finance*, 33, 2333–2345.

Appendices

A State Space Representation

For completeness, we present the full specification of the state space form of the model. We give a detailed explanation of these equations in sections 2.1 and 2.2 of the main text. Consider a set of bond yields $y_t = \{y_t(1), ..., y_t(N)\}'$. τ_j is the maturity in months of bond yield j and λ is the point of maximal curvature.

$$y_t = \begin{pmatrix} 1 & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} - e^{\lambda\tau_1} \\ . & . & . \\ . & . & . \\ 1 & \frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} & \frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} - e^{\lambda\tau_N} \end{pmatrix} \begin{pmatrix} f_{l,t} \\ f_{s,t} \\ f_{c,t} \end{pmatrix} + \epsilon_t, \quad \epsilon_t \sim N(0, Q) \tag{A.1}$$

$$\log(RV_t) = \beta + \Lambda_h \widetilde{h}_t + \zeta_t, \quad \zeta_t \sim N(0, S) \tag{A.2}$$

$$f_t = (I_3 - \Phi_f)\mu_f + \Phi_f f_{t-1} + \eta_t, \quad \eta_t \sim N\left(0,\ diag\left(e^{[h_{l,t}, h_{s,t}, h_{c,t}]}\right)\right) \tag{A.3}$$

$$h_{i,t} = \mu_{i,h}(1 - \phi_{i,h}) + \phi_{i,h} h_{i,t-1} + \nu_{i,t}, \quad \nu_{i,t} \sim N(0,\ \sigma^2_{i,h}) \tag{A.4}$$

for $i = l, s, c$ and Q and S are diagonal matrix. We write $\widetilde{h}_t$ as a vector of the demeaned volatility processs with individual elements $\widetilde{h}_{i,t} = h_{i,t} - \mu_{i,h}$.

B Measurement Equation for RV: Derivation and Approximation

Equation A.2 is the linearized version of the nonlinear measurement equation that comes from adding realized volatility information to the dynamic factor model. We perform a

first-order approximation of the logarithm of the following equation

$$\begin{aligned} RV_t \approx Var_{t-1}(y_t) &= diag(\Lambda_f H_t \Lambda_f' + Q) \\ &= diag(\widetilde{\Lambda}_f \widetilde{H}_t \widetilde{\Lambda}_f' + Q) \end{aligned} \tag{A.5}$$

where we write the logarithm of volatility in deviation form $\widetilde{h}_{i,t} = h_{i,t} - \mu_{i,h}$ for $i = l, s, c$. Then $\widetilde{H}_t$ is a 3×3 diagonal matrix with each element corresponding to $e^{\widetilde{h}_{i,t}}$ and $\widetilde{\Lambda}_f = \Lambda_f[e^{\mu_l/2}, e^{\mu_s/2}, e^{\mu_c/2}]'$. We first derive the nonlinear measurement equation that links the realized volatility with underlying factor volatility. Our derivation is similar to Maheu and McCurdy (2011) but we derive it under the dynamic factor model framework. Then, we describe the approximation to get the linearized measurement equation for RV_t[21].

Derivation of the measurement equation. We assume that RV_t is a noisy measure for the true conditional volatility. Following Maheu and McCurdy (2011), we assume that RV_t follows a log-normal distribution conditional on past histories of bond yields and bond yield realized volatilities:

$$RV_t = diag(\widetilde{\Lambda}_f \widetilde{H}_t \widetilde{\Lambda}_f' + Q) Z_t,$$

where Z_t is a 17×1 vector and each element in Z_t follows a log-normal distribution with $E_{t-1}[Z_{i,t}] = 1$ for $i = 1, ..., 17$. Taking logarithm on both sides gives

$$\log(RV_t) = \widetilde{\beta} + \log(diag(\widetilde{\Lambda}_f \widetilde{H}_t \widetilde{\Lambda}_f' + Q)) + \widetilde{S}\zeta_t, \quad \zeta_t \sim \mathcal{N}(0, I_{17}). \tag{A.6}$$

where $\widetilde{\beta}$ is a 17×1 vector, $\widetilde{S}$ is a 17×17 diagonal matrix, and I_{17} is a 17×17 identity matrix. We separately estimate $\widetilde{\beta}$ and $\widetilde{S}$ where $\widetilde{\beta}$ is a conditional variance of Z_t plus a bias correction term (e.g. Takahashi et al., 2009, for the univariate case).

[21] The derivation and approximation of the measurement equations when we add stochastic volatility to the measurement errors Q_t follows the same basic procedure, so we suppress the details.

Linearization. We present the derivation of equation 5. We linearize equation A.6 for the ith element around $\widetilde{h}_{j,t} = 0$ for $j = 1, 2, 3$ and $\widetilde{S}_{i,i} = 0$ for $i = 1, ..., 17$ to get

$$\log(RV_{i,t}) = \beta_i + \nu_i \left(\sum_{j=1}^{3} \widetilde{\Lambda}^2_{f,i,j} \widetilde{h}_{j,t} \right) + \widetilde{S}\zeta_t, \quad \zeta_t \sim \mathcal{N}(0, I_{17}).$$

where $\widetilde{\Lambda}^2_{f,i,j}$ is the square of the (i,j)th element of $\widetilde{\Lambda}_f$ in equation 5 and

$$\beta_i = \widetilde{\beta}_i + \log \left(\sum_{j=1}^{3} \widetilde{\Lambda}^2_{f,i,j} + Q_{i,i} \right)$$
$$\nu_i = \frac{1}{\left(\sum_{j=1}^{3} \widetilde{\Lambda}^2_{f,i,j} + Q_{i,i} \right)}.$$

Note that the term $\nu_i \left(\sum_{j=1}^{3} \widetilde{\Lambda}^2_{f,i,j} \widetilde{h}_{j,t} \right)$ is linear in $\widetilde{h}_t$ and we can write it as $\Lambda_{i,h} \widetilde{h}_t$. In addition, we write $S = \widetilde{S}\widetilde{S}'$. This completes the derivation of the RV measurement equation (equation A.2).

