Interest Rate Uncertainty and Economic Fluctuations *

Drew D. Creal
Chicago Booth

Jing Cynthia Wu
Chicago Booth and NBER

First draft: November 1, 2013
This draft: October 2, 2014

Abstract

Uncertainty associated with the monetary policy transmission mechanism is a key driving force of business cycles. To investigate this link, we propose a new term structure model that allows the volatility of the yield curve to interact with macroeconomic indicators. The data favors a model with two volatility factors that capture short-term and long-term interest rate uncertainty. Increases in either of them lead higher unemployment rates, but they interact with inflation in opposite directions.

Keywords: interest rate uncertainty; macroeconomic fluctuations; affine term structure models; stochastic volatility; Bayesian estimation.

*We thank Torben Andersen, Peter Christoffersen, Steve Davis, Marty Eichenbaum, Jim Hamilton, Lars Hansen, Steve Heston, Dale Rosenthal, Dora Xia, Lan Zhang and seminar and conference participants at Chicago Booth, Northwestern, Ohio State, UIC, Conference in Honor of James Hamilton, and Fifth Risk Management Conference. Drew Creal gratefully acknowledges financial support from the William Ladany Faculty Scholar Fund at the University of Chicago Booth School of Business. Cynthia Wu gratefully acknowledges financial support from the IBM Faculty Research Fund at the University of Chicago Booth School of Business. This paper was formerly titled "Term Structure of Interest Rate Volatility and Macroeconomic Uncertainty". Correspondence: dcreal@chicagobooth.edu, cynthia.wu@chicagobooth.edu.

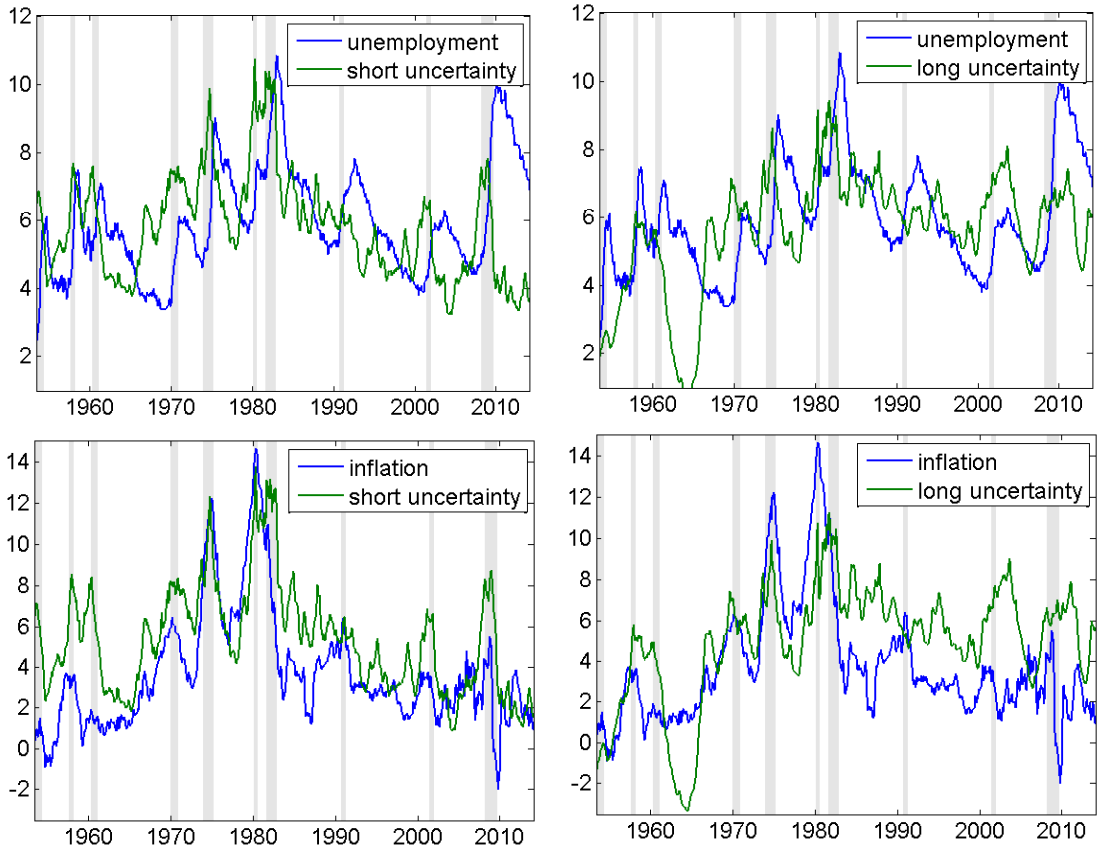
1 Introduction

Does interest rate uncertainty contribute to economic fluctuations and business cycles? If so, by how much? While numerous studies have focused on monetary policy and its transmission mechanism, less attention has been placed on understanding how uncertainty surrounding the monetary policy transmission may impact the economy. We investigate this relationship by introducing a new term structure model: uncertainty is extracted from the volatility (second moments) of yields yet it also impacts the first moments of macroeconomic variables in a vector autoregression (VAR). Combining these two features constitutes the first contribution of our paper, and sets our paper apart from the rest of the literature studying uncertainty in a VAR.

Second, we also contribute to the term structure literature by devising a no-arbitrage model with multiple stochastic volatility factors. The factors driving volatility are distinct from the factors driving yields, which is necessary for capturing both characteristics of the data. The data suggests that two factors are needed to capture the term structure of interest rate volatility. We show that these two dimensions, short-term and long-term interest rate uncertainty, have distinct economic implications too.

Figure 1 plots the short-term (left) and long-term (right) uncertainties in green, and how they relate to the fluctuations of the unemployment rate (top) and inflation rate (bottom) in blue. The top panels illustrate the common business cycle feature of the two uncertainties. Increases in interest rate uncertainty precede economic downturns. Peaks of uncertainty typically lead postwar recessions and lead up-turns in unemployment rates by several months. Despite the common movements across business cycles, these two uncertainties have a distinct dynamic relationship with inflation as seen in the bottom panels of Figure 1. Short-term uncertainty resembles the dynamics of inflation, especially during the high inflation era of the 1970s and early 1980s. This is consistent with the view that monetary policy responds aggressively to battle inflation, for example, Taylor (1993). Greater movements in short-term interest rates that are a response to higher inflation imply higher short-term volatility. The

Figure 1: Uncertainty and macroeconomic variables



Blue: unemployment in the top panels, inflation in the bottom, with the scales on the y-axis. Green: short-term uncertainty in the left panels and long-term uncertainty in the right panels. Estimated uncertainties are scaled to fit the graph. Shaded area: NBER recessions.

relationship between long-term uncertainty and inflation is not uniform. For the first half of the sample, they co-move. In the second half, they move in opposite directions.

To formally capture the dynamic correlations in Figure 1, we document the impulse responses of macroeconomic variables to uncertainty shocks surrounding the yield curve. Consistent with the lead-lag relationship in the top row of Figure 1, the unemployment rate increases in response to a positive shock to either short-term or long-term uncertainty. In contrast, shocks to these uncertainties impact inflation in the opposite directions: inflation decreases in response to a shock to short-term uncertainty and it increases after a positive shock to long-term uncertainty. This is because inflation is neither pro-cyclical nor counter-

cyclical: some recessions are associated with high inflation, e.g. the recessions of the 1970's and early 1980's; others are associated with concerns over deflation, e.g. the Great Recession.

Our two measures of uncertainty have competing effects on inflation. It is not a priori clear about the overall effect. Moreover, although there is no ambiguity about the direction, the magnitude of the effect uncertainty has on unemployment might also differ from a recession to an expansion. Time-varying impulse response functions provide a more comprehensive analysis for both. Different from VARs with homoskedastic or heteroskedastic shocks, impulse responses from our model vary through time in both scale and shape. This is a unique feature of our model due to the fact that both the first and second moments are functions of uncertainty. It allows us to study the nature of uncertainty at different episodes in depth. For example, the Great Recession is associated with downward pressure on inflation and a larger than normal increase in the unemployment rate. This reflects concerns over deflation and a jobless recovery. A similar experience happens during Volcker's tenure, indicating his position as an inflation hawk. In contrast, the Great Inflation is associated with upward pressure on inflation. Movements in the impulse responses for the Great Moderation are less extreme than other time periods and lie in-between the impulse responses from other episodes.

As our model is a non-linear, non-Gaussian state space model whose likelihood is not known in closed-form, we use Bayesian methods to efficiently estimate the parameters and state variables of the model. We develop a Markov chain Monte Carlo (MCMC) algorithm to estimate the parameters and a particle filter for calculating the likelihood and filtered estimates of the state variables. Our MCMC and particle filtering algorithms leverage the fact that the model can be written as a conditionally, linear Gaussian state space model. Consequently, we can use the Kalman filter to draw many of the state variables and parameters efficiently.

The remainder of the paper is organized as follows. We describe our relationship to the relevant literatures in Subsection 1.1. Section 2 presents the new term structure model.

Section 3 describes the MCMC and particle filtering algorithms used for estimation. In Section 4, we study the economic implications of interest rate uncertainty. Section 5 demonstrates how a collection of models with different specifications fit the yield curve. Section 6 concludes.

1.1 Related Literature

Our paper is closely related to recent advances in the literatures on uncertainty and the term structure of interest rates. First, our paper contributes to the fast growing literature on the role that uncertainty shocks play in macroeconomic fluctuations, asset prices and monetary policy; see, e.g. Bloom(2013) for a survey; Baker, Bloom, and Davis(2013), Jurado, Ludvigson, and Ng(2013), Bekaert, Hoerova, and Lo Duca(2013), and Aastveit, Natvik, and Sola(2013) for empirical evidence; and Ulrich(2012), Pastor and Veronesi(2012) and Pastor and Veronesi(2013) for theoretical models.

We differ from the empirical papers in the uncertainty literature in the following ways: (i) We internalize the uncertainty: in our model, uncertainty serves both as the second moment of yields (the factor driving the volatility of interest rates) and it directly impacts the first moment of macroeconomic variables. In contrast, the uncertainty literature usually extracts an estimate of uncertainty in a data pre-processing step, and then uses this estimate in a second step as an observable variable in a homoskedastic vector autoregression. (ii) To our knowledge, we are the first to advocate the two dimensions of uncertainty, and our analysis shows that they impact inflation in opposite directions. (iii) Different from the rest of the literature, we focus on interest rate uncertainty.

Our paper is also related to the VAR literature with stochastic volatility; see Cogley and Sargent(2001), Cogley and Sargent(2005), and Primiceri(2005) for examples. We adopt a similar approach to modeling the time-varying covariance matrix as Primiceri(2005). We differ from this literature in that the volatility factors not only determine the size of the shocks, but they also enter the conditional mean and have a first order impact. The latter is

absent from existing models in this literature, and we show its importance through impulse responses. Another difference between our model and those in the VAR literature is that shocks to volatility are correlated with shocks to macroeconomic variables and the yield factors. For example, shocks to macroeconomic variables can increase volatility.

Finally, we contribute to the term structure literature by introducing a flexible way to simultaneously fit yields and their volatilities at different maturities. Much emphasis of the earlier literature – e.g. Dai and Singleton(2000) and Duffee(2002) – has been placed on fitting the cross section of yields at the expense of fitting the volatilities poorly. To address this limitation, Collin-Dufresne and Goldstein(2002) proposed the class of unspanned stochastic volatility (USV) models which separate the dynamics of volatility from yield factors.[1] Creal and Wu(2014) showed that USV models do improve the fit of volatility, but restrict the cross-sectional fit of yields at the same time. More importantly, the existing literature on USV models typically stops at one volatility factor. In contrast, we show that the data suggests multiple factors.

2 Models

This section proposes a new macro finance term structure model to capture the dynamic relationship between interest rate uncertainty and the macroeconomy. Our model has the following unique features. First, uncertainty – originated from the volatility of the yield curve – has a first order impact on the macroeconomy. Second, our model captures multiple dimensions of yield volatility in a novel way. In our setting, fitting the yield volatility does not constrain bond prices. Besides the flexibility of fitting the volatility, our pricing formula remains simple and straightforward.

[1]There is some prior empirical work that studies whether volatility is priced using interest rate derivatives or high frequency data. Examples include Bikbov and Chernov(2009), Andersen and Benzoni(2010), Joslin(2010), Mueller, Vedolin, and Yen(2011), Cieslak and Povala(2013), and Christensen, Lopez, and Rudebusch(2014).

2.1 Dynamics

The model has an $M \times 1$ vector of macroeconomic variables m_t, and a $G \times 1$ vector of Gaussian yield factors g_t that drive bond prices. The $H \times 1$ vector of factors h_t determine the volatility of yields, and provide us measures of interest rate uncertainty.

The factors jointly follow a vector autoregression with stochastic volatility. Specifically, the macro factors follow

$$m_{t+1} = \mu_m + \Phi_m m_t + \Phi_{mg} g_t + \Phi_{mh} h_t + \Sigma_m \varepsilon_{m,t+1}. \quad (1)$$

The dynamics for the yield factors are

$$g_{t+1} = \mu_g + \Phi_{gm} m_t + \Phi_g g_t + \Phi_{gh} h_t + \Sigma_{gm} \varepsilon_{m,t+1} + \Sigma_g D_t \varepsilon_{g,t+1}, \quad (2)$$

where the diagonal time-varying volatility is a function of h_t

$$D_t = \text{diag}\left(\exp\left(\frac{\Gamma_0 + \Gamma_1 h_t}{2}\right)\right). \quad (3)$$

The volatility factors h_t have dynamics

$$h_{t+1} = \mu_h + \Phi_h h_t + \Sigma_{hm} \varepsilon_{m,t+1} + \Sigma_{hg} D_t \varepsilon_{g,t+1} + \Sigma_h \varepsilon_{h,t+1}. \quad (4)$$

To ensure stability of the system, the conditional mean of h_{t+1} does not depend on g_t and m_t. The shocks are jointly i.i.d. normal $(\varepsilon'_{m,t+1}, \varepsilon'_{g,t+1}, \varepsilon'_{h,t+1})' \sim N(0, I)$, with the contemporaneous correlations captured through the matrices Σ_{gm}, Σ_{hm} and Σ_{hg}.

The identifying assumptions between the macro and yield factors are similar to standard assumptions made in the VAR literature; i.e. macro variables are slow moving and do not react to contemporaneous monetary policy shocks; but monetary policy does respond to contemporaneous macroeconomic shocks; see, e.g. Christiano, Eichenbaum, and Evans(1999),

Stock and Watson(2001), Bernanke, Boivin, and Eliasz(2005), and Wu and Xia(2014). We make an additional assumption about the volatility factors: only the lagged volatility factors affect the macro variables and yield curve; on the other hand, volatility does react to monetary policy and macroeconomic shocks contemporaneously. Note that our results are robust to reordering of the variables.

Of critical importance for our analysis, interest rate uncertainty h_t impacts the macroeconomy through the conditional mean term $\Phi_{mh}h_t$ in (1), which is identified from the conditional variance of the yield curve through D_t. This unique combination unifies two literatures; the literature on vector autoregressions with stochastic volatility (e.g. Cogley and Sargent(2001), Cogley and Sargent(2005), and Primiceri(2005)) and the more recent uncertainty literature that uses VARs to study uncertainty shocks and fluctuations in the macroeconomy and/or asset-prices (e.g. Baker, Bloom, and Davis(2013), Jurado, Ludvigson, and Ng(2013), Bekaert, Hoerova, and Lo Duca(2013), and Aastveit, Natvik, and Sola(2013)).

By choosing a log-normal process for the volatility in (3), we can utilize popular algorithms for Bayesian estimation of stochastic volatility models; see, e.g. Kim, Shephard, and Chib(1998). The matrices Γ_0 and Γ_1 permit a factor structure within the covariance matrix and allow us to estimate models where the number of volatility factors and yield factors may differ with $G \neq H$.

Our dynamic setup is related to the GARCH-M (GARCH-in-mean) literature within a VAR; see, e.g. Engle, Lilien, and Robins(1987) and Elder(2004). The difference is that we use stochastic volatility instead of GARCH to model time-varying variances, meaning that volatility is not a deterministic function of past levels, but has its own innovations. This is important because with our framework, we can use tools from the VAR literature such as impulse responses, which are not as easily defined in GARCH-M type models because there is no separate shock to volatility. Jo(2013) is similar to our paper in this spirit: while her focus is oil price shocks, we focus on uncertainty shocks from the term structure of interest rates.

2.2 Bond prices

Zero coupon bonds are priced to permit no arbitrage opportunities. The literature on affine term structure models demonstrates that to have h_t realistically capture yield volatility, it cannot price bonds; see, e.g. Creal and Wu(2014). In our model, this means that the yield factors g_t summarize all the information for the cross section of the yield curve.

The short rate is an affine function of g_t:

$$r_t = \delta_0 + \delta_1' g_t. \tag{5}$$

The risk neutral $\mathbb{Q}$ measure adjusts the probability distribution in (2) to incorporate investors' risk premium, and is defined such that the price of an asset is equal to the present value of its expected payoff. For an n-period zero coupon bond,

$$P_t^n = \mathbb{E}_t^{\mathbb{Q}} \left[\exp\left(-r_t\right) P_{t+1}^{n-1} \right], \tag{6}$$

where the risk-neutral expectation is taken under the autonomous VAR(1) process for g_t:

$$g_{t+1} = \mu_g^{\mathbb{Q}} + \Phi_g^{\mathbb{Q}} g_t + \Sigma_g^{\mathbb{Q}} \varepsilon_{g,t+1}^{\mathbb{Q}}, \quad \varepsilon_{g,t+1}^{\mathbb{Q}} \sim \mathrm{N}\left(0, I\right). \tag{7}$$

As a result, zero-coupon bonds are an exponential affine function of the Gaussian state variables

$$P_t^n = \exp\left(\bar{a}_n + \bar{b}_n' g_t\right). \tag{8}$$

The bond loadings $\bar{a}_n$ and $\bar{b}_n$ can be expressed recursively as

$$\begin{aligned} \bar{a}_n &= -\delta_0 + \bar{a}_{n-1} + \mu_g^{\mathbb{Q}\prime} \bar{b}_{n-1} + \frac{1}{2} \bar{b}_{n-1}' \Sigma_g^{\mathbb{Q}} \Sigma_g^{\mathbb{Q}\prime} \bar{b}_{n-1}, \\ \bar{b}_n &= -\delta_1 + \Phi_g^{\mathbb{Q}\prime} \bar{b}_{n-1}, \end{aligned}$$

with initial conditions $\bar{a}_1 = -\delta_0$ and $\bar{b}_1 = -\delta_1$. Bond yields $y_t^n \equiv -\frac{1}{n}\log(P_t^n)$ are linear in the factors

$$y_t^n = a_n + b_n' g_t \tag{9}$$

with $a_n = -\frac{1}{n}\bar{a}_n$ and $b_n = -\frac{1}{n}\bar{b}_n$.

Our model introduces a novel approach to incorporating volatility factors in no-arbitrage term structure models while keeping bond prices simple through the assumptions (5) and (7). Most term structure models impose the covariance matrices under the $\mathbb{P}$ and $\mathbb{Q}$ measures to be the same in order to guarantee no-arbitrage in continuous time. Consequently, volatility factors enter the variance of g_t under $\mathbb{Q}$ and hence bond prices in general. To cancel the volatility factors out of the pricing equation, unspanned stochastic volatility (USV) models impose restrictions on the $\mathbb{Q}$ parameters that subsequently constrain the cross-sectional fit of the model. In our model, h_t does not enter the $\mathbb{Q}$ dynamics and it is not priced by construction. Consequently, our model does not impose any restrictions on the cross section of the yield curve. An advantage of working with discrete time models is that it allows different variance-covariance matrices under $\mathbb{P}$ and $\mathbb{Q}$, while still preserving no arbitrage.[2] We will show the no arbitrage condition – the equivalence of the two probability measures – by deriving the Radon-Nikodym derivative in Subsection 2.3. Because only g_t is priced, we do not need to specify the dynamics of m_t and h_t under $\mathbb{Q}$. Nor are they identified using bond prices alone even if we do.

The benefits of our specification are twofold. First, our dynamics for g_t under $\mathbb{Q}$ and hence the bond pricing formula are the same as in a Gaussian ATSM. Second, the separation of the covariance matrices under the two measures allows a more flexible $\mathbb{P}$ dynamics, since we are not limited to the functional forms that achieve analytical bond prices.

[2] In concurrent and independent work, Ghysels, Le, Park, and Zhu(2014) propose a term structure model where the pricing factors g_t have Gaussian VAR dynamics and whose covariance matrices under the $\mathbb{P}$ and $\mathbb{Q}$ measures are different. Their covariance matrix under $\mathbb{P}$ uses GARCH instead of stochastic volatility. Stochastic volatility allows us to study the impact of uncertainty shocks, which is the focus of this paper.

2.3 Stochastic discount factor

The pricing equation in (6) can be equivalently written as

$$P_t^n = \mathbb{E}_t \left[\mathcal{M}_{t+1} P_{t+1}^{n-1} \right]. \tag{10}$$

Instead of incorporating the risk premium in the probability distribution through a change of measure, we capture it by the stochastic discount factor (SDF) $\mathcal{M}_{t+1}$ defined as

$$\mathcal{M}_{t+1} = \frac{\exp(-r_t)\, p^{\mathbb{Q}}(g_{t+1}|\mathcal{I}_t;\theta)}{p(g_{t+1}|\mathcal{I}_t;\theta)}$$

where $\mathcal{I}_t$ contains all the information up to and including t, and θ summarizes all the parameters in the model.

3 Bayesian estimation

The ATSM with stochastic volatility is a non-linear, non-Gaussian state space model whose log-likelihood is not known in closed-form. We estimate the model by Bayesian methods using Markov chain Monte Carlo (MCMC) for the parameters and a particle filter to calculate the log-likelihood and filtered estimates of the state variables.[3] Our estimation method is highly tractable because we formulate the model as a conditionally, linear Gaussian state space model, which allows us to use the Kalman filter for most of our Gibbs sampler. We outline the basic ideas of the MCMC algorithm in this section and provide full details in Appendix B. Our paper contributes to the literature on Bayesian estimation of term structure models; see, e.g. Chib and Ergashev(2009) and Bauer(2014) for Gaussian affine term structure models. The MCMC algorithms developed in this paper are efficient and can handle a wide range of models.

[3]Particle filters are simulation based algorithms for handling non-linear, non-Gaussian state space models; see Creal(2012) for a survey.

3.1 State space forms

Stack the yields y_t^n from (9) in order of increasing maturity for $n = n_1, n_2, ..., n_N$, and assume that all yields are observed with Gaussian measurement errors:

$$y_t = A + Bg_t + \eta_t, \qquad \eta_t \sim \mathrm{N}(0, \Omega), \tag{11}$$

where $A = (a_{n_1}, \ldots, a_{n_N})'$, $B = (b'_{n_1}, ..., b'_{n_N})'$. Under the assumption that all yields are measured with error, both the yield factors $g_{1:T} = (g_1, \ldots, g_T)$ and the volatility factors $h_{0:T} = (h_0, \ldots, h_T)$ are latent state variables. Let $y_{1:T} = (y_1, \ldots, y_T)$ and $m_{1:T} = (m_1, \ldots, m_T)$. Using data augmentation and the Gibbs sampler, we draw from the joint posterior distribution $p(g_{1:T}, h_{0:T}, \theta | y_{1:T}, m_{1:T})$. The Gibbs sampler iterates between drawing from the full conditional distributions of the yield factors $p(g_{1:T} | y_{1:T}, m_{1:T}, h_{0:T}, \theta)$, the volatility factors $p(h_{0:T} | y_{1:T}, m_{1:T}, g_{1:T}, \theta)$, and the parameters θ. At each step of the algorithm, we formulate a linear, Gaussian state space model that conditions on the most recent draw of the yield factors $g_{1:T}$ or the volatility factors $h_{0:T}$. Next, we outline the two linear Gaussian state space forms used in our algorithm.

State space form I conditional on $h_{0:T}$ Conditional on the most recent draw of $h_{0:T}$, the model has a linear Gaussian state space form: the state variable g_t has a transition equation in (2), and the observation equations for this state space model combine yields y_t in (11), the macroeconomic variables m_t in (1), and the volatility factors h_t in (4). Using this representation, we draw the latent yield factors $g_{1:T}$ and many of the parameters that enter the dynamics of g_t.

State space form II conditional on $g_{1:T}$ Conditional on the most recent draw of $g_{1:T}$, we have a state space model with observation equations for the macro variables and yields in (1) and (2) and transition equation for h_t in (4); the observation equation for yields makes the system non-linear. We use the approach of Kim, Shephard, and Chib(1998) to transform the

non-linear observation equation in (2) into an (approximate) linear, Gaussian model. They transform the nonlinear observation equation (2) into a linear equation with non-Gaussian shocks. The non-Gaussian shocks are then approximated by a mixture of normals, where each component of the mixture is indexed by a latent variable s_{it} for $i = 1, \ldots, G$ and $t = 1, \ldots, T$. The details of this transformation are in Appendix B.1.2.

3.2 MCMC and particle filter

Our MCMC algorithm alternates between the two state space forms. We split the parameters into blocks $\theta = (\theta_g', \theta_h', \theta_r')'$, where we draw θ_g from the state space model when g_t is latent, θ_h from the state space model when h_t is latent, θ_r conditional on $g_{1:T}$ and $h_{0:T}$. Here we sketch the rough steps and leave the details to Appendix B.

1. Conditional on $h_{0:T}$, use the linear Gaussian state space model I.
   (a) Draw θ_g using the Kalman filter without conditioning on $g_{1:T}$.
   (b) Draw $g_{1:T}$ using forward filtering and backward sampling; see, e.g. de Jong and Shephard(1995), Durbin and Koopman(2002).
2. Draw the mixture indicators s_{it} for $i = 1, \ldots, G$ and $t = 1, \ldots, T$ as in Kim, Shephard, and Chib(1998), see Appendix B.1.2.
3. Conditional on $g_{1:T}$ and $s_{1:T}$, use the linear Gaussian state space model II.
   (a) Draw θ_h using the Kalman filter without conditioning on $h_{0:T}$.
   (b) Draw $h_{0:T}$ using forward filtering and backward sampling; see, e.g. de Jong and Shephard(1995), Durbin and Koopman(2002).
4. Draw any remaining parameters in θ_r conditional on both $g_{1:T}$ and $h_{0:T}$.

Iterating on these steps produces a Markov chain whose stationary distribution is the posterior distribution $p\left(g_{1:T}, h_{0:T}, \theta | y_{1:T}, m_{1:T}\right)$.

The particle filter that we implement is the mixture Kalman filter (MKF); see Chen and Liu(2000).[4] Similar to our MCMC algorithm, it utilizes the conditionally linear, Gaussian state space form for statistical efficiency. Intuitively, if the volatilities $h_{0:T}$ were known, then the Kalman filter would calculate the filtered estimates of $g_{1:T}$ and the likelihood of the model exactly. In practice, the value of the volatilities are not known. The MKF calculates a weighted average of Kalman filter estimates of $g_{1:T}$ where each Kalman filter is run with a different value of the volatilities. This integrates out the uncertainty associated with the volatilities. The statistical efficiency gains come from the fact that the Kalman filter integrates out the Gaussian state variables exactly once we condition on any one path of the volatilities.

4 Economic implications

Does interest rate uncertainty contribute to economic fluctuations, particularly at the business cycle frequency? If so, by how much? How does this behavior vary through time? This section investigates these questions by exploiting the modeling and estimation tools described in the previous sections. We focus on two aspects of interest rate uncertainty: short-term and long-term uncertainty. We illustrate that, although they both contribute negatively to economic activity by inducing higher unemployment, their impacts on inflation have opposite signs. The usefulness of two factors in capturing yield volatility is documented in Section 5.

Data, model and estimates We use the Fama-Bliss zero-coupon yields available from the Center for Research in Securities Prices (CRSP) with maturities $n = (1, 3, 12, 24, 36, 48, 60)$ months. We use consumer price index inflation and the unemployment rate as our macroeconomic variables, which were downloaded from the FRED database at the Federal Reserve Bank of St. Louis. Inflation is measured as the annual percentage change. We scale all

[4]The MKF has recently been applied in economics by Creal, Koopman, and Zivot(2010), Creal(2012), and Shephard(2013).