C Estimation Procedure

Presented is the algorithm for the posterior sampler. We draw $15,000$ samples, saving every 5th draw, with the first $5,000$ draws as burn-in. The priors we choose for the model are presented in Table A-1. We describe the algorithm for our main model, DNS-RV. Estimation of other model specifications is a straightforward simplification or generalization. We collect the parameters on which we would like to perform inference in one vector and write it as $\Theta^* = \{Q, \beta, S, f_{1:T}, \mu_f, \phi_f, h_{1:T}, \mu_h, \phi_h, \sigma^2_h\}$[22]. In addition, we denote Θ^*_{-x} as the parameter vector of all elements in Θ^* except x and *data* as U.S. bond yields and their corresponding yield realized volatilities used in the estimation. The posterior sampler iterates the following

[22] While some authors have estimated the λ, we fix it at 0.0609, noting from Diebold and Li (2006) and others that the value does not move around too much across time and that its estimation does not seem to affect the results.

Table A-1 Prior Distribution

Parameter	Description	Dim.	Dist.	Para(1)	Para(2)
H	Variance of the measurement error (y_t).	17×1	IG	0	0.001
μ_f	Long-run mean parameter for f_t.	3×1	N	0	100
ϕ_f	AR(1) coefficient for f_t.	3×1	N	0.8	100
μ_h	Long-run mean parameter for h_t.	3×1	N	0	100
ϕ_h	AR(1) coefficient for h_t.	3×1	N	0.8	100
σ_h^2	Variance of the innovation for the h_t.	3×1	IG	0.01	2
β	Intercepts in the RV measurement equation. Only used for $\mathcal{M}_{RV}$	17×1	N	0	100
S	Variance of the measurement error RV. Only used for $\mathcal{M}_{RV}$	17×1	IG	0	0.001
σ_f^2	Variance of the innovation for the f_t. Only used for models without time-varying volatility	3×1	IG	0.1	2

Note: a) All prior distributions are independent. For example, prior distributions for elements in H are independent from each other and follow the inverse gamma distribution.
b) Dim: Dimension of the parameters.
c) IG: Inverse gamma distribution. Para(1) and Para(2) mean scale and shape parameters, respectively.
d) N: Normal distribution. Para(1) and Para(2) stand for mean and variance, respectively.
e) Priors for ϕ_f and ϕ_h are truncated so that the processes for factors and volatilities are stationary.
f) $\mathcal{M}_{RV}$ is the set of models with realized volatility data.
g) For the DNS-RV-FULL model We use a normal prior centered around 0 with variance 100 for each element of $\mu_{i,h}$. For Φ_h and Σ_h, we use a natural conjugate prior of the form $\Sigma_h^{-1} \sim W(3, diag([0.01, 0.01, 0.01]))$ and $vec(\Phi_h) \sim N(vec(diag([0.8, 0.8, 0.8])), \Sigma_h \otimes diag([100, 100, 100]))$.

steps starting with an initial value Θ^0 and $s = 1$:

1. **(Drawing $Q|data, \Theta_{-Q}^{s-1}$):** Since Q is diagonal, we draw the diagonal elements one at a time. Note that without the RV measurement equation, each element of the diagonal term on Q is distributed as an inverse gamma distribution. For DNS-RV, Q enters in the realized measurement equation (equation A.6). In this case, we draw Q using the Metropolis-Hastings algorithm with a proposal distribution as an inverse gamma distribution with the same moments as in the conditional posterior distribution of Q without the RV measurement equation. Set $\Theta^{s-1} = \{Q^*, \beta^{s-1}, S^{s-1}, f_{1:T}^{s-1}, \mu_f^{s-1}, \phi_f^{s-1}, h_{1:T}^{s-1}, \mu_h^{s-1}, \phi_h^{s-1}, (\sigma_h^2)^{s-1}\}$ where Q^* is a new draw from this Metropolis-Hastings

sampler.

2. **(Drawing $\beta, S|data, \Theta^{s-1}_{-\beta,S}$):** We can likewise draw β and the diagonal elements of S equation-by-equation. It is a standard linear regression normal-inverse gamma framework. Set $\Theta^{s-1} = \{Q^*, \beta^*, S^*, f^{s-1}_{1:T}, \mu^{s-1}_f, \phi^{s-1}_f, h^{s-1}_{1:T}, \mu^{s-1}_h, \phi^{s-1}_h, (\sigma^2_h)^{s-1}\}$ where β^* and S^* is a new draw from the conditional posterior distribution of β and S.

3. **(Drawing $f_{1:T}|data, \Theta^{s-1}_{-f_{1:T}}$):** The Carter and Kohn (1994) multi-move Gibbs sampling procedure with stochastic volatility can be used to draw the level, slope, and curvature factors. Set $\Theta^{s-1} = \{Q^*, \beta^*, S^*, f^*_{1:T}, \mu^{s-1}_f, \phi^{s-1}_f, h^{s-1}_{1:T}, \mu^{s-1}_h, \phi^{s-1}_h, (\sigma^2_h)^{s-1}\}$ where $f^*_{1:T}$ is a new draw from the multi-move Gibbs sampler.

4. **(Drawing $\mu_f, \phi_f|data, \Theta^{s-1}_{-\mu_f,\phi_f}$):** Because we specify the factors and stochastic volatilities to have independent AR(1) processes, we can separate the drawing of the parameters for each factor. Drawing the parameters equation-by-equation is possible through the linear regression framework allowing for stochastic volatility. We generate draws from the conditional distribution using the Gibbs sampler laid out in Bianchi et al. (2009) and Hautsch and Yang (2012). Then, set $\Theta^{s-1} = \{Q^*, \beta^*, S^*, f^*_{1:T}, \mu^*_f, \phi^*_f, h^{s-1}_{1:T}, \mu^{s-1}_h, \phi^{s-1}_h, (\sigma^2_h)^{s-1}\}$.

5. **(Drawing $h_{1:T}|data, \Theta^{s-1}_{-h_{1:T}}$):** We have a measurement equation made up of two parts. The first part uses the Kim et al. (1998) method to transform the level, slope, and curvature factor equations (equation A.3). These measurement equations are

$$\log\left((f_{i,t} - (1-\phi_{f,i})\mu_{f,i} - \phi_{f,i}f_{i,t-1})^2\right) = h_{i,t} + \log(\eta^2_{i,t}) \tag{A.7}$$

for $i = l, s, c$. The second part is the realized volatility measurement equation (equation A.2). Because of our linear approximation of the nonlinear RV measurement equation, we can simply use the standard Carter and Kohn (1994) multi-move Gibbs sampler in conjunction with the transition equation for the volatility processes (equation A.4). This step can be viewed as an extension of the Kim et al. (1998)'s sampler with an

extra linear measurement equations. We set $\Theta^{s-1} = \{Q^*, \beta^*, S^*, f^*_{1:T}, \mu^*_f, \phi^*_f, h^*_{1:T}, \mu^{s-1}_h, \phi^{s-1}_h, (\sigma^2_h)^{s-1}\}$ where $h^*_{1:T}$ is a new draw from the sampler.

6. **(Drawing $\mu_h, \phi_h, \sigma^2_h | data, \Theta^{s-1}_{-\mu_h,\phi_h,\sigma^2_h}$):** Conditional on the volatility process series, $h^*_{1:T}$, we can use a standard linear regression normal-inverse gamma framework to draw $\mu_{i,h}$, $\phi_{i,h}$, and $\sigma^2_{i,h}$ equation-by-equation for $i = l, s, c$. Set

$$\Theta^s = \{Q^*, \beta^*, S^*, f^*_{1:T}, \mu^*_f, \phi^*_f, h^*_{1:T}, \mu^*_h, \phi^*_h, (\sigma^2_h)^*\}$$

where $\mu^*_{i,h}$, $\phi^*_{i,h}$, and $(\sigma^2_{i,h})^*$ are newly drawn parameters from the conditional posterior distribution.

7. Go to the step 1 with $s \leftarrow s + 1$.

D Forecasting Procedure

Presented in equations A.8 - A.10 is the forecasting algorithm that we use. Because we are performing Bayesian analysis, we explicitly take into account the parameter uncertainty when generating our forecasts. We first draw parameters from the relevant posterior distributions (j) and then simulate 10 trajectories of data given the parameter values (k). We do so for $2,000$ parameter draws for a total of $20,000$ simulated data chains from which to compare to the realized data (Del Negro and Schorfheide, 2013). Note that for the DNS-C model, we would not have equation A.10 and the H_t would become H.

$$\hat{y}^{j,k}_t = \Lambda_f f^{j,k}_t + \tilde{\epsilon}^{j,k}_t, \quad \tilde{\epsilon}^{j,k}_t \sim N\left(0, Q^j\right) \tag{A.8}$$

$$f^{j,k}_t = (I_3 - \Phi^j_f)\mu_f + \Phi^j_f f^{j,k}_{t-1} + \tilde{\eta}^{j,k}_t, \quad \tilde{\eta}^{j,k}_t \sim N\left(0,\ diag\left(e^{[h^{j,k}_{l,t}, h^{j,k}_{s,t}, h^{j,k}_{c,t}]}\right)\right) \tag{A.9}$$

$$h^{j,k}_{i,t} - \mu^j_{i,h} = \phi^j_{i,h}(h^{j,k}_{i,t-1} - \mu^j_{i,h}) + \tilde{e}^{j,k}_{i,t}, \quad \tilde{e}^{j,k}_{i,t} \sim N\left(0, \left(\sigma^j_{i,h}\right)^2\right) \tag{A.10}$$

for $j = 1, ..., 2,000$, $k = 1, ..., 10$, and $t = T, ..., T + 12$ where T is the beginning of the forecasting period.

E Supporting Material

This section contains additional tables and figures that support our claims in the main text. First, we present descriptive statistics of the data that we used in the estimation and forecasting exercises (section E.1). Then, we report parameter estimates (section E.2) and estimated bond yields factors f_t (section E.3). In section E.4, we report the relative importance (ratio in %) of variation between the measurement error and f_t. Finally, as a robustness check for the density forecasting evaluation conducted in section 5.3, we compute and present the model confidence set of Hansen et al. (2011) in section E.5. The model confidence set results list a subset of forecasting models that includes the best models in terms of the log predictive score at the 5% confidence level.