Table 1: Estimates for the benchmark macro model

$\delta_0 \times 1200$							Γ_0	$\Gamma_1 \times \frac{1}{1200}$					
-0.135							0	1	0				
(0.013)							–	–	–				
							0	0	1				
$\phi^{\mathbb{Q}}$							–	–	–				
0.996	0.949	0.717					-1.978	0.495	0.498				
(0.0005)	(0.003)	(0.020)					(0.121)	(0.012)	(0.012)				
$\bar{\mu}_m \times 1200$		$\bar{\mu}_g \times 1200$			$\bar{\mu}_h \times 1200$								
4.343	6.826	5.701	6.862	6.128	0.107	2.733							
(0.542)	(0.541)	(0.722)	(0.758)	(0.759)	(0.627)	(0.131)							
Φ_m		Φ_{mg}			Φ_{mh}		$\Sigma_m \times 1200$						
0.983	-0.012	0.039	-0.067	0.036	-0.036	0.019	0.371						
(0.008)	(0.013)	(0.048)	(0.033)	(0.068)	(0.020)	(0.013)	(0.010)						
0.011	0.975	0.044	0.032	-0.089	0.054	-0.010	-0.008	0.187					
(0.005)	(0.007)	(0.028)	(0.019)	(0.040)	(0.012)	(0.011)	(0.008)	(0.006)					
Φ_{gm}		Φ_g			Φ_{gh}		$\Sigma_{gm} \times 1200$		$\Sigma_g \times 1200$				
0.011	-0.010	0.706	-0.065	0.332	-0.008	0.002	0.015	-0.025	0.360				
(0.006)	(0.006)	(0.037)	(0.014)	(0.043)	(0.003)	(0.002)	(0.008)	(0.009)	–				
0.010	-0.003	0.010	0.965	0.014	-0.009	0.002	0.022	-0.023	0.184	0.360			
(0.006)	(0.007)	(0.041)	(0.021)	(0.051)	(0.003)	(0.002)	(0.010)	(0.008)	(0.012)	–			
0.012	-0.011	-0.138	-0.026	1.144	-0.012	0.003	0.016	-0.037	0.326	0.181	0.360		
(0.006)	(0.006)	(0.042)	(0.018)	(0.051)	(0.004)	(0.003)	(0.009)	(0.009)	(0.009)	(0.008)	–		
					Φ_h		$\Sigma_{hm} \times 1200$		$\Sigma_{hg} \times 1200$			$\Sigma_h \times 1200$	
					0.958	0.011	0.072	-0.051	0.033	-0.114	0.547	0.356	
					(0.015)	(0.010)	(0.034)	(0.033)	(0.033)	(0.068)	(0.191)	(0.052)	
					0.001	0.994	0.008	0.005	0.060	-0.033	0.098	0.140	0.230
					(0.012)	(0.006)	(0.028)	(0.028)	(0.029)	(0.058)	(0.162)	(0.050)	(0.032)

Posterior mean and standard deviations for the benchmark macroeconomic plus yields model. The parameters $\phi^{\mathbb{Q}}$ are the eigenvalues of $\Phi_g^{\mathbb{Q}}$ while $\bar{\mu}_m, \bar{\mu}_g$, and $\bar{\mu}_h$ are the unconditional means of the factors.

the variables by 1/1200. For example, a 5% annual interest rate is 5/1200. Our sample is monthly from June 1953 to December 2013.

Our model has $G = 3$ yield factors which are rotated to capture the 3 month, 5 year, and 1 year yields excluding measurement errors.[5] The $H = 2$ volatility factors capture the volatility of the 3 month yield and 5 year yield, and we label them short term uncertainty and long term uncertainty. Posterior means and standard deviations for the model's parameters are reported in Table 1.

4.1 Impulse responses

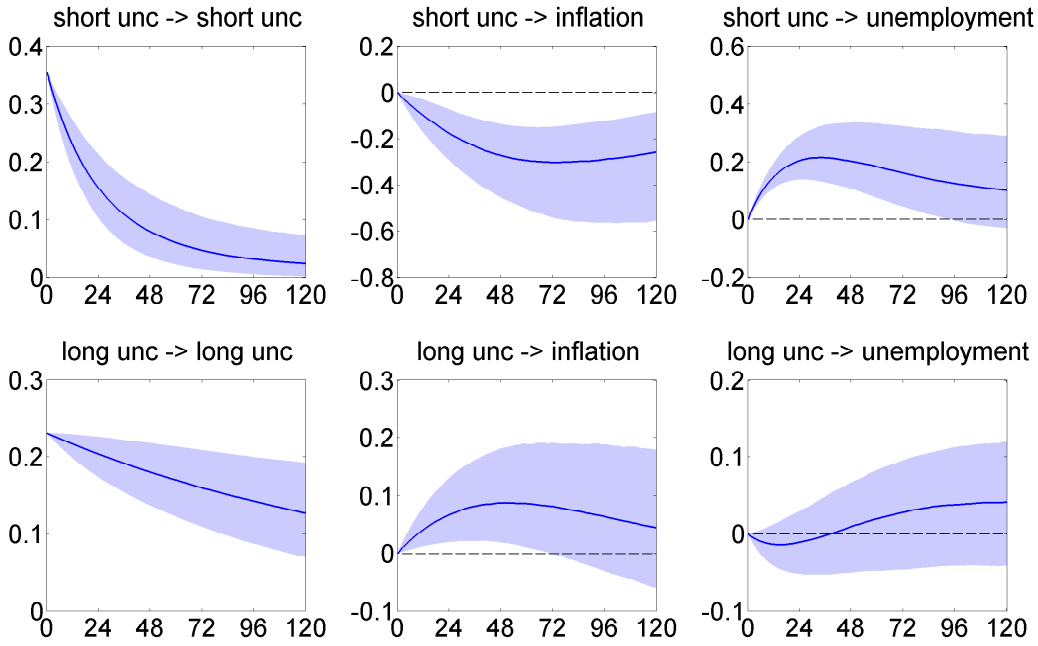
We focus on the dynamic response of macroeconomic variables m_t to uncertainty shocks ε_{ht}.

[5]This normalization is chosen to create economically meaningful Gaussian factors. As in a Gaussian ATSM, rotation of the factors does not have an economic impact on the model itself.

Constant impulse response In a VAR with homoskedastic shocks – where uncertainty only has a first moment effect and does not enter the conditional volatility, the impulse response to either a one standard deviation shock or a one unit shock is a constant function through time. It does not depend on the state variables. In a VAR with heteroskedastic shocks – where uncertainty only effects the second moment, the impulse responses to a one unit shock are a constant function although for a one standard deviation shock the scale varies through time because of the stochastic volatility. In either of these cases, the shape of the impulse response remains the same. For our model, we can define a time-invariant version of an impulse response by assuming that $\varepsilon_{g,t} = 0 \; \forall \; t$; see Appendix B.3 for details. This assumption eliminates the non-linear effect a shock to h_t has on g_{t+1} through D_t in (2). This constant impulse response provides a summary of the average of the time-varying impulse responses that we will discuss below.

We plot the median constant impulse response in Figure 2, with the [10%, 90%] highest posterior density intervals calculated from our MCMC draws in the shaded areas. A one standard deviation shock to short-term uncertainty is about 1/15 of the change in uncertainty immediately before the Great Recession. Short-term interest rate uncertainty dies out faster than long-term uncertainty. Both of them have a negative impact on the real economy: higher uncertainty is associated with higher future unemployment rates. The difference is the impact of the short-term uncertainty shock dies out faster, peaking at 0.2% in about 2 years. Long-term uncertainty leads to higher unemployment but at a much slower pace and the impact is not statistically significant. The major difference between the two interest rate uncertainties with different horizons is their relationship with inflation. A one time shock to short-term interest rate uncertainty leads to lower inflation, peaking at -0.3%. In contrast, a shock to long-term uncertainty is associated with higher inflation at a maximum of 0.1%. Both effects are statistically significant. These two effects compete with each other, and it is not a priori clear about the sign of the overall effect on inflation.

Figure 2: Constant impulse response functions

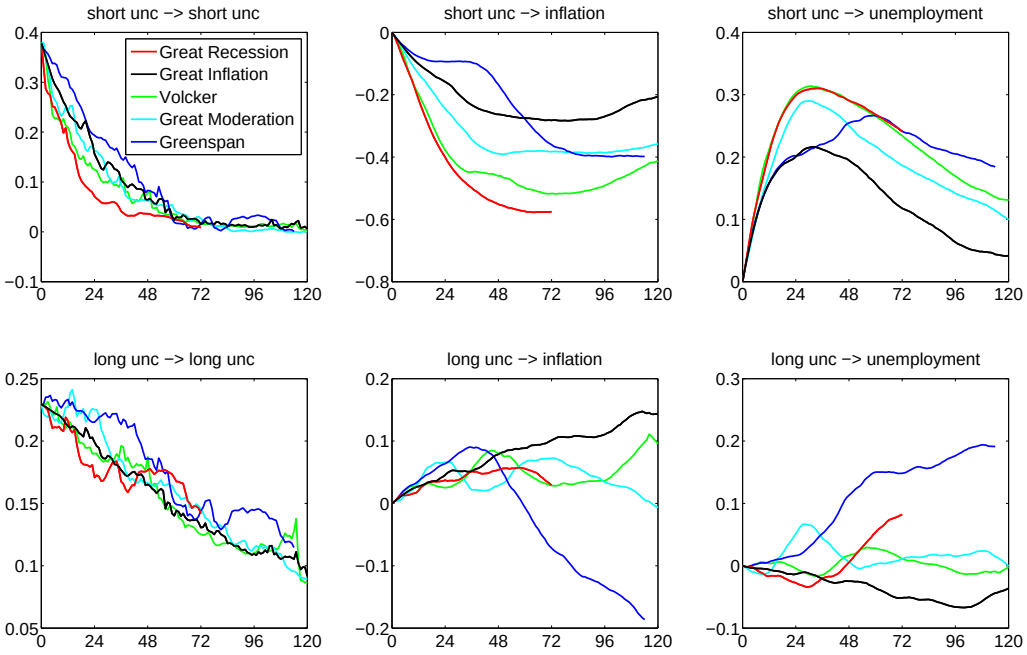


Constant impulse responses to a one standard deviation shock to interest rate uncertainty. Top: short-term uncertainty; bottom: long-term uncertainty. Left: uncertainty; middle: inflation; right: unemployment rate. The [10%, 90%] highest posterior density intervals are shaded. Sample: June 1953 - December 2013.

Time-varying impulse response A unique feature of our hybrid model – jointly capturing the first and second moment effects of uncertainty – is the existence of a state-dependent impulse response, with both the scale and shape varying across time. Our model can distinguish periods like the Great Recession (2007 - 2009) from the Great Moderation (1985 - 2007). It also helps us to aggregate the overall effect uncertainty has on inflation for a specific time period. Neither a standard VAR with homoskedasticity or with heteroskedasticity has this property. This feature usually only exists in VARs with time-varying autoregressive coefficients. Our model gets the same benefit without explicitly introducing many more state variables as time-varying parameters.

We plot the median impulse response to a one standard deviation shock in Figure 3 for the following economically significant time periods: the Great Recession in red from December 2007, the Great Inflation in black from 1965, Volcker's tenure in green from August 1979, the

Figure 3: Time-varying impulse response functions



Impulse responses to a one standard deviation shock to interest rate uncertainty. Top: short-term uncertainty; bottom: long-term uncertainty. Left: uncertainty; middle: inflation; right: unemployment rate. Sample periods: the Great Recession in red from December 2007, the Great Inflation in black from 1965, the start of Volcker's tenure in blue from August 1979, the Great Moderation in turquoise from 1985, and Greenspan's conundrum in pink from June 2004.

Great Moderation in turquoise from 1985, and Greenspan's conundrum in blue from June 2004. See Appendix B.3 for the calculations.

The responses of uncertainty to their own shocks in the first column are relatively homogeneous across each of the episodes. The impulse responses for inflation and unemployment are different. During the Great Recession, short-term interest rate uncertainty has a bigger negative impact on the unemployment rate. This period is associated with increased uncertainty over unconventional monetary policy – quantitative easing and the Fed's exit strategy for example. The larger response of the unemployment rate coincides with the jobless recovery. At the same time, the same uncertainty leads to lower inflation with a bigger magnitude than other periods. This reflects the publics' concerns over deflation during this time period.

In contrast, the Great Inflation is a period associated with a positive inflation response to

uncertainty: increased long-term uncertainty leads to increases in inflation, higher than any other time with an increasing trend 10 years out; increased short-term uncertainty still leads to a decrease in inflation, but with a relatively milder magnitude. The subsequent Volker tenure contrasts with the Great Inflation in that the short-term uncertainty (uncertainty about monetary policy) now decreases inflation with twice the magnitude. This reflects Volcker's reputation as an inflation hawk. At the same time, short-term uncertainty increases unemployment more so than before. This reinforces the view that Volcker placed more weight on inflation and tolerated more negative real activity than his predecessor.

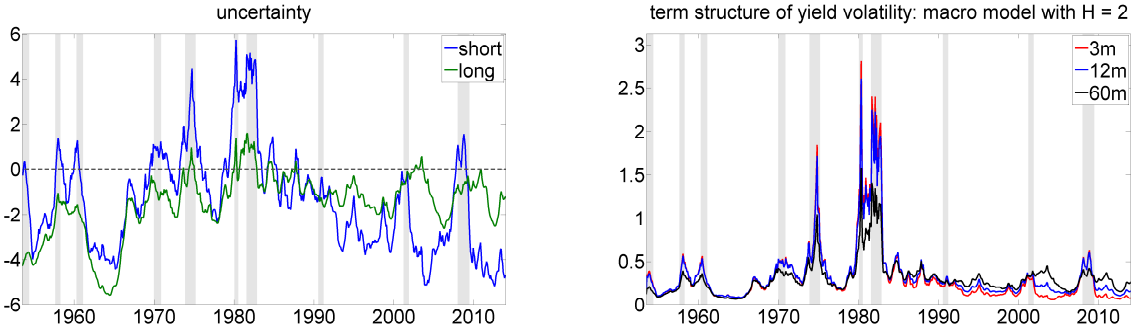
The Great Moderation is a calm period for economic activity and the impulse responses (turquoise lines) lie in-between the other time-periods. Inflation and unemployment do not respond aggressively to uncertainty shocks. Greenspan's conundrum describes the period where the Fed is actively increasing the short-term interest rate but the long-term rates are not responsive, or even move in the opposite direction. The impulse responses show a big drop in inflation in response to long-term interest rate uncertainty on top of the negative response of inflation to short-term uncertainty. This movement in future inflation is priced in bonds, by forward-looking agents putting downward pressure on long-term interest rates.

4.2 Estimates of uncertainty

We plot short-term and long-term interest rate uncertainties in the left panel of Figure 4. Both uncertainties increased in the first half of the sample, and peaked in the early 1980s. The short-term uncertainty displayed a decreasing trend since, although it increased again right before the Great Recession, and settled down at the end of the sample. Whereas long-term uncertainty remained relatively stable for the second half of the sample, and started increasing again recently.

Magnitude of uncertainty The left panel of Figure 4 provides a visual inspection of the magnitude of a one standard deviation shock as discussed in Subsection 4.1. One standard

Figure 4: Uncertainty and volatility



Left: short-term and long-term uncertainty from the macro plus yields model. Right: conditional (smoothed) volatility of yields for the 3, 12, and 60 month maturities.

deviation of the short-term uncertainty shock is about 0.4. The change in uncertainty immediately before the Great Recession is about 6 (15 times the shock size), and the change is about 8 before the 1980 recession (20 standard deviations). One standard deviation for the long-term uncertainty shock is about 0.23. The hike of long-term uncertainty before the Great Recession is about 11 times this magnitude, slightly smaller than the movement for the short-term uncertainty relative to their respective standard deviations.