E.1 Descriptive statistics of data

Table A-2 Descriptive Statistics (Yields)

Maturity	mean	std	min	max	$\hat{\rho}(1)$	$\hat{\rho}(12)$	$\hat{\rho}(24)$
3	5.35	3.14	0.04	16.02	0.97	0.65	0.39
6	5.52	3.17	0.15	16.48	0.98	0.66	0.40
9	5.64	3.19	0.19	16.39	0.98	0.67	0.42
12	5.75	3.19	0.25	16.10	0.98	0.69	0.44
15	5.87	3.21	0.38	16.06	0.98	0.70	0.46
18	5.95	3.20	0.44	16.22	0.98	0.71	0.47
21	6.03	3.19	0.53	16.17	0.98	0.71	0.49
24	6.06	3.15	0.53	15.81	0.98	0.72	0.50
30	6.18	3.11	0.82	15.43	0.98	0.73	0.52
36	6.29	3.08	0.98	15.54	0.98	0.74	0.54
48	6.48	3.02	1.02	15.60	0.98	0.75	0.57
60	6.60	2.94	1.56	15.13	0.98	0.76	0.59
72	6.73	2.92	1.53	15.11	0.98	0.77	0.61
84	6.81	2.84	2.18	15.02	0.98	0.77	0.61
96	6.90	2.81	2.11	15.05	0.98	0.78	0.63
108	6.95	2.79	2.15	15.11	0.98	0.78	0.63
120	6.95	2.72	2.68	15.19	0.98	0.77	0.63

Notes: For each maturity we present mean, standard deviation, minimum, maximum and the j-th order autocorrelation coefficients for j = 1, 12, and 24.

Table A-3 Descriptive Statistics (log realized volatility)

Maturity	mean	std	min	max	$\hat{\rho}(1)$	$\hat{\rho}(12)$	$\hat{\rho}(24)$
3	-3.05	1.29	-5.33	1.10	0.65	0.31	0.19
6	-3.35	1.21	-6.66	0.64	0.69	0.28	0.18
9	-3.24	1.10	-5.90	0.47	0.66	0.23	0.10
12	-3.11	1.02	-5.76	0.41	0.63	0.21	0.04
15	-2.99	0.96	-5.68	0.36	0.61	0.20	0.01
18	-2.89	0.93	-5.59	0.29	0.61	0.20	-0.01
21	-2.81	0.90	-5.52	0.20	0.60	0.20	-0.02
24	-2.75	0.88	-5.46	0.11	0.60	0.20	-0.03
30	-2.66	0.85	-5.34	-0.04	0.59	0.20	-0.03
36	-2.61	0.83	-5.23	-0.12	0.60	0.21	-0.03
48	-2.57	0.79	-5.02	-0.19	0.61	0.22	-0.02
60	-2.57	0.77	-4.84	-0.22	0.61	0.23	-0.01
72	-2.58	0.76	-4.67	-0.25	0.61	0.25	0.00
84	-2.59	0.76	-4.66	-0.21	0.62	0.27	0.00
96	-2.61	0.75	-4.66	-0.18	0.63	0.29	0.01
108	-2.62	0.75	-4.66	-0.17	0.64	0.30	0.01
120	-2.63	0.75	-4.65	-0.18	0.66	0.31	0.01

Notes: For each maturity we present mean, standard deviation, minimum, maximum and the j-th order autocorrelation coefficients for j = 1, 12, and 24.

E.2 In-sample estimation (posterior moments)

We denote $\{y_1, y_2, y_3, ..., y_{16}, y_{17}\}$ = monthly U.S. Treasury yields with maturities of (3, 6, 9, 12, 15, 18, 21 months and 2, 2.5, 3, 4, 5, 6, 7, 8, 9, 10 years.

Table A-4 Posterior moments of H

		RW-C	RW-SV	RW-RV	DNS-C	DNS-SV	DNS-RV	RW-SV-RW	DNS-SV-RW
y_1	5%	0.064	0.063	0.060	0.063	0.063	0.488	0.063	0.063
	50%	0.074	0.072	0.068	0.072	0.072	0.610	0.072	0.072
	95%	0.085	0.083	0.080	0.083	0.082	0.760	0.083	0.083
y_2	5%	0.008	0.008	0.010	0.008	0.008	0.157	0.008	0.008
	50%	0.010	0.009	0.012	0.009	0.009	0.202	0.009	0.009
	95%	0.012	0.011	0.015	0.011	0.011	0.284	0.011	0.011
y_3	5%	0.002	0.002	0.004	0.002	0.002	0.043	0.002	0.002
	50%	0.003	0.003	0.005	0.002	0.002	0.067	0.002	0.003
	95%	0.004	0.003	0.006	0.003	0.003	0.138	0.003	0.003
y_4	5%	0.003	0.003	0.004	0.003	0.003	0.018	0.003	0.003
	50%	0.004	0.004	0.005	0.004	0.004	0.038	0.004	0.004
	95%	0.005	0.005	0.005	0.005	0.005	0.097	0.005	0.005
y_5	5%	0.006	0.006	0.006	0.006	0.007	0.017	0.006	0.006
	50%	0.007	0.007	0.007	0.007	0.007	0.035	0.007	0.007
	95%	0.008	0.009	0.008	0.009	0.009	0.086	0.009	0.009
y_6	5%	0.006	0.006	0.005	0.006	0.006	0.016	0.006	0.006
	50%	0.006	0.006	0.006	0.006	0.007	0.034	0.007	0.007
	95%	0.007	0.008	0.007	0.007	0.008	0.083	0.008	0.008
y_7	5%	0.004	0.004	0.004	0.004	0.004	0.015	0.004	0.004
	50%	0.005	0.005	0.005	0.005	0.005	0.033	0.005	0.005
	95%	0.006	0.006	0.005	0.006	0.006	0.080	0.006	0.006
y_8	5%	0.002	0.002	0.002	0.002	0.002	0.014	0.002	0.002
	50%	0.002	0.002	0.003	0.002	0.002	0.033	0.002	0.002
	95%	0.003	0.003	0.003	0.003	0.002	0.082	0.002	0.003
y_9	5%	0.002	0.002	0.002	0.002	0.002	0.013	0.002	0.002
	50%	0.002	0.002	0.003	0.002	0.002	0.029	0.002	0.002
	95%	0.003	0.003	0.003	0.003	0.003	0.075	0.003	0.003
y_{10}	5%	0.002	0.002	0.003	0.002	0.002	0.011	0.002	0.002
	50%	0.003	0.003	0.003	0.003	0.003	0.026	0.003	0.003
	95%	0.003	0.003	0.003	0.003	0.003	0.068	0.003	0.003
y_{11}	5%	0.004	0.004	0.004	0.004	0.004	0.009	0.004	0.004
	50%	0.004	0.004	0.004	0.004	0.004	0.021	0.004	0.004
	95%	0.005	0.005	0.005	0.005	0.005	0.059	0.005	0.005
y_{12}	5%	0.003	0.003	0.004	0.003	0.003	0.007	0.003	0.003
	50%	0.004	0.004	0.004	0.004	0.004	0.017	0.004	0.004
	95%	0.004	0.004	0.005	0.004	0.004	0.049	0.004	0.004
y_{13}	5%	0.004	0.004	0.005	0.004	0.004	0.006	0.004	0.004
	50%	0.005	0.005	0.006	0.005	0.005	0.013	0.005	0.005
	95%	0.005	0.005	0.006	0.005	0.005	0.038	0.005	0.005
y_{14}	5%	0.003	0.003	0.004	0.003	0.003	0.004	0.003	0.003
	50%	0.004	0.004	0.005	0.004	0.004	0.011	0.004	0.004
	95%	0.005	0.004	0.006	0.004	0.004	0.032	0.004	0.004
y_{15}	5%	0.003	0.002	0.004	0.003	0.002	0.004	0.002	0.002
	50%	0.003	0.003	0.005	0.003	0.003	0.011	0.003	0.003
	95%	0.004	0.004	0.006	0.004	0.004	0.031	0.004	0.004
y_{16}	5%	0.006	0.007	0.007	0.006	0.007	0.006	0.007	0.007
	50%	0.007	0.008	0.008	0.008	0.008	0.012	0.008	0.008
	95%	0.009	0.009	0.010	0.009	0.009	0.028	0.009	0.009
y_{17}	5%	0.012	0.012	0.012	0.012	0.013	0.016	0.013	0.012
	50%	0.014	0.014	0.014	0.014	0.014	0.029	0.014	0.014
	95%	0.016	0.016	0.016	0.016	0.017	0.053	0.017	0.017

Notes: Variance of the measurement error. Not applicable to DNS-ME-SV and DNS-ME-RV. Estimation sample is from January 1981 to November 2009.

Table A-5 Posterior moments of parameters related to f_t

		RW-C	RW-SV	RW-RV	DNS-C	DNS-SV	DNS-RV	RW-SV-RW	DNS-SV-RW	DNS-ME-SV	DNS-ME-RV
$\mu_{f,l}$	5%	0.00	0.00	0.00	-7.28	-5.83	-17.48	0.00	-5.44	-7.20	-7.40
	50%	0.00	0.00	0.00	4.63	4.57	-5.47	0.00	4.71	4.41	4.37
	95%	0.00	0.00	0.00	7.86	7.17	2.15	0.00	6.94	7.83	7.79
$\mu_{f,s}$	5%	0.00	0.00	0.00	-4.28	-4.04	-1.14	0.00	-4.28	-4.47	-4.49
	50%	0.00	0.00	0.00	-2.54	-0.88	2.92	0.00	-0.98	-2.55	-2.59
	95%	0.00	0.00	0.00	-1.43	6.93	13.81	0.00	6.46	-1.50	-1.43
$\mu_{f,c}$	5%	0.00	0.00	0.00	-2.64	-2.13	-6.31	0.00	-1.78	-2.97	-2.91
	50%	0.00	0.00	0.00	-0.82	-0.52	0.16	0.00	-0.42	-0.87	-0.79
	95%	0.00	0.00	0.00	0.92	0.84	8.66	0.00	0.70	1.26	1.33
$\phi_{f,l}$	5%	1.00	1.00	1.00	0.98	0.98	0.99	1.00	0.98	0.98	0.98
	50%	1.00	1.00	1.00	0.99	0.99	1.00	1.00	0.99	0.99	0.99
	95%	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$\phi_{f,s}$	5%	1.00	1.00	1.00	0.94	0.97	0.99	1.00	0.97	0.94	0.94
	50%	1.00	1.00	1.00	0.96	0.99	1.00	1.00	0.99	0.96	0.97
	95%	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00	0.99	0.99
$\phi_{f,c}$	5%	1.00	1.00	1.00	0.92	0.91	0.97	1.00	0.90	0.92	0.93
	50%	1.00	1.00	1.00	0.95	0.95	0.99	1.00	0.94	0.96	0.96
	95%	1.00	1.00	1.00	0.99	0.99	1.00	1.00	0.98	0.99	0.99
$\sigma_{f,l}$	5%	0.10			0.10	0.98	0.99		0.98	0.09	0.09
	50%	0.12			0.11	0.99	1.00		0.99	0.11	0.10
	95%	0.13			0.13	1.00	1.00		1.00	0.12	0.12
$\sigma_{f,s}$	5%	0.16			0.16					0.16	0.15
	50%	0.19			0.19					0.18	0.17
	95%	0.22			0.21					0.20	0.19
$\sigma_{f,c}$	5%	0.42			0.43					0.40	0.37
	50%	0.49			0.51					0.46	0.44
	95%	0.57			0.59					0.54	0.51

Notes: Estimation sample is from January 1981 to November 2009.