Uncertainty and recession The left panel of Figure 4 shows that both short-term and long-term interest rate uncertainties are counter-cyclical: they increase drastically before almost every recession and remain high throughout recessions; when the economy recovers, both measures of uncertainty drop. This pattern is more pronounced for short-term uncertainty than long-term uncertainty.

To reflect this observation statistically, we use the following simple regression:

$$h_{jt} = \alpha + \beta \mathbb{1}_{recession,t} + u_{jt}, \tag{12}$$

where $\mathbb{1}_{recession,t}$ is a recession dummy, and takes a value of 1 if time t is dated within a recession by the NBER. The coefficients are 2.3 for short-term uncertainty meaning that

uncertainty is 2.3 units higher during the recessions as opposed to expansions, and the difference is 0.6 for long-term uncertainty. Both coefficients are statistically significant with p-values numerically at zero.

5 Model comparison

5.1 Model specifications

Identification and other model restrictions All the models we study have $G = 3$ yield factors, as this is a well-established choice in the term structure literature. We normalize the yield factors to be the model-implied short-term, long-term, and medium-term (3, 60, and 12 months) yields to study the term structure of interest rate uncertainty. Our method to achieve this rotation is similar to Hamilton and Wu(2012). The differences are that our model has heteroskedastic shocks, and that we assume that all yields are observed with measurement error. Parameter restrictions are: (i) Σ_g is lower triangular; and (ii) $\Phi_g^{\mathbb{Q}}$ has three free parameters that are its eigenvalues labeled $\phi^{\mathbb{Q}}$. Further details on this rotation are in Appendix A. These restrictions prevent rotation of the yield-factor state variables g_t.

To prevent the volatility factors from rotating and scaling, we impose (i) Σ_h is lower triangular; (ii) Γ_1 has an identity matrix $1200 \times I_H$ as its top block; scaling the volatility factors h_t up by 1200 gives them a similar scale as the observed variables m_t and y_t; (iii) when $H > 0$, the diagonal elements of Σ_g are fixed at 0.0003, which is the same magnitude as the estimates of Σ_g from the Gaussian ATSM when $H = 0$; (iv) the first H elements of Γ_0 are equal to zero.

We use the procedures of Kim, Shephard, and Chib(1998) as part of our MCMC algorithm, which requires the restriction $\Phi_{gh} = \Sigma_{gm}\Sigma_m^{-1}\Phi_{mh}$, as explained in Appendix B. We also restrict the covariance matrix under $\mathbb{Q}$ to be equal to the long-run mean under $\mathbb{P}$; i.e. $\Sigma_g^{\mathbb{Q}} = \Sigma_g \text{diag}\left(\exp\left[\frac{\Gamma_0+\Gamma_1\bar{\mu}_h}{2}\right]\right)$. This restriction implies that our model nests Gaussian ATSMs as $\Sigma_h \to 0$.

Models To understand the factor structure in volatility, on top of the macro model studied in Section 4, we compare yield only models $M = 0$ with $H = 0, 1, 2, 3$ volatility factors. We label these models $\mathbb{H}_H$. Yields in the $\mathbb{H}_1$ model share one common volatility factor. Our choice of Γ_1 implies that the volatility factors in the $\mathbb{H}_2$ model capture the short-term (3 month) and long-term (60 month) volatilities. The $\mathbb{H}_3$ model adds another degree of freedom for the medium-term (12 month) volatility.

Estimates Estimates of the posterior mean and standard deviation for the parameters of all four yields only models as well as the log-likelihood and BIC (evaluated at the posterior mean) are reported in Table 2. Priors for all parameters of the model are discussed in Appendix C. The parameter estimates for the $\mathbb{H}_0(3)$ model are typical of those found in the literature on Gaussian ATSMs, see Hamilton and Wu(2012). We also note that with the introduction of stochastic volatility the estimated values of the autoregressive matrix Φ_g become more persistent. The modulus of the eigenvalues of this matrix are larger for all the stochastic volatility models. Due to the increased persistence, the long-run mean parameters $\bar{\mu}_g$ of the yield factors are larger than for the $\mathbb{H}_0$ model and closer to the unconditional sample mean of yields.

5.2 Yield volatilities

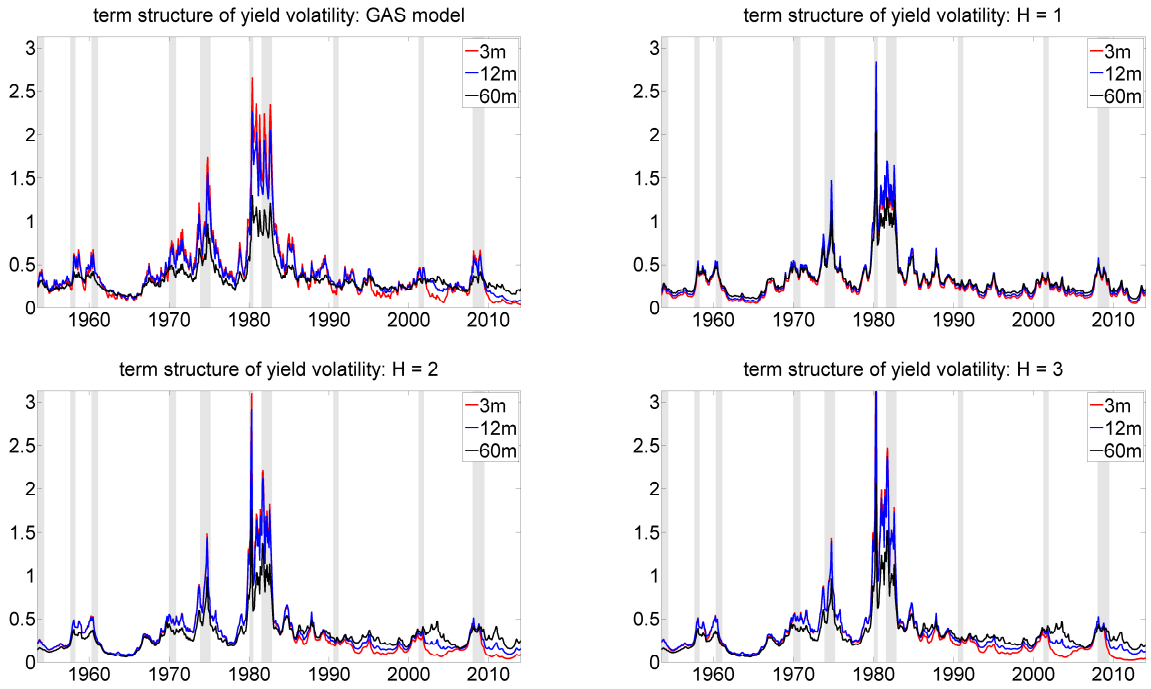
Yield only models We first compare yield only models $\mathbb{H}_H$ in terms of fitting the yield volatility, and select the number of volatility factors needed to describe the data. Table 2 shows that the introduction of the first stochastic volatility factor causes an enormous increase in the log-likelihood from 37581.8 for the $\mathbb{H}_0$ model to 38091.1 for the $\mathbb{H}_1$ model. The addition of a second volatility factor that captures movements in long-term yields adds another 100 points to the likelihood of the $\mathbb{H}_2$ model to 38191.3. Adding a third volatility factor increases the likelihood by less than 30 points to 38227.0. As the number of volatility factors increases, the number of parameters also increases. The BIC penalizes the log-likelihood for

Table 2: Estimates for yields-only term structure models

$\mathbb{H}_0$			$\mathbb{H}_1$				$\mathbb{H}_2$					$\mathbb{H}_3$					
	log-like:	37581.1			log-like:	38091.1				log-like:	38191.3					log-like:	38227.0
	BIC:	-74735			BIC:	-75662				BIC:	-75837					BIC:	-75827
$\delta_0 \times 1200$			$\delta_0 \times 1200$				$\delta_0 \times 1200$					$\delta_0 \times 1200$					
-0.100			-0.123				-0.138					-0.134					
(0.012)			(0.017)				(0.013)					(0.013)					
$\phi^{\mathbb{Q}}$			$\phi^{\mathbb{Q}}$				$\phi^{\mathbb{Q}}$					$\phi^{\mathbb{Q}}$					
0.995	0.956	0.669	0.995	0.949	0.679		0.995	0.949	0.712			0.996	0.949	0.720			
(0.0006)	(0.003)	(0.021)	(0.0005)	(0.004)	(0.024)		(0.0005)	(0.003)	(0.020)			(0.0005)	(0.003)	(0.019)			
$\bar{\mu}_g \times 1200$			$\bar{\mu}_g \times 1200$				$\bar{\mu}_g \times 1200$					$\bar{\mu}_g \times 1200$					
3.003	3.879	3.291	3.847	4.188	4.014		3.824	4.196	4.014			3.443	3.842	3.627			
(0.508)	(0.457)	(0.513)	(1.259)	(1.324)	(1.331)		(0.976)	(0.921)	(1.010)			(0.953)	(0.911)	(0.987)			
Φ_g			Φ_g			Γ_0	Φ_g			Γ_0		Φ_g			Γ_0		
0.818	-0.033	0.199	0.716	-0.062	0.3275	0	0.712	-0.047	0.317	0		0.745	-0.033	0.273	0		
(0.058)	(0.028)	(0.074)	(0.039)	(0.017)	(0.047)	—	(0.036)	(0.011)	(0.041)	—		(0.035)	(0.011)	(0.039)	—		
0.119	1.006	-0.125	0.017	0.979	0.001	-0.374	-0.006	0.985	0.014	0		0.007	0.992	-0.004	0		
(0.042)	(0.021)	(0.054)	(0.045)	(0.020)	(0.056)	(0.118)	(0.037)	(0.016)	(0.045)	—		(0.036)	(0.016)	(0.044)	—		
0.009	0.014	0.970	-0.122	-0.025	1.1372	-2.213	-0.135	-0.012	1.135	-1.8938		-0.118	-0.0004	1.109	0		
(0.054)	(0.027)	(0.070)	(0.045)	(0.017)	(0.054)	(0.136)	(0.040)	(0.013)	(0.046)	(0.118)		(0.038)	(0.013)	(0.044)	—		
$\Sigma_g \times 1200$			$\Sigma_g \times 1200$			$\Gamma_1 \times \frac{1}{1200}$	$\Sigma_g \times 1200$			$\Gamma_1 \times \frac{1}{1200}$		$\Sigma_g \times 1200$			$\Gamma_1 \times \frac{1}{1200}$		
0.388	0	0	0.360	0	0	1	0.360	0	0	1	0	0.360	0	0	1	0	0
(0.012)	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
0.190	0.206	0	0.230	0.360	0	0.557	0.191	0.360	0	0	1	0.190	0.360	0	0	1	0
(0.011)	(0.008)	—	(0.027)	—	—	(0.101)	(0.011)	—	—	—	—	(0.011)	—	—	—	—	—
0.338	0.124	0.080	0.343	0.179	0.360	0.998	0.332	0.175	0.360	0.494	0.499	0.333	0.184	0.360	0	0	1
(0.012)	(0.007)	(0.005)	(0.011)	(0.009)	—	(0.008)	(0.009)	(0.009)	—	(0.012)	(0.0120)	(0.008)	(0.009)	—	—	—	—
			$\Sigma_{hg} \times 1200$			$\bar{\mu}_h \times 1200$	$\Sigma_{hg} \times 1200$			$\bar{\mu}_h \times 1200$		$\Sigma_{hg} \times 1200$			$\bar{\mu}_h \times 1200$		
			0.059	-0.117	0.354	2.439	0.104	-0.019	0.382	0.330		0.121	-0.010	0.278	0.386		
			(0.034)	(0.068)	(0.166)	(1.410)	(0.027)	(0.071)	(0.180)	(0.602)		(0.029)	(0.076)	(0.203)	(0.678)		
							0.045	-0.123	0.252	2.670		0.049	-0.133	0.205	2.754		
							(0.024)	(0.059)	(0.148)	(0.136)		(0.024)	(0.062)	(0.152)	(0.132)		
												0.068	0.027	0.647	-0.510		
												(0.047)	(0.091)	(0.239)	(0.485)		
			Φ_h			$\Sigma_h \times 1200$	Φ_h			$\Sigma_h \times 1200$		Φ_h			$\Sigma_h \times 1200$		
			0.984			0.504	0.981	0.003		0.370	0	0.981	-0.004	0.012	0.378	0	0
			(0.017)			(0.070)	(0.012)	(0.007)		(0.058)	—	(0.021)	(0.031)	(0.055)	(0.067)	—	—
							0.008	0.991		0.141	0.237	0.013	0.992	-0.008	0.102	0.241	0
							(0.010)	(0.006)		(0.054)	(0.033)	(0.015)	(0.019)	(0.035)	(0.059)	(0.034)	—
												0.042	0.064	0.869	0.036	0.133	0.272
												(0.024)	(0.034)	(0.057)	(0.101)	(0.075)	(0.060)

Posterior mean and standard deviations for four yield only models with $H = 0, 1, 2, 3$ volatility factors. The log-likelihood and BIC are evaluated at the posterior mean.

Figure 5: Estimated conditional volatility of yields from four models



Estimated conditional volatility of $3, 12, 60$ *month yields from four different models. Top left: univariate generalized autoregressive score model; Top right:* $\mathbb{H}_1(4)$ *model; Bottom left:* $\mathbb{H}_2(5)$ *model; Bottom right:* $\mathbb{H}_3(6)$ *model.*

these added parameters. It selects the $\mathbb{H}_2$ model with two volatility factors as the best model for its overall fit.

In Figure 5, we compare the estimated volatilities from the three yields only models with the generalized autoregressive score volatility model from Creal, Koopman, and Lucas(2011) and Creal, Koopman, and Lucas(2013).[6] The top left panel plots the conditional volatility of the $3, 12$, and 60 month yields from the univariate GAS models. This graph illustrates a factor structure for yield volatilities of different maturities. There is a common element driving the co-movement between different maturities. At the same time, they have distinct behavior across time. In the first half of the sample, the term structure of yield volatilities

[6]For each maturity n, we estimate an AR(1) model for the conditional mean of yields and the Student's t generalized autoregressive score (GAS) model of Creal, Koopman, and Lucas(2011) and Creal, Koopman, and Lucas(2013) for the conditional volatility.

sloped downward as the volatility of short-term interest rates was higher than long-term rates. This reflects uncertainty about monetary policy. Short term rates became less volatile than long-term rates after the early 1980's and the term structure of volatility sloped upward on average. This may reflect efforts by the monetary authorities to make policy more transparent and better anchor agents' expectations. In the mid-2000s, the volatility of long and short rates moved in opposite directions with long-term volatility increasing at the same time that short-term volatility is declining.

The remaining panels in Figure 5 plot the conditional volatility from the ATSM's. In the $\mathbb{H}_1$ model, movements in yield volatility are nearly perfectly correlated. With only one factor, the model is not flexible enough to capture the idiosyncratic movements across different maturities that are observed in the data. This is consistent with the findings in Creal and Wu(2014). The $\mathbb{H}_2$ model adds flexibility through a second factor that drives the difference in volatility between the short-term and long-term yields. This model captures all the key features in the data we describe above. Although adding a third volatility factor provides more flexibility, the key economically meaningful movements are already captured by the first and second factors. Moreover, in the $\mathbb{H}_3$ model, the second and third volatility factors h_{2t} and h_{3t} are highly correlated with a value of 0.76, while the correlation between the first and second factors h_{1t} and h_{2t} is 0.4. Overall, both economic and statistical evidence points to two volatility factors.

Macro model Our benchmark macro model studied in Section 4 adds two macro variables – inflation and the unemployment rate – into the $\mathbb{H}_2$ model selected above. The conditional volatilities from this model are plotted in the right panel of Figure 4. The estimated conditional volatilities from the macro model are nearly identical to the estimates from the yields-only $\mathbb{H}_2$ model and capture all the characteristics of yield volatility discussed above. Overall, our benchmark macro model fits the yield volatility similarly to the preferred yield only model.

Table 3: Pricing errors relative to the Gaussian model

	$\mathbb{H}_0$	$\mathbb{H}_1$	$\mathbb{H}_2$	$\mathbb{H}_3$	macro
1m	0.2524	0.9917	1.0170	1.0539	1.0059
3m	0.1283	0.7155	0.6539	0.6196	0.7007
12m	0.1262	0.7726	0.7599	0.7583	0.7995
24m	0.0941	1.0499	0.9904	0.9586	0.9894
36m	0.0781	0.9577	0.8912	0.8489	0.8822
48m	0.1070	0.9103	0.8748	0.8598	0.8804
60m	0.0841	0.9382	0.8644	0.8728	0.9298

First column: Posterior mean estimates of the pricing errors $\sqrt{\text{diag}\,(\Omega)} \times 1200$ *for Gaussian the* $\mathbb{H}_0$ *model. Column 2-5: ratios of pricing errors of other models relative to the* $\mathbb{H}_0$ *model.*

5.3 Cross section of the yield curve

Our term structure models are designed to capture the volatility of yields while not sacrificing their ability to fit the cross section of the yield curve. In fact, we find that by introducing stochastic volatility it improves their ability to fit the yield curve at the same time. In Table 3, we report the average pricing errors across the seven maturities for the Gaussian $\mathbb{H}_0$ model in the first column. The next three columns report the ratios of pricing errors for the yields-only models with $H = 1, 2, 3$ relative to the $\mathbb{H}_0$ model for the same maturity. In general, the pricing errors decrease as more volatility factors are added. Relative to the Gaussian $\mathbb{H}_0$ model, the largest improvement comes with the addition of the first volatility factor. Adding a second volatility factor reduces the pricing errors further at all maturities but the one month yield. Smaller pricing errors are expected when we introduce stochastic volatility into the model. Stochastic volatility allows the signal-to-noise ratio to vary through time. During periods of low conditional volatility, the data is highly informative and estimates of the yield factors g_t will weigh the recent data more heavily. Conversely, estimates of the yield factors will place more weight on past data when conditional yield volatility is higher. This has the effect of stabilizing the estimator of g_t.

The last column of Table 3 shows the ratio of pricing errors in our benchmark macro model relative to the $\mathbb{H}_0$ model. The average pricing errors are smaller than the $\mathbb{H}_0$ model –

the most widely used model in the literature – and are comparable to the 2-factor $\mathbb{H}_2$ model. Adding the macro factors into the model does not affect the fit of the cross-section of yields.

6 Conclusion

We developed a new macro finance affine term structure model with stochastic volatilities to study the importance of interest rate uncertainty. In our model, the volatility factor serves two roles: it is the volatility of the yield curve, and it also measures interest rate uncertainty which directly interacts with macro variables in a VAR. Our model allows multiple volatility factors, which are determined separately from the yield factors. With two volatility factors and three traditional yield factors, our model can capture both aspects of the data.

The two volatility factors capture two aspects of interest rate uncertainty: short-term and long-term uncertainty. Although they both contribute negatively to the real economy, their influence over inflation has distinct signs: a shock to the short term interest rate uncertainty leads to lower inflation, whereas a higher long term uncertainty leads to higher inflation. Our model provides us time-varying impulse responses to study relative importance of these two competing effect at different point of time. For example, during Volcker's tenure, Greenspan's Conundrum and the Great Recession, uncertainty shocks are associated with downward pressure on inflation; whereas Great Inflation is associated with upward inflation pressure.

References

Aastveit, Knut Are, Gisle James Natvik, and Sergio Sola (2013) "Economic uncertainty and the effectiveness of monetary policy." Unpublished manuscript, Norges Bank.

Andersen, Torben, and Luca 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.

Baker, Scott R., Nicholas Bloom, and Steven J. Davis (2013) "Measuring economic policy uncertainty." University of Chicago, Booth School of Business, Working paper.

Bauer, Michael D. (2014) "Bayesian Estimation of Dynamic Term Structure Models under Restrictions on Risk Pricing" Federal Reserve Bank of San Francisco Working Paper 2011-03.

Bekaert, Geert, Marie Hoerova, and Marco Lo Duca (2013) "Risk, uncertainty, and monetary policy." Working paper, European Central Bank.

Bernanke, Ben S., Jean Boivin, and Piotr Eliasz (2005) "Measuring monetary policy: a factor augmented vector autoregressive (FAVAR) approach." *Quarterly Journal of Economics* 120, 387–422.

Bikbov, Ruslan, and Mikhail Chernov (2009) "Unspanned stochastic volatility in affine models: evidence from eurodollar futures and options." *Management Science* 55, 1292–1305.

Bloom, Nicholas (2013) "Fluctuations in uncertainty." *Journal of Economic Perspectives* forthcoming.

Chen, Rong, and Jun S. Liu (2000) "Mixture Kalman filters." *Journal of the Royal Statistical Society, Series B* 62, 493–508.

Chib, Siddhartha, and Bakhodir Ergashev (2009) "Analysis of multifactor affine yield curve models." *Journal of the American Statistical Association* 104, 1324–1337.

Chib, Siddhartha, and Edward Greenberg (1994) "Bayes inference in regression models with ARMA(p,q) errors." *Journal of Econometrics* 64, 183–206.

Christensen, Jens H.E., Jose A. Lopez, and Glenn D. Rudebusch (2014) "Can spanned term structure factors drive stochastic yield volatility?" Working paper, Federal Reserve Bank of San Francisco.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans (1999) "Monetary policy shocks: What have we learned and to what end?" in *Handbook of macroeconomics* Elsevier.

Cieslak, Anna, and Pavol Povala (2013) "Information in the term structure of yield curve volatility." Working Paper, Northwestern University.

Cogley, Timothy, and Thomas J. Sargent (2001) "Evolving post-World War II U.S. inflation dynamics" in *NBER Macroeconomics Annual*, edited by Ben S. Bernanke and Kenneth Rogoff pages 331–388.

Cogley, Timothy, and Thomas J. Sargent (2005) "Drift and volatilities: monetary policies and outcomes in the post WWII U.S." *Review of Economic Dynamics* 8, 262–302.

Collin-Dufresne, Pierre, and Robert S. Goldstein (2002) "Do bonds span the fixed income markets? Theory and evidence for unspanned stochastic volatility." *The Journal of Finance* 57, 1685–1730.

Creal, Drew D. (2012) "A survey of sequential Monte Carlo methods for economics and finance." *Econometric Reviews* 31, 245–296.

Creal, Drew D., Siem Jan Koopman, and André Lucas (2011) "A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations." *Journal of Business and Economic Statistics* 29, 552–563.

Creal, Drew D., Siem Jan Koopman, and André Lucas (2013) "Generalized autoregressive score models with applications." *Journal of Applied Econometrics* 28, 777–795.

Creal, Drew D., Siem Jan Koopman, and Eric Zivot (2010) "Extracting a robust U.S. business cycle using a time-varying multivariate model-based bandpass filter." *Journal of Applied Econometrics* 25, 695–719.

Creal, Drew D., and Jing Cynthia Wu (2014) "Estimation of affine term structure models with spanned or unspanned stochastic volatility." Working paper, University of Chicago, Booth School of Business.

Dai, Qiang, and Kenneth J. Singleton (2000) "Specification analysis of affine term structure models." *The Journal of Finance* 55, 1943–1978.

de Jong, Piet, and Neil Shephard (1995) "The simulation smoother for time series models" *Biometrika* 82, 339–350.

Del Negro, Marco, and Frank Schorfheide (2011) "Bayesian Macroeconometrics." in *Handbook of Bayesian Econometrics*, edited by John Geweke, Gary Koop, and Herman K. van Dijk Oxford University Press, Oxford pages 293–389.

Duffee, Gregory R. (2002) "Term premia and interest rate forecasts in affine models" *The Journal of Finance* 57, 405–443.

Durbin, James, and Siem Jan Koopman (2002) "A simple and efficient simulation smoother for state space time series analysis." *Biometrika* 89, 603–616.

Elder, John (2004) "Another perspective on the effects of inflation uncertainty." *Journal of Money, Credit, and Banking* 36, 911–928.

Engle, Robert F., David M Lilien, and Russel P Robins (1987) "Estimating time varying risk premia in the term structure: the ARCH-M model." *Econometrica* 55, 391–407.

Ghysels, Eric, Ahn Le, Sunjin Park, and Haoxiang Zhu (2014) "Risk and return trade-off in the U.S. treasury market." Working Paper, Department of Economics, University of North Carolina.

Hamilton, James D., and Jing Cynthia Wu (2012) "Identification and estimation of

Gaussian affine term structure models." *Journal of Econometrics* 168, 315–331.

Hamilton, James D., and Jing Cynthia Wu (2014) "Testable implications of affine term structure models." *Journal of Econometrics* 178, 231–242.

Jo, Soojin (2013) "The effects of oil price uncertainty on global real economic activity." Working paper, Bank of Canada.

Joslin, Scott (2010) "Pricing and hedging volatility risk in fixed income markets." Working paper, MIT Sloan School of Management.

Jurado, Kyle, Sydney C. Ludvigson, and Serena Ng (2013) "Measuring uncertainty." Working paper, Columbia University, Department of Economics.

Kim, Sangjoon, Neil Shephard, and Siddhartha Chib (1998) "Stochastic volatility: likelihood inference and comparison with ARCH models." *The Review of Economic Studies* 65, 361–393.

Liu, Jun S., and Rong Chen (1998) "Sequential Monte Carlo computation for dynamic systems." *Journal of the American Statistical Association* 93, 1032–1044.

Mueller, Phillipe, Andrea Vedolin, and Yu-min Yen (2011) "Bond variance risk premia" Unpublished Manuscript, London School of Economics.

Omori, Yasuhiro, Siddhartha Chib, Neil Shephard, and Jouchi Nakajima (2007) "Stochastic volatility with leverage: fast and efficient likelihood inference." *Journal of Econometrics* 140, 425–449.

Pastor, Lubos, and Pietro Veronesi (2012) "Uncertainty about government policy and stock prices." *The Journal of Finance* 67, 1219–1264.

Pastor, Lubos, and Pietro Veronesi (2013) "Political uncertainty and risk premia." *Journal of Financial Economics* 110, 520–545.

Primiceri, Giorgio E. (2005) "Time varying structural vector autoregressions and monetary policy" *The Review of Economic Studies* 72, 821–852.

Shephard, Neil (2013) "Martingale unobserved components models." Working paper, Department of Economics, University of Oxford.

Stock, James H., and Mark W. Watson (2001) "Vector autoregressions" *Journal of Economic Perspectives* 15, 101–115.

Ulrich, Maxim (2012) "Economic policy uncertainty and asset price volatility" Unpublished manuscript, Department of Economics, Columbia University.

Wu, Jing Cynthia, and Dora Fan Xia (2014) "Measuring the macroeconomic impact of monetary policy at the zero lower bound." Working paper, University of Chicago, Booth School of Business.

Appendix A Bond loadings

To ensure that the bond loadings satisfy the conditions $S_1 A = 0$ and $S_1 B' = I_G$ a priori for a given value of the parameters $(\delta_0, \phi^{\mathbb{Q}}, \Sigma_g^{\mathbb{Q}})$, we use Proposition 2 of Hamilton and Wu(2014). Their proposition provides an algorithm for imposing these conditions given an $G \times N$ matrix S_1. In our work, S_1 selects out the 3, 60, and 12 month maturities (the 2nd, 7th, and 3rd elements of the vector of yields y_t) and is equal to

$$S_1 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}, \qquad y_t' = \begin{pmatrix} y_t^{(1)} & y_t^{(3)} & y_t^{(12)} & y_t^{(24)} & y_t^{(36)} & y_t^{(48)} & y_t^{(60)} \end{pmatrix}.$$

Using this rotation, the yield factors are equal to the model implied yields $g_t = \left(\hat{y}_t^{(3)}, \hat{y}_t^{(60)}, \hat{y}_t^{(12)}\right)$ because we maintain the assumption that all yields are measured with error.

We make a slight modification to the algorithm of Hamilton and Wu(2014). We note that the vector of bond loadings A is a linear function of the scalar parameter δ_0. The loadings can therefore be decomposed as $A = A_{\delta_0} + C_{\delta_0} \times \delta_0$, where A_{δ_0} and C_{δ_0} are both $N \times 1$ vectors. In Appendix B, we use this decomposition in our state space model and place the parameter δ_0 in the state vector.