Table A-6 Posterior moments of parameters related to h_t

		RW-C	RW-SV	RW-RV	DNS-C	DNS-SV	DNS-RV	RW-SV-RW	DNS-SV-RW
$\mu_{l,h}$	5%		-4.75	-4.11		-4.40	-4.19	0.00	0.00
	50%		-2.55	-3.85		-2.56	-3.92	0.00	0.00
	95%		-1.11	-1.76		-1.48	-2.21	0.00	0.00
$\mu_{s,h}$	5%		-3.39	-3.22		-3.41	-3.22	0.00	0.00
	50%		-2.30	-2.82		-2.30	-2.85	0.00	0.00
	95%		-1.69	-1.74		-1.66	-1.94	0.00	0.00
$\mu_{c,h}$	5%		-1.47	-2.25		-1.48	-2.18	0.00	0.00
	50%		-0.97	-1.93		-0.96	-1.88	0.00	0.00
	95%		-0.53	-1.45		-0.47	-1.39	0.00	0.00
$\phi_{l,h}$	5%		0.93	0.59		0.93	0.58	1.00	1.00
	50%		0.98	0.66		0.98	0.66	1.00	1.00
	95%		1.00	0.73		1.00	0.73	1.00	1.00
$\phi_{s,h}$	5%		0.92	0.56		0.92	0.57	1.00	1.00
	50%		0.97	0.65		0.96	0.64	1.00	1.00
	95%		0.99	0.72		1.00	0.72	1.00	1.00
$\phi_{c,h}$	5%		0.74	0.40		0.81	0.39	1.00	1.00
	50%		0.91	0.50		0.92	0.49	1.00	1.00
	95%		0.98	0.61		0.98	0.60	1.00	1.00
$\sigma_{l,h}$	5%		0.01	0.39		0.01	0.41	0.01	0.01
	50%		0.03	0.72		0.03	0.75	0.01	0.01
	95%		0.08	0.88		0.08	0.92	0.03	0.03
$\sigma_{s,h}$	5%		0.01	1.29		0.01	1.27	0.01	0.01
	50%		0.03	1.75		0.04	1.72	0.02	0.03
	95%		0.10	3.00		0.10	2.64	0.06	0.05
$\sigma_{c,h}$	5%		0.01	1.41		0.02	1.30	0.01	0.01
	50%		0.12	2.00		0.09	1.78	0.03	0.03
	95%		0.40	4.00		0.27	4.30	0.09	0.09

Notes: Parameters related to h_t. Not applicable to DNS-C, RW-C, DNS-ME-SV and DNS-ME-RV. Estimation sample is from January 1981 to November 2009.

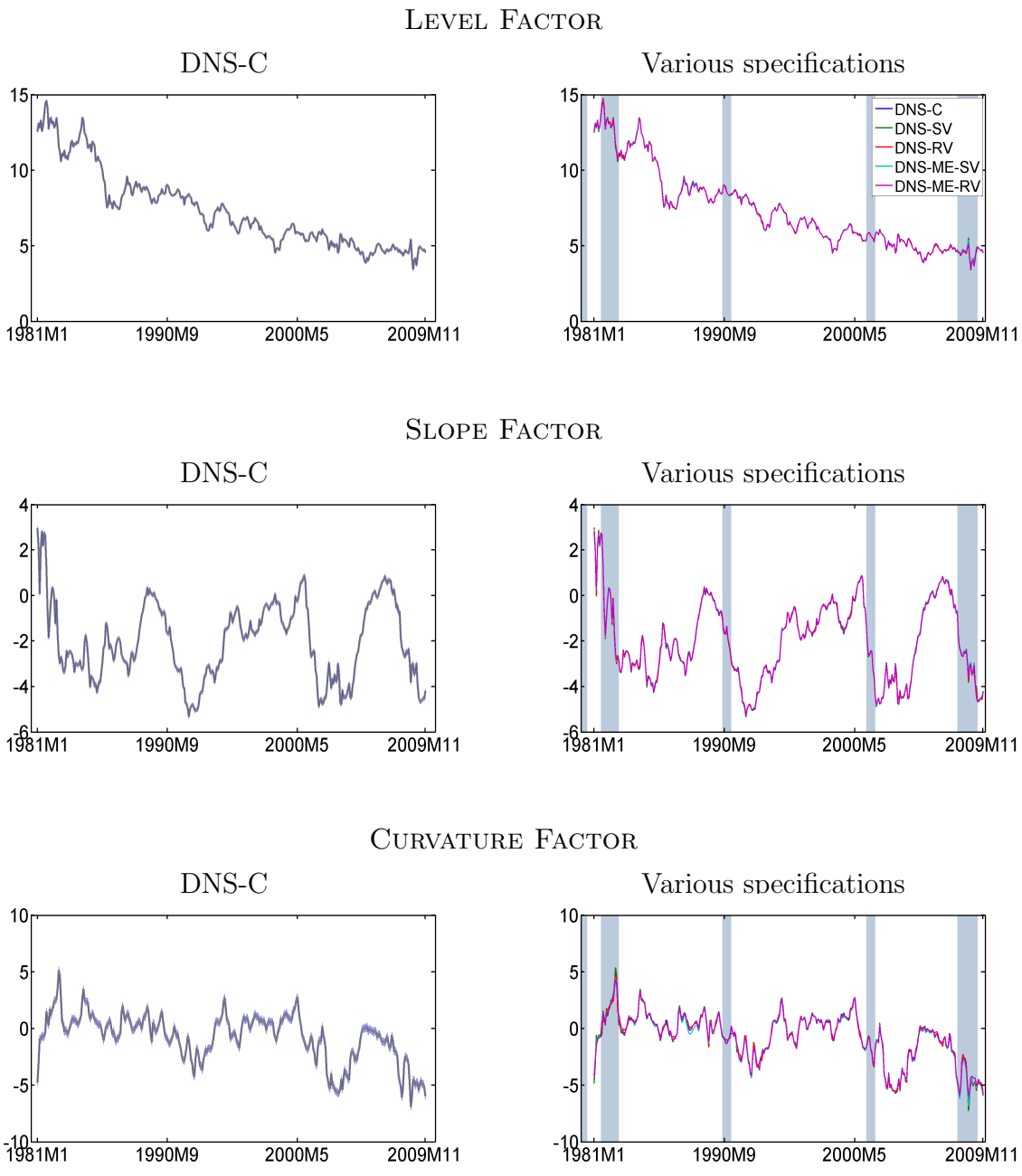
Table A-7 Posterior moments of parameters related to h_t for ME-SV model