Let $\phi^{\mathbb{Q}}$ denote the vector of ordered eigenvalues $\phi^{\mathbb{Q}} = \left(\phi_1^{\mathbb{Q}}, \ldots, \phi_G^{\mathbb{Q}}\right)$. Define the function $\gamma_n(x) = \frac{1}{n}\sum_{j=0}^{n-1} x^j$. The algorithm for imposing these conditions is:

1. Calculate $\Phi_g^{\mathbb{Q}} = \left[K\left(\phi^{\mathbb{Q}}\right) S_1'\right]^{-1} \left[V\left(\phi^{\mathbb{Q}}\right)\right] \left[K\left(\phi^{\mathbb{Q}}\right) S_1'\right]$ and $\delta_{1,g} = \left[K\left(\phi^{\mathbb{Q}}\right) S_1'\right]^{-1} \iota_G$ where

$$\underset{(G\times N)}{K\left(\phi^{\mathbb{Q}}\right)} = \begin{pmatrix} \gamma_{n_1}\left(\phi_1^{\mathbb{Q}}\right) & \gamma_{n_2}\left(\phi_1^{\mathbb{Q}}\right) & \cdots & \gamma_{n_N}\left(\phi_1^{\mathbb{Q}}\right) \\ \gamma_{n_1}\left(\phi_2^{\mathbb{Q}}\right) & \gamma_{n_2}\left(\phi_2^{\mathbb{Q}}\right) & \cdots & \gamma_{n_N}\left(\phi_2^{\mathbb{Q}}\right) \\ \vdots & \vdots & \cdots & \vdots \\ \gamma_{n_1}\left(\phi_G^{\mathbb{Q}}\right) & \gamma_{n_2}\left(\phi_G^{\mathbb{Q}}\right) & \cdots & \gamma_{n_N}\left(\phi_G^{\mathbb{Q}}\right) \end{pmatrix} \qquad \underset{(G\times G)}{V\left(\phi^{\mathbb{Q}}\right)} = \begin{pmatrix} \phi_1^{\mathbb{Q}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \phi_G^{\mathbb{Q}} \end{pmatrix}$$

2. Given $\Phi_g^{\mathbb{Q}}$ and $\delta_{1,g}$, run the recursions

$$\begin{aligned} \bar{\psi}_n &= \bar{\psi}_{n-1} - \frac{1}{2}\bar{b}_{n-1}' \Sigma_g^{\mathbb{Q}} \Sigma_g^{\mathbb{Q}\prime} \bar{b}_{n-1} \\ \bar{\zeta}_n &= \bar{\zeta}_{n-1} - \bar{b}_{n-1} \\ \bar{b}_n &= \Phi_g^{\mathbb{Q}\prime} \bar{b}_{n-1} - \delta_{1,g} \end{aligned}$$

where the initial conditions are $\bar{\psi}_1 = 0, \bar{\zeta}_1 = 0_{G\times 1}$ and $\bar{b}_1 = -\delta_{1,g}$. Let $\zeta_n = \frac{1}{n}\bar{\zeta}_n$ and $\psi_n = \frac{1}{n}\bar{\psi}_n$. The matrix B is formed by stacking b_n' in order of maturity.

3. Calculate $\mu_{g,1}^{\mathbb{Q}} = -\left[S_1\zeta\left(\phi^{\mathbb{Q}}\right)\right]^{-1}\left[S_1\psi\left(\phi^{\mathbb{Q}},\Sigma_g^{\mathbb{Q}}\right)\right]$ and $\mu_{g,2}^{\mathbb{Q}} = -\left[S_1\zeta\left(\phi^{\mathbb{Q}}\right)\right]^{-1}\left[S_1\iota_G\right]$ where

$$\underset{N\times G}{\zeta\left(\phi^{\mathbb{Q}}\right)} = \begin{pmatrix} \zeta_{n_1}\left(\phi^{\mathbb{Q}}\right)' \\ \vdots \\ \zeta_{n_N}\left(\phi^{\mathbb{Q}}\right)' \end{pmatrix} \qquad \underset{N\times 1}{\psi\left(\phi^{\mathbb{Q}},\Sigma_g^{\mathbb{Q}}\right)} = \begin{pmatrix} \psi_{n_1}\left(\phi^{\mathbb{Q}},\Sigma_g^{\mathbb{Q}}\right) \\ \vdots \\ \psi_{n_N}\left(\phi^{\mathbb{Q}},\Sigma_g^{\mathbb{Q}}\right) \end{pmatrix}$$

4. Given $\mu_{g,1}^{\mathbb{Q}}$ and $\mu_{g,2}^{\mathbb{Q}}$, run the recursions

$$\begin{aligned} \bar{a}_{n,\delta_0} &= \bar{a}_{n-1,\delta_0} + \bar{b}_{n-1}'\mu_{g,1}^{\mathbb{Q}} + \frac{1}{2}\bar{b}_{n-1}'\Sigma_g^{\mathbb{Q}}\Sigma_g^{\mathbb{Q}\prime}\bar{b}_{n-1} \\ \bar{c}_{n,\delta_0} &= \bar{c}_{n-1,\delta_0} + \bar{b}_{n-1}'\mu_{g,2}^{\mathbb{Q}} \end{aligned}$$

where the initial conditions are $\bar{a}_{1,\delta_0} = 0$ and $\bar{c}_{1,\delta_0} = 0$. Let $a_{n,\delta_0} = -\frac{1}{n}\bar{a}_{n,\delta_0}$ and $c_{n,\delta_0} = -\frac{1}{n}\bar{c}_{n,\delta_0}$. The vectors A_{δ_0} and C_{δ_0} are formed by stacking a_{n,δ_0},and c_{n,δ_0} in order of maturity.

Forming the matrices in this way automatically satisfies the conditions $S_1B' = I_G$ and $S_1A = 0$. They also imply a simple decomposition of $A = A_{\delta_0} + C_{\delta_0}\times\delta_0$, which is useful for placing the model in state space form. Also, note that $\mu_g^{\mathbb{Q}} = \mu_{g,1}^{\mathbb{Q}} + \mu_{g,2}^{\mathbb{Q}}\times\delta_0$.

Appendix B MCMC and particle filtering algorithms

Appendix B.1 MCMC algorithm

In the appendix, we use the notation $x_{t:t+k} = (x_t,\ldots,x_{t+k})$ to denote a sequence of variables from time t to time $t+k$. We formulate the model in terms of deviations from the means, with $\bar{\mu}_x$ denoting the unconditional mean of the variable x_t and $\bar{x}_t = x_t - \bar{\mu}_x$ being the demeaned variable.

Our Gibbs sampling algorithm iterates between three basic steps: (i) drawing the latent yield factors $g_{1:T}$ conditional on the volatilities $h_{0:T}$ and parameters θ; (ii) drawing the volatilities $h_{0:T}$ conditional on $g_{1:T}$ and θ; (iii) and then drawing the parameters of the model θ given the state variables. The MCMC algorithm is designed to minimize the amount that we condition on the latent variables by using the Kalman filter to marginalize over the state variables. We will use two different state space representations as described in Subsection 3.1. We write the model using the following state space form

$$Y_t = Z_t x_t + d_t + \eta_t^* \qquad \eta_t^* \sim \mathrm{N}\left(0,H_t\right), \tag{B.1}$$

$$x_{t+1} = T_t x_t + c_t + R_t\varepsilon_{t+1}^* \qquad \varepsilon_{t+1}^* \sim \mathrm{N}\left(0,Q_t\right), \tag{B.2}$$

Appendix B.1.1 State space form I conditional on $h_{0:T}$

The intercept A in (11) is a linear function of δ_0, and we can decompose it into $A = A_{\delta_0} + C_{\delta_0} \times \delta_0$. For better mixing behavior of the MCMC algorithm, we place the unconditional means $\bar{\mu}_m, \bar{\mu}_g$ and δ_0 in the state vector: we can draw them jointly with the yield factors $\bar{g}_{1:T}$ using simulation smoothing algorithms (forward-filtering backward sampling). We can also marginalize over these parameters when drawing other parameters of the model. We fit the model into the state space form (B.1) and (B.2) by defining the state space matrices as

$$Y_t = \begin{pmatrix} m_t \\ y_t \\ \bar{h}_t \end{pmatrix}, \quad Z_t = \begin{pmatrix} I & 0 & 0 & I & 0 & 0 \\ 0 & B & 0 & 0 & B & C_{\delta_0} \\ 0 & 0 & I & 0 & 0 & 0 \end{pmatrix}, \quad d_t = \begin{pmatrix} 0 \\ A_{\delta_0} \\ 0 \end{pmatrix}, \quad H_t = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \Omega & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$x_t = \begin{pmatrix} \bar{m}_t \\ \bar{g}_t \\ \bar{h}_t \\ \bar{\mu}_m \\ \bar{\mu}_g \\ \delta_0 \end{pmatrix}, \quad T_t = \begin{pmatrix} \Phi_m & \Phi_{mg} & \Phi_{mh} & 0 & 0 & 0 \\ \Phi_{gm} & \Phi_g & \Phi_{gh} & 0 & 0 & 0 \\ 0 & 0 & \Phi_h & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & 0 & I \end{pmatrix}, \quad R_t = \begin{pmatrix} \Sigma_m & 0 & 0 \\ \Sigma_{gm} & \Sigma_g D_t & 0 \\ \Sigma_{hm} & \Sigma_{hg} D_t & \Sigma_h \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Q_t = I,$$

and $c_t = 0$. The priors for the parameters are $\delta_0 \sim \mathrm{N}\left(\underline{\delta_0}, V_{\delta_0}\right)$, $\bar{\mu}_g \sim \mathrm{N}\left(\underline{\bar{\mu}_g}, V_{\bar{\mu}_g}\right)$, and $\bar{\mu}_m \sim \mathrm{N}\left(\underline{\bar{\mu}_m}, V_{\bar{\mu}_m}\right)$. The initial conditions for $x_1 \sim \mathbf{N}\left(\underline{x}_1, P_1\right)$ are

$$\underline{x}_1 = \begin{pmatrix} \Phi_m \bar{m}_0 + \Phi_{mh} \bar{h}_0 \\ \Phi_m \bar{m}_0 + \Phi_{gh} \bar{h}_0 \\ \bar{h}_0 \\ \underline{\bar{\mu}_m} \\ \underline{\bar{\mu}_g} \\ \underline{\delta_0} \end{pmatrix}, \qquad R_1 = \begin{pmatrix} \Sigma_m & 0 & 0 & 0 & 0 & 0 \\ \Sigma_{gm} & \Sigma_g D_0 & 0 & 0 & 0 & 0 \\ \Sigma_{hm} & \Sigma_{hg} D_0 & \Sigma_h & 0 & 0 & 0 \\ 0 & 0 & 0 & V_{\bar{\mu}_m}^{\frac{1}{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & V_{\bar{\mu}_g}^{\frac{1}{2}} & 0 \\ 0 & 0 & 0 & 0 & 0 & V_{\delta_0}^{\frac{1}{2}} \end{pmatrix},$$

where $\underline{x}_1$ and $P_1 = R_1 Q_1 R_1'$ are observable.

Appendix B.1.2 State space form II conditional on $g_{1:T}$

First, we stack the dynamics of $\bar{m}_t$, $\bar{g}_t$, and $\bar{h}_t$ implied by (1), (2), and (4) together as a VAR(1). Next, we multiply both sides of the VAR by

$$\begin{pmatrix} \Sigma_m & 0 & 0 \\ 0 & I_G & 0 \\ 0 & 0 & \Sigma_h \end{pmatrix} \begin{pmatrix} \Sigma_m & 0 & 0 \\ \Sigma_{gm} & \Sigma_g & 0 \\ \Sigma_{hm} & \Sigma_{hg} & \Sigma_h \end{pmatrix}^{-1}.$$

This rotates the model so that the shocks $(\varepsilon_{mt}, \varepsilon_{gt}, \varepsilon_{ht})$ are only present in their own respective equations. The model can be written as

$$\begin{aligned}
\bar{m}_t &= \Phi_m \bar{m}_{t-1} + \Phi_{mg} \bar{g}_{t-1} + \Phi_h \bar{h}_{t-1} + \Sigma_m \varepsilon_{mt} \\
\Sigma_g^{-1} \bar{g}_t &= \Sigma_g^{-1} \Sigma_{gm} \Sigma_m^{-1} \bar{m}_t + \Sigma_g^{-1} \left[\Phi_{gm} - \Sigma_{gm} \Sigma_m^{-1} \Phi_m \right] \bar{m}_{t-1} + \Sigma_g^{-1} \left[\Phi_g - \Sigma_{gm} \Sigma_m^{-1} \Phi_{mg} \right] \bar{g}_{t-1} + D_{t-1} \varepsilon_{gt} \\
\bar{h}_t &= \left[\Sigma_{hm} - \Sigma_{hg} \Sigma_g^{-1} \Sigma_{gm} \right] \Sigma_m^{-1} \bar{m}_t + \Sigma_{hg} \Sigma_g^{-1} \bar{g}_t - \left[\left(\Sigma_{hm} - \Sigma_{hg} \Sigma_g^{-1} \Sigma_{gm} \right) \Sigma_m^{-1} \Phi_m + \Sigma_{hg} \Sigma_g^{-1} \Phi_{gm} \right] \bar{m}_{t-1} \\
&\quad - \left[\left(\Sigma_{hm} - \Sigma_{hg} \Sigma_g^{-1} \Sigma_{gm} \right) \Phi_{mg} + \Sigma_{hg} \Sigma_g^{-1} \Phi_g \right] \bar{g}_{t-1} + \left[\Phi_h - \Sigma_{hm} \Sigma_m^{-1} \Phi_{mh} \right] \bar{h}_{t-1} + \Sigma_h \varepsilon_{ht}
\end{aligned}$$

where we have imposed the restriction that $\Phi_{gh} = \Sigma_{gm} \Sigma_m^{-1} \Phi_{mh}$. This restriction means that $\bar{g}_t$ does not depend on $\bar{h}_{t-1}$. Consequently, we can use the approach of Kim, Shephard, and Chib(1998) to reformulate the stochastic volatility model for $\bar{g}_t$ as an (approximate) linear, Gaussian model. Define the following variables

$$\begin{aligned}
g_t^* &= \log\left(\tilde{g}_t \odot \tilde{g}_t \right) \\
\tilde{g}_t &= \Sigma_g^{-1} \left(\bar{g}_t - \Sigma_{gm} \Sigma_m^{-1} \bar{m}_t - \left[\Phi_{gm} - \Sigma_{gm} \Sigma_m^{-1} \Phi_m \right] \bar{m}_{t-1} - \left[\Phi_g - \Sigma_{gm} \Sigma_m^{-1} \Phi_{mg} \right] \bar{g}_{t-1} \right)
\end{aligned}$$

By squaring both sides of $\tilde{g}_{it}$ for $i = 1, \ldots, G$ equation-by-equation and then taking logarithms, we can write the model as

$$g_t^* = \Gamma_0 + \Gamma_1 h_{t-1} + \varepsilon_{gt}^*$$

where the shocks are $\varepsilon_{gt}^* = \log\left(\varepsilon_{gt} \odot \varepsilon_{gt} \right)$. Let $s_t = (s_{1t}, \ldots, s_{G,t})$ denote a $G \times 1$ vector of discrete variables indexing each component of the normal mixture $p(\varepsilon_{i,gt}^*) \approx \sum_{j=1}^{10} \pi_j N(e_j, \sigma_j^2)$, where $\mathbb{E}\left[\varepsilon_{i,gt}^* | s_{it} = j \right] = e_j$ and $\mathbb{V}\left[\varepsilon_{i,gt}^* | s_{it} = j \right] = \sigma_j^2$. In practice, we use the means and variances from the 10 component mixture provided by Omori, Chib, Shephard, and Nakajima(2007). However, we do not use their approach to estimate

the leverage parameters because our model is naturally written in terms of uncorrelated shocks.

The state space matrices are

$$Y_t = \begin{pmatrix} \bar{m}_t \\ g_t^* \end{pmatrix}, \quad Z_t = \begin{pmatrix} \Phi_{mh} & 0 \\ \Gamma_1 & 0 \end{pmatrix}, \quad d_t = \begin{pmatrix} \Phi_m \bar{m}_{t-1} + \Phi_{mg} \bar{g}_{t-1} \\ \Gamma_0 + \Gamma_1 \bar{\mu}_h + e_{s_t} \end{pmatrix}, \quad H_t = \begin{pmatrix} \Sigma_m \Sigma_m' & 0 \\ 0 & \sigma_{s_t}^2 \end{pmatrix},$$

$$x_t = \begin{pmatrix} \bar{h}_{t-1} \\ \text{vec}\left(\Sigma_{hg}\right) \end{pmatrix}, \quad T_t = \begin{pmatrix} \Phi_h - \Sigma_{hm} \Sigma_m^{-1} \Phi_{mh} & \left(\check{g}_t' \otimes I\right) \\ 0 & I \end{pmatrix}, \quad R_t = \begin{pmatrix} \Sigma_h \\ 0 \end{pmatrix}, \quad c_t = \begin{pmatrix} \check{m}_t \\ 0 \end{pmatrix},$$

and with $Q_t = I$. The priors for the parameters are $\text{vec}\left(\Sigma_{hg}\right) \sim \text{N}\left(\underline{\Sigma_{hg}}, V_{\Sigma_{hg}}\right)$. The initial conditions for $x_1 \sim \mathbf{N}\left(\underline{x}_1, P_1\right)$ are

$$\underline{x}_1 = \begin{pmatrix} 0 \\ \underline{\Sigma_{hg}} \end{pmatrix}, \qquad P_1 = \begin{pmatrix} \Sigma_h \Sigma_h' & 0 \\ 0 & V_{\Sigma_{hg}} \end{pmatrix},$$

where $\bar{m}_0$ is observable at time $t = 0$ and we define

$$\begin{aligned} \check{m}_t &= \Sigma_{hm} \left[\Sigma_m^{-1} \bar{m}_t - \Sigma_m^{-1} \Phi_m \bar{m}_{t-1} - \Phi_{mg} \bar{g}_{t-1}\right] \\ \check{g}_t &= \Sigma_g^{-1} \left[\bar{g}_t - \Sigma_{gm} \Sigma_m^{-1} \left(\bar{m}_t - \Phi_m \bar{m}_{t-1}\right) - \Phi_{gm} \bar{m}_{t-1} + \Sigma_{gm} \Phi_{mg} \bar{g}_{t-1} - \Phi_g \bar{g}_{t-1}\right] \end{aligned}$$

By placing $\text{vec}\left(\Sigma_{hg}\right)$ in the state vector, we can draw these parameters jointly with $\bar{h}_{0:T}$. Note that the unconditional mean of volatility $\bar{\mu}_h$ also enters the bond loadings A in (11) and cannot be placed in the state vector.

Appendix B.1.3 The IMH algorithm

In our MCMC algorithm, we draw as many parameters as possible without conditioning on the state variables and other parameters. In the algorithm in Appendix B.1.4, we will repeatedly apply the independence Metropolis Hastings (IMH) algorithm along the lines of Chib and Greenberg(1994), in a combination with Kalman filter to marginalize out the state variables based on the state space representations above. Here is how it works. Suppose we separate the parameter vector $\theta = (\psi, \psi^-)$ and we want to draw a subset of the parameters ψ conditional on the remaining parameters ψ^-.

- Maximize the log-posterior $p\left(\psi | Y_{1:T}, \psi^-\right) \propto p\left(Y_{1:T} | \psi, \psi^-\right) p\left(\psi\right)$, where the likelihood is computed using Kalman filter. Let $\hat{\psi}$ be the posterior mode and H_ψ^{-1} be the inverse Hessian at the mode.
- Draw a proposal $\psi^* \sim t_5\left(\hat{\psi}, H_\psi^{-1}, \right)$ from a Student's t distribution with mean $\hat{\psi}$, scale matrix H_ψ^{-1}, and 5 degrees of freedom.

- The proposal ψ^* is accepted with probability $\alpha = \frac{p\left(Y_{1:T}|\psi^*,\psi^-\right)p(\psi^*)q\left(\psi^{(j-1)}\right)}{p\left(Y_{1:T}|\psi^{(j-1)},\psi^-\right)p\left(\psi^{(j-1)}\right)q(\psi^*)}$.

A similar algorithm has been used by Chib and Ergashev(2009) for Gaussian ATSMs with macroeconomic factors but no stochastic volatility.

Appendix B.1.4 MCMC algorithm

We use the notation $\theta^{(-)}$ to denote all the remaining parameters in θ other than the parameters being drawn in that step. Our MCMC algorithm proceeds as follows:

1. **Draw** $\Sigma_m, \Sigma_g, \Sigma_{gm}$: Conditional on $h_{0:T}$ and the remaining parameters of the model, write the model as in state space form I. Draw parameters listed below using the IMH algorithm as explained in Appendix B.1.3.
   - vech (Σ_m)
   - free parameters in Σ_g: note the diagonal elements of Σ_g are fixed.
   - vec (Σ_{gm})
2. **Draw** Ω **(Option #1):** Conditional on $h_{0:T}$ and the remaining parameters of the model $\theta^{(-)}$, write the model as in state space form 1. Draw the elements of Ω using the IMH algorithm as explained above. In practice, not all the parameters can be drawn simultaneously and we randomly block them into smaller groups at each iteration.
3. **Draw** $\left(\bar{g}_{1:T}, \bar{\mu}_g, \bar{\mu}_m, \delta_0, \phi^{\mathbb{Q}}\right)$ **jointly in one block.**
   - **Draw** $\phi^{\mathbb{Q}}$ **from** $p\left(\phi^{\mathbb{Q}}|Y_{1:T}, \theta^{(-)}\right)$**:** Conditional on $h_{0:T}$ and the remaining parameters of the model $\theta^{(-)}$, write the model in state space form I. Draw the elements of $\phi^{\mathbb{Q}}$ using the IMH algorithm as explained above.
   - **Draw** $(\bar{g}_{1:T}, \bar{\mu}_g, \bar{\mu}_m, \delta_0)$ **jointly from** $p\left(\bar{g}_{1:T}, \bar{\mu}_g, \bar{\mu}_m, \delta_0|Y_{1:T}, \theta^{(-)}, \phi^{\mathbb{Q}}\right)$**:** Conditional on $\phi^{\mathbb{Q}}$, draw $(\bar{g}_{1:T}, \bar{\mu}_g, \bar{\mu}_m, \delta_0)$ using the simulation smoother of Durbin and Koopman(2002).
4. **Draw** Ω **(Option #2):** Conditional on the draw of $g_{1:T}$, we can draw from the full conditional distribution of Ω when it has an inverse Wishart prior. This comes from the regression model $y_t = A + Bg_t + \eta_t$ where A and B are the bond loadings.
5. **Draw** $(\Phi_m, \Phi_{mg}, \Phi_{gm}, \Phi_g)$ **conditional on** $\bar{m}_{1:T}, \bar{g}_{1:T}, h_{0:T}$**:** In this step, we write the model as a vector autoregression with a known form of conditional heteroskedasticity (we are conditioning on $h_{0:T}$). Under a multivariate normal prior, the full conditional posterior for $(\Phi_m, \Phi_{mg}, \Phi_{gm}, \Phi_g)$ is multivariate normal. The draw is standard for a Bayesian VAR; see, e.g. Del Negro and Schorfheide(2011).

6. **Draw the mixture indicators s_{it} for $i = 1, \ldots, G$ and $t = 1, \ldots, T$:** Draw mixture indicators s_{it} to approximate the distribution of the errors $p\left(\varepsilon^*_{i,gt}\right)$. This step uses the 10 component mixture from Omori, Chib, Shephard, and Nakajima(2007).

7. **Draw $(\Gamma_1, \Phi_{mh}, \Phi_h, \Sigma_h)$:** Using state space form II, draw the following parameters conditional on the mixture indicators $s_{1:T}$. We use the IHM algorithm as described in Appendix B.1.3..

   - Draw the free parameters in Γ_1 if $H < G$.
   - $\text{vec}\left(\Sigma_{hm}\right)$.
   - Φ_{mh}, Φ_h and $\text{vech}\left(\Sigma_h\right)$: If the number of volatility factors $H \geq 2$, we split the total number of parameters into sub-groups of even size (roughly 5 parameters per MH step).

8. **Draw from the joint distribution $\left(\bar{h}_{0:T}, \Sigma_{hg}, \bar{\mu}_h, \Gamma_0\right)$:**

   - **Draw $\bar{\mu}_h$ and Γ_0 from $p\left(\bar{\mu}_h, \Gamma_0 | Y_{1:T}, \theta^{(-)}\right)$.** Recall that the parameters $\bar{\mu}_h$ and Γ_0 enter the bond loadings A in (11) from our assumption that the covariance matrix under $\mathbb{Q}$ is equal to the (long-run) covariance matrix under $\mathbb{P}$. This adds an additional term to the target distribution. The Metropolis-Hastings acceptance probability is

$$\alpha = \frac{p\left(Y_{1:T} | \bar{\mu}^*_h, \Gamma^*_0, \theta^{(-)}\right) p\left(y_{1:T} | g_{1:T}, \bar{\mu}^*_h, \Gamma^*_0, \theta^{(-)}\right) p\left(\bar{\mu}^*_h, \Gamma^*_0\right) q\left(\psi^{(j-1)}\right)}{p\left(Y_{1:T} | \bar{\mu}^{(j-1)}_h, \Gamma^{(j-1)}_0, \theta^{(-)}\right) p\left(y_{1:T} | g_{1:T}, \bar{\mu}^{(j-1)}_h, \Gamma^{(j-1)}_0, \theta^{(-)}\right) p\left(\bar{\mu}^{(j-1)}_h, \Gamma^{(j-1)}_0\right) q\left(\bar{\mu}^*_h, \Gamma^*_0\right)}$$

     where $p\left(Y_{1:T} | \bar{\mu}_h, \Gamma_0, \theta^{(-)}\right)$ is the likelihood of the state space form II calculated by the Kalman filter and

$$\log p\left(y_{1:T} | g_{1:T}, \bar{\mu}_h, \Gamma_0, \theta^{(-)}\right) \propto -0.5 * (\nu_\omega + T) * \log|S_\omega + \sum_{t=1}^{T} (y_t - A - Bg_t)(y_t - A - Bg_t)'|.$$

     Here, ν_ω and S_ω are the prior hyperparameters of Ω, which we have integrated out of the MH acceptance ratio.

   - **Draw $\left(\bar{h}_{0:T}, \Sigma_{hg}\right)$ jointly from $p\left(\bar{h}_{0:T}, \Sigma_{hg} | Y_{1:T}, \theta^{(-)}, \bar{\mu}_h, \Gamma_0\right)$:** Using the state space form II, draw $\bar{h}_{0:T}$ and Σ_{hg} using the simulation smoother of Durbin and Koopman(2002).

Appendix B.2 Particle filter

The particle filter we implement is the mixture Kalman filter of Chen and Liu(2000). For a survey of particle filtering see Creal(2012). Let $x_{t|t-1}$ denote the conditional mean and $P_{t|t-1}$ the conditional covariance matrix

of the one-step ahead predictive distribution $p(x_t|Y_{1:t-1}, h_{0:t-1}; \theta)$ of a conditionally linear, Gaussian state space model. Similarly, let $x_{t|t}$ denote the conditional mean and $P_{t|t}$ the conditional covariance matrix of the filtering distribution $p(x_t|Y_{1:t}, h_{0:t}; \theta)$. Conditional on the volatilities $h_{0:T}$, these quantities can be calculated by the Kalman filter.

We define the parameters in the state space form (B.1) and (B.2) as follows

$$Y_t = \begin{pmatrix} m_t \\ y_t \end{pmatrix}, \quad Z_t = \begin{pmatrix} I & 0 \\ 0 & B \end{pmatrix}, \quad d_t = \begin{pmatrix} \bar{\mu}_m \\ A + B\bar{\mu}_g \end{pmatrix}, \quad H_t = \begin{pmatrix} 0 & 0 \\ 0 & \Omega \end{pmatrix},$$

$$x_t = \begin{pmatrix} \bar{m}_t \\ \bar{g}_t \end{pmatrix}, \quad T_t = \begin{pmatrix} \Phi_m & \Phi_{mg} \\ \Phi_{gm} & \Phi_g \end{pmatrix}, \quad R_t = \begin{pmatrix} \Sigma_m & 0 \\ \Sigma_{gm} & \Sigma_g D_t \end{pmatrix}, \quad c_t = \begin{pmatrix} \Phi_{mh}\bar{h}_t \\ \Phi_{gh}\bar{h}_t \end{pmatrix},$$

and where $Q_t = I$. Let $N_{mn} = M + N$ denote the dimension of the observation vector Y_t and J the number of particles. Within the particle filter, we use the residual resampling algorithm of Liu and Chen(1998). The particle filter then proceeds as follows:

At $t = 0$, for $i = 1, \ldots, J$, set $w_0^{(i)} = \frac{1}{J}$ and

- Draw $\bar{h}_0^{(i)} \sim p\left(\bar{h}_0; \theta\right)$ and calculate $\left(D_0 D_0'\right)^{(i)} = \text{diag}\left[\exp\left(\Gamma_0 + \Gamma_1\bar{\mu}_h + \Gamma_1\bar{h}_0^{(i)}\right)\right]$.
- Set $x_{1|0}^{(i)} = \begin{pmatrix} \Phi_m\bar{m}_0 + \Phi_{mh}\bar{h}_0^{(i)} \\ \Phi_{gm}\bar{m}_0 + \Phi_{gh}\bar{h}_0^{(i)} \end{pmatrix}$, $P_{1|0}^{(i)} = \begin{pmatrix} \Sigma_m\Sigma_m' & \Sigma_m\Sigma_{gm}' \\ \Sigma_{gm}\Sigma_m' & \Sigma_{gm}\Sigma_{gm}' + \Sigma_g\left(D_0 D_0'\right)^{(i)}\Sigma_g' \end{pmatrix}$,
- Set $\ell_0 = 0$.

For $t = 1, \ldots, T$ do:

STEP 1: For $i = 1, \ldots, J$:

- Calculate $c_t^{(i)}$ and $R_t^{(i)}$ using $\bar{h}_t^{(i)}$.