	μ_h			ϕ_h			σ_h^2		
		DNS-ME-SV	DNS-ME-RV		DNS-ME-SV	DNS-ME-RV		DNS-ME-SV	DNS-ME-RV
y_1	5%	-4.80	-4.28	5%	0.91	0.91	5%	0.14	0.04
	50%	-3.61	-3.45	50%	0.95	0.95	50%	0.23	0.08
	95%	-2.50	-2.92	95%	0.99	0.98	95%	0.36	0.13
y_2	5%	-5.75	-5.79	5%	0.89	0.91	5%	0.03	0.01
	50%	-4.88	-5.03	50%	0.95	0.95	50%	0.07	0.03
	95%	-4.14	-4.62	95%	0.99	0.99	95%	0.14	0.06
y_3	5%	-6.28	-5.77	5%	0.53	0.88	5%	0.00	0.00
	50%	-5.48	-5.55	50%	0.96	0.93	50%	0.01	0.01
	95%	-4.96	-5.35	95%	1.00	0.98	95%	0.04	0.02
y_4	5%	-6.52	-5.60	5%	0.24	0.87	5%	0.00	0.00
	50%	-5.43	-5.43	50%	0.96	0.92	50%	0.01	0.01
	95%	-4.65	-5.26	95%	1.00	0.97	95%	0.02	0.01
y_5	5%	-7.89	-5.35	5%	0.95	0.86	5%	0.01	0.00
	50%	-5.40	-5.12	50%	0.98	0.92	50%	0.02	0.01
	95%	-4.15	-4.93	95%	1.00	0.97	95%	0.04	0.02
y_6	5%	-10.34	-5.69	5%	0.97	0.88	5%	0.00	0.01
	50%	-5.95	-5.39	50%	0.99	0.93	50%	0.01	0.02
	95%	-5.14	-5.14	95%	1.00	0.97	95%	0.02	0.04
y_7	5%	-9.71	-6.07	5%	0.97	0.89	5%	0.00	0.01
	50%	-6.16	-5.67	50%	0.99	0.94	50%	0.01	0.03
	95%	-5.46	-5.37	95%	1.00	0.98	95%	0.01	0.06
y_8	5%	-8.41	-6.01	5%	0.92	0.84	5%	0.00	0.00
	50%	-5.92	-5.85	50%	0.98	0.91	50%	0.01	0.01
	95%	-5.03	-5.68	95%	1.00	0.96	95%	0.01	0.02
y_9	5%	-7.06	-6.29	5%	0.93	0.87	5%	0.00	0.01
	50%	-6.11	-5.98	50%	0.97	0.93	50%	0.01	0.03
	95%	-5.63	-5.73	95%	1.00	0.97	95%	0.03	0.06
y_{10}	5%	-8.33	-5.89	5%	0.95	0.83	5%	0.00	0.00
	50%	-5.91	-5.71	50%	0.98	0.90	50%	0.01	0.01
	95%	-5.00	-5.56	95%	1.00	0.95	95%	0.01	0.02
y_{11}	5%	-7.80	-5.88	5%	0.95	0.86	5%	0.01	0.02
	50%	-5.53	-5.55	50%	0.98	0.92	50%	0.02	0.05
	95%	-0.82	-5.27	95%	1.00	0.96	95%	0.05	0.09
y_{12}	5%	-8.28	-5.68	5%	0.95	0.84	5%	0.00	0.01
	50%	-5.69	-5.50	50%	0.98	0.91	50%	0.01	0.01
	95%	-4.91	-5.31	95%	1.00	0.96	95%	0.02	0.04
y_{13}	5%	-8.00	-6.12	5%	0.95	0.88	5%	0.01	0.03
	50%	-5.48	-5.60	50%	0.99	0.94	50%	0.03	0.06
	95%	0.85	-5.20	95%	1.00	0.98	95%	0.06	0.10
y_{14}	5%	-9.18	-6.51	5%	0.97	0.92	5%	0.01	0.03
	50%	-5.94	-5.77	50%	0.99	0.96	50%	0.02	0.05
	95%	-2.26	-5.14	95%	1.00	0.99	95%	0.04	0.08
y_{15}	5%	-8.72	-6.07	5%	0.96	0.88	5%	0.00	0.01
	50%	-5.95	-5.69	50%	0.99	0.93	50%	0.01	0.03
	95%	-2.15	-5.36	95%	1.00	0.98	95%	0.03	0.07
y_{16}	5%	-9.02	-6.04	5%	0.96	0.89	5%	0.01	0.03
	50%	-5.84	-5.42	50%	0.99	0.94	50%	0.01	0.06
	95%	-4.57	-4.92	95%	1.00	0.98	95%	0.04	0.11
y_{17}	5%	-6.53	-5.66	5%	0.93	0.91	5%	0.04	0.04
	50%	-5.02	-4.87	50%	0.97	0.95	50%	0.07	0.07
	95%	-3.06	-4.25	95%	1.00	0.99	95%	0.14	0.12

Notes: Parameters related to h_t. For DNS-ME-SV and DNS-ME-RV. Estimation sample is from January 1981 to November 2009.