- Run the Kalman filter:

$$\begin{aligned} v_t^{(i)} &= Y_t - Z_t x_{t|t-1}^{(i)} - d_t \\ F_t^{(i)} &= Z_t P_{t|t-1}^{(i)} Z_t' + H_t \\ K_t^{(i)} &= P_{t|t-1}^{(i)} Z_t' \left(F_t^{(i)}\right)^{-1} \\ x_{t|t}^{(i)} &= x_{t|t-1}^{(i)} + K_t^{(i)} v_t^{(i)} \\ P_{t|t}^{(i)} &= P_{t|t-1}^{(i)} - K_t^{(i)} Z_t P_{t|t-1}^{(i)} \\ x_{t+1|t}^{(i)} &= T_t x_{t|t}^{(i)} + c_t^{(i)} \\ P_{t+1|t}^{(i)} &= T_t P_{t|t}^{(i)} T_t' + R_t^{(i)} Q_t R_t^{(i)\prime} \end{aligned}$$

- Draw from the transition density: $\bar{h}_{t+1}^{(i)} \sim p(\bar{h}_{t+1}|\bar{m}_{t+1}, \bar{g}_{t+1}^{(i)}, \bar{h}_t^{(i)}, \bar{m}_t, \bar{g}_t^{(i)}; \theta)$ given by:

$$\bar{h}_{t+1} = \Phi_h \bar{h}_t + \Sigma_{hm}\varepsilon_{m,t+1} + \Sigma_{hg} D_t \varepsilon_{g,t+1} + \Sigma_h \varepsilon_{h,t+1} \qquad \varepsilon_{h,t+1} \sim \mathrm{N}(0, I)$$

- Calculate the weight: $\log\left(w_t^{(i)}\right) = \log\left(\hat{w}_{t-1}^{(i)}\right) - 0.5 N_{nm} \log(2\pi) - 0.5 \log|F_t^{(i)}| - \frac{1}{2} v_t^{(i)\prime} \left(F_t^{(i)}\right)^{-1} v_t^{(i)}$.

STEP 2: Calculate an estimate of the log-likelihood: $\ell_t = \ell_{t-1} + \log\left(\sum_{i=1}^J w_t^{(i)}\right)$.

STEP 3: For $i = 1, \ldots, J$, calculate the normalized importance weights: $\hat{w}_t^{(i)} = \frac{w_t^{(i)}}{\sum_{j=1}^J w_t^{(j)}}$.

STEP 4: Calculate the effective sample size $E_t = \frac{1}{\sum_{j=1}^J \left(\hat{w}_t^{(j)}\right)^2}$.

STEP 5: If $E_t < 0.5J$, resample $\left\{x_{t+1|t}^{(i)}, P_{t+1|t}^{(i)}, \bar{h}_{t+1}^{(i)}\right\}_{i=1}^J$ with probabilities $\hat{w}_t^{(i)}$ and set $\hat{w}_t^{(i)} = \frac{1}{J}$.

STEP 6: Increment time and return to STEP 1.

Appendix B.3 Impulse response functions

We summarize (1) - (4) by

$$x_t = \mu + \Phi x_{t-1} + \Sigma_{t-1}\varepsilon_t.$$

where

$$x_t = \begin{pmatrix} m_t \\ g_t \\ h_t \end{pmatrix} \quad \mu = \begin{pmatrix} \mu_m \\ \mu_g \\ \mu_h \end{pmatrix}, \quad \Phi = \begin{pmatrix} \Phi_m & \Phi_{mg} & \Phi_{mh} \\ \Phi_{gm} & \Phi_g & \Phi_{gh} \\ 0 & 0 & \Phi_h \end{pmatrix}, \quad \Sigma_{t-1} = \begin{pmatrix} \Sigma_m & 0 & 0 \\ \Sigma_{gm} & \Sigma_g D_{t-1} & 0 \\ \Sigma_{hm} & \Sigma_{hg} D_{t-1} & \Sigma_h \end{pmatrix}.$$

For each draw $\left\{\theta^{(k)}, x_{0:T}^{(k)}\right\}$ in the MCMC algorithm where $x_t^{(k)} = \left(m_t^{(k)\prime}, g_t^{(k)\prime}, h_t^{(k)\prime}\right)'$, we calculate the implied value of the shocks $\varepsilon_t^{(k)} = \left(\varepsilon_{mt}^{(k)}, \varepsilon_{gt}^{(k)}, \varepsilon_{ht}^{(k)}\right)$ for $t = 1, \ldots, T$. Note that m_t is observable and does not change from one draw to another.

Appendix B.3.1 Time-varying impulse responses

Let $\tilde{x}_t^{(k)} = \left(\tilde{m}_t^{(k)\prime}, \tilde{g}_t^{(k)\prime}, \tilde{h}_t^{(k)\prime}\right)'$ denote the implied value of the state vector for the k-th draw assuming that the j-th shock at the time of impact s is given by $\tilde{\varepsilon}_{js}^{(k)} = \varepsilon_{js}^{(k)} + 1$, and keeping all other shocks at their values implied by the data. Then, given an initial condition of the state vector $\tilde{x}_{s-1}^{(k)} = x_{s-1}^{(k)}$, for a τ period impulse response, we iterate forward on the dynamics of the vector autoregression

$$\tilde{x}_t^{(k)} = \mu^{(k)} + \Phi^{(k)} \tilde{x}_{t-1}^{(k)} + \Sigma_{t-1}^{(k)} \tilde{\varepsilon}_t^{(k)}, \qquad t = s, \ldots, s+\tau.$$

The impulse response at time s, for a horizon τ, variable i, shock j, and draw k is defined as

$$\Upsilon_{s,ij,\tau}^{(k)} = \tilde{x}_{i,s+\tau}^{(k)} - x_{i,s+\tau}^{(k)}.$$

We then calculate the median and quantiles of $\Upsilon_{s,ij,\tau}^{(k)}$ across the draws $k = 1, \ldots, M$.

Appendix B.3.2 Constant impulse responses

Let $\dot{x}_t^{(k)} = \left(\dot{m}_t^{(k)\prime}, \dot{g}_t^{(k)\prime}, \dot{h}_t^{(k)\prime}\right)'$ denote the implied value of the state vector for the k-th draw assuming that the j-th shock at the time of impact s is given by $\dot{\varepsilon}_{js}^{(k)} = \varepsilon_{js}^{(k)} + 1$, $\varepsilon_{g,t}^{(k)} = 0$ for all t, and keeping all other shocks at their values implied by the data. And, let $\ddot{x}_t^{(k)} = \left(\ddot{m}_t^{(k)\prime}, \ddot{g}_t^{(k)\prime}, \ddot{h}_t^{(k)\prime}\right)'$ denote the implied value of the state vector for the k-th draw assuming that the j-th shock at the time of impact s is given by $\ddot{\varepsilon}_{js}^{(k)} = \varepsilon_{js}^{(k)}$, $\ddot{\varepsilon}_{gt}^{(k)} = 0$ for all t and keeping all other shocks at their values implied by the data.[7] Then, given an initial condition of the state vector $\dot{x}_{s-1}^{k} = \ddot{x}_{s-1}^{k} = x_{s-1}^{k}$, we iterate forward on the dynamics of the vector autoregression

$$\begin{aligned}
\dot{x}_t^{(k)} &= \mu^{(k)} + \Phi^{(k)} \dot{x}_{t-1}^{(k)} + \Sigma_{t-1}^{(k)} \dot{\varepsilon}_t^{(k)}, \qquad t = s, \ldots, s+\tau, \\
\ddot{x}_t^{(k)} &= \mu^{(k)} + \Phi^{(k)} \ddot{x}_{t-1}^{(k)} + \Sigma_{t-1}^{(k)} \ddot{\varepsilon}_t^{(k)}, \qquad t = s, \ldots, s+\tau.
\end{aligned}$$

The impulse response for a horizon τ, variable i, shock j, and draw k is defined as

$$\Psi_{ij,\tau}^{(k)} = \dot{x}_{i,s+\tau}^{(k)} - \ddot{x}_{i,s+\tau}^{(k)}$$

[7] Due to the linear nature, what the other shocks except for ε_{gt} are evaluated at will not change the value of the constant impulse response.

We then calculate the median and quantiles of $\Psi_{ij,\tau}$ across the draws $k = 1, \ldots, M$.

Appendix C Prior distributions

We use proper priors for all parameters of the model but select the hyperparameters of each distribution to be non-informative. Throughout this discussion, a normal distribution is defined as $x \sim \mathrm{N}(\mu_x, V_x)$, where μ_x is the mean and V_x is the covariance matrix. The inverse Wishart distribution $X \sim \mathrm{InvWishart}(\nu, S)$ is defined for a random $k \times k$ matrix X such that $\mathbb{E}[X] = \frac{S}{\nu - k - 1}$. Recall that we have divided all observed variables by 1200 and this is reflected in the scale of the hyperparameters. Let ι_k denote a $k \times 1$ vector of ones.

- $\delta_0 \sim \mathrm{N}\left(0.001, 5 \times 10^{-5}\right)$.
- The eigenvalues $\phi^{\mathbb{Q}}$ of $\Phi_g^{\mathbb{Q}}$ each have a beta distribution $\phi_i^{\mathbb{Q}} \sim \mathrm{beta}(a_i, b_i)$ for $i = 1, \ldots G$. For the three factor model, we set $a_1 = 100, b_1 = 1, a_2 = 200, b_2 = 10, a_3 = 100, b_3 = 60$.
- $\bar{\mu}_m \sim \mathrm{N}\left((4,6)'/1200, I_M \times (4/1200)^2\right)$.
- $\bar{\mu}_g \sim \mathrm{N}\left((4.8 \times \iota_{G\times 1})/1200, I_G \times 5 \times 10^{-6}\right)$.
- $\bar{\mu}_h \sim \mathrm{N}\left(0_{H\times 1}, I_H \times 10^{-6}\right)$.
- $\mathrm{vec}(\Sigma_{mg}) \sim \mathrm{N}\left(0, I_{M\times G} \times 10^{-6}\right)$, $\mathrm{vec}(\Sigma_{hm}) \sim \mathrm{N}\left(0, I_{H\times M} \times 10^{-6}\right)$,$\mathrm{vec}(\Sigma_{hg}) \sim \mathrm{N}\left(0, I_{H\times G} \times 10^{-6}\right)$
- For the autoregressive parameters, the prior is a conditional prior $p(\Phi_m, \Phi_{mg}, \Phi_g, \Phi_{gm}, \Phi_{mh}, \Phi_h | \Sigma_m, \Sigma_{gm})$. Define the transition matrix Φ as

$$\Phi = \begin{pmatrix} \Phi_m & \Phi_{mg} & \Phi_{mh} \\ \Phi_{gm} & \Phi_g & \Sigma_{gm}\Sigma_m^{-1}\Phi_{mh} \\ 0 & 0 & \Phi_h \end{pmatrix}$$

  where have substituted the restriction $\Phi_{gh} = \Sigma_{gm}\Sigma_m^{-1}\Phi_{mh}$. The distribution of the free parameters of $\mathrm{vec}(\Phi)$ is truncated normal. We truncate it to the stationarity region. The distribution of each sub-matrix is $\mathrm{vec}(\Phi_{mg}) \sim \mathrm{N}\left(0, I_{M\times G} \times 10^{-1}\right)$, $\mathrm{vec}(\Phi_{gm}) \sim \mathrm{N}\left(0, I_{M\times G} \times 10^{-1}\right)$, $\mathrm{vec}(\Phi_{mh}) \sim \mathrm{N}\left(0, I_{M\times H} \times 10^{-1}\right)$, $\mathrm{vec}(\Phi_m) \sim \mathrm{N}\left(\mathrm{vec}(I_M \times 0.95), I_{M^2} \times 10^{-1}\right)$, $\mathrm{vec}(\Phi_g) \sim \mathrm{N}\left(\mathrm{vec}(I_G \times 0.95), I_{G^2} \times 10^{-1}\right)$, $\mathrm{vec}(\Phi_h) \sim \mathrm{N}\left(\mathrm{vec}(I_H \times 0.985), I_{H^2} \times 10^{-1}\right)$.
- $\Sigma_m\Sigma_m' \sim \mathrm{InvWishart}(\nu_{\sigma_m}, S_{\sigma_m})$ with $\nu_{\sigma_m} = M + 2$ and $S_{\sigma_m} = (\nu_{\sigma_m} - M - 1) \times I_M \times 10^{-6}$.
- $\Sigma_h\Sigma_h' \sim \mathrm{InvWishart}\left(\nu_{\sigma_g}, S_{\sigma_g}\right)$ with $\nu_{\sigma_g} = H + 2$ and $S_{\sigma_g} = (\nu_{\sigma_g} - H - 1) \times I_H \times 10^{-6}$.

- For the models with $H = 1, 2$ factors, we estimate parameters in the matrix Γ_0. For $H = 1$, $\Gamma_0' = \begin{pmatrix} 1200 & \gamma_{2,1} & \gamma_{3,1} \end{pmatrix}$ where $\gamma_{i,1} \sim \mathrm{N}(0, 1)$ for $i = 2, 3$. For $H = 2$, $\Gamma_0' = \begin{pmatrix} 0 & 0 & \gamma_{3,1} \end{pmatrix}$ with $\gamma_{3,1} \sim \mathrm{N}(0, 1)$. For $H = 3$, $\Gamma_0 = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}$ with no free parameters.

- For the models with $H = 1, 2$ factors, we estimate parameters in the matrix Γ_1. For $H = 1$, $\Gamma_1' = \begin{pmatrix} 1200 & \gamma_{2,1} & \gamma_{3,1} \end{pmatrix}$ where $\gamma_{i,1} \sim \mathrm{N}(1200, 250)$ for $i = 2, 3$. For $H = 2$, $\Gamma_1' = \begin{pmatrix} 1200 & 0 & \gamma_{3,1} \\ 0 & 1200 & \gamma_{3,2} \end{pmatrix}$ with $\gamma_{3,i} \sim \mathrm{N}(600, 250)$ for $i = 2, 3$. For $H = 3$, $\Gamma_1 = I_H \times 1200$ with no free parameters. The scale 1200 is chosen so that h_t has the same scale as the yield factors g_t.

- The off-diagonal, lower triangular elements of Σ_g are $\Sigma_{g,ij} \sim \mathrm{N}\left(0.0001, 10^{-6}\right)$. For identification of the models with stochastic volatility, the diagonal values of Σ_g are fixed at 0.0003.

- $\Omega \sim \mathrm{InvWishart}(\nu_\omega, S_\omega)$, where $\nu_\omega = N + 3$ and $S_\omega = (\nu_\omega - N - 1) \times I_N \times 5 \times 10^{-7}$.