E.3 Extracted factors (f_t)

Figure A-1 Extracted factors



Notes: Left columns: Factors estimated from the DNS-C model with 80% credible intervals. Right column: Estimated factors from the various specifications. Shaded bars on the right panel are NBER recession dates. 1) Factors are very similar to each other. 2) Factors are very accurately estimated. Estimation sample is from January 1981 to November 2009.

E.4 Relative importance (ratio in %) of variation between the measurement error and f_t.

Table A-8 Relative importance (ratio in %) of variation between the measurement error and f_t.

Maturity	DNS-C	DNS-SV	DNS-RV	DNS-ME-SV	DNS-ME-RV
3	7.72	7.73	7.77	7.95	7.66
6	2.54	2.54	2.80	2.78	2.56
9	1.13	1.12	1.60	1.55	1.51
12	1.80	1.81	1.90	1.73	1.85
15	2.24	2.26	2.22	2.13	2.09
18	2.05	2.06	2.04	2.00	1.94
21	1.77	1.78	1.75	1.75	1.74
24	1.29	1.29	1.35	1.36	1.34
30	1.31	1.28	1.44	1.43	1.41
36	1.43	1.42	1.52	1.39	1.42
48	2.04	2.04	2.11	2.08	2.10
60	1.70	1.71	1.85	1.72	1.69
72	2.18	2.18	2.43	2.32	2.25
84	1.96	1.95	2.24	2.13	2.19
96	1.48	1.45	2.01	1.51	1.64
108	2.74	2.78	2.92	2.57	2.66
120	4.04	4.07	4.11	3.92	4.16

Notes: We calculate $\frac{std(e_t)}{std(f_t)} * 100$ where e_t is measurement error and f_t is a vector level, slope, and curvature factor. This table is to show that variation in the measurement equation is relatively smaller than variation in the factor component. Sizes of variation from the measurement error is about $1\% \sim 8$ % of the variation from the factors. Mostly they are below 3% except 3 month and 10 year bond yields. This evidence supports that time-varying volatility in the transition equation (factor equation) plays much larger role in prediction. Calculated at the posterior median. Estimation sample is from January 1981 to November 2009.

E.5 Model confidence set

Table A-9 Model Confidence Set (5%) Based on the log predictive score

Maturity	List of Models									
1-step-ahead prediction										
3	DNS-RV									
12	RW-RV	DNS-RV								
36	DNS-RV									
60	DNS-SV	DNS-RV	RW-SV							
120	DNS-RV	RW-SV	RW-RV	DNS-SV						
3-step-ahead prediction										
3	RW-RV	DNS-RV								
12	DNS-SV	RW-SV-RW	RW-SV	RW-RV	DNS-RV					
36	RW-C	DNS-SV	DNS-ME-RV	DNS-ME-SV	RW-SV-RW	RW-SV	RW-RV	DNS-RV		
60	DNS-SV-RW	DNS-SV	RW-C	DNS-C	RW-SV-RW	DNS-ME-RV	DNS-ME-SV	RW-RV	RW-SV	DNS-RV
120	RW-RV	DNS-ME-SV	DNS-ME-RV	RW-SV	DNS-RV	DNS-SV	RW-SV-RW	DNS-SV-RW		
6-step-ahead prediction										
3	DNS-SV	DNS-ME-SV	RW-SV	DNS-ME-RV	RW-RV	DNS-RV				
12	DNS-SV-RW	DNS-SV	DNS-C	RW-C	RW-SV-RW	DNS-ME-SV	RW-SV	DNS-ME-RV	RW-RV	DNS-RV
36	DNS-SV-RW	DNS-SV	RW-SV-RW	RW-SV	DNS-C	RW-C	DNS-ME-SV	DNS-ME-RV	RW-RV	DNS-RV
60	DNS-SV-RW	RW-SV-RW	DNS-SV	RW-SV	RW-C	RW-RV	DNS-C	DNS-ME-RV	DNS-ME-SV	DNS-RV
120	RW-SV-RW	DNS-C	RW-SV	DNS-SV-RW	DNS-ME-RV	DNS-ME-SV	DNS-SV	RW-RV	DNS-RV	
12-step-ahead prediction										
3	RW-RV	DNS-RV	DNS-ME-SV	DNS-C	DNS-ME-RV					
12	DNS-SV-RW	DNS-SV	RW-SV-RW	RW-SV	RW-RV	RW-C	DNS-RV	DNS-C	DNS-ME-SV	DNS-ME-RV
36	DNS-SV-RW	DNS-SV	RW-SV-RW	RW-SV	RW-RV	RW-C	DNS-C	DNS-RV	DNS-ME-SV	DNS-ME-RV
60	DNS-ME-SV	RW-RV	DNS-ME-RV	DNS-RV						
120	DNS-ME-RV	DNS-ME-SV	RW-SV	RW-SV-RW	DNS-SV-RW	DNS-SV	RW-RV	DNS-RV		

Notes: This table lists a subset of forecasting models that includes the best models (in terms of the log predictive score) at the 5% confidence level. Specifically, we define the difference in the log predictive score for model i and j as

$$d_{ij,t} = LPS_{i,t} - LPS_{j,t}$$

and define $\mu_{ij} = E[d_{ij,t}]$. Then, the set of best forecasts is defined as,

$$\mathcal{M}^* = \{i \in \mathcal{M} : \mu_{ij} \geq 0, \quad \forall j \in \mathcal{M}\}.$$

We follow Hansen et al. (2011) to construct the model confidence set. We construct p-values using the stationary bootstrap with 10,000 replications and the average window length 12. Computation is based on the MFE Toolbox provided by Kevin Sheppard.