NBER WORKING PAPER SERIES

# MONETARY POLICY UNCERTAINTY AND ECONOMIC FLUCTUATIONS

Drew D. Creal
Jing Cynthia Wu

Working Paper 20594
http://www.nber.org/papers/w20594

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2014

Previously circulated as "Interest Rate Uncertainty and Economic Fluctuations." We thank Torben Andersen, Peter Christoffersen, Todd Clark, Steve Davis, Marty Eichenbaum, Bjorn Eraker, Jim Hamilton, Lars Hansen, Steve Heston, Jim Nason, Dale Rosenthal, Dora Xia, Lan Zhang and seminar and conference participants at Chicago Booth, Northwestern, UCL, Ohio State, NC State, Cleveland Fed, Illinois, Indiana, Texas A&M, Houston, Bank of England, Deutsche Bundesbank, Conference in Honor of James Hamilton, Annual Econometric Society Winter Meetings, Fifth Risk Management Conference, MFA, Midwest Econometrics, CFE. 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” and “Interest Rate Uncertainty and Economic Fluctuations”. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Monetary Policy Uncertainty and Economic Fluctuations
Drew D. Creal and Jing Cynthia Wu
NBER Working Paper No. 20594
October 2014, Revised January 2016
JEL No. C5,E4


## ABSTRACT

We investigate the relationship between uncertainty about monetary policy and its transmission mechanism, and economic fluctuations. We propose a new term structure model where the second moments of macroeconomic variables and yields can have a first-order effect on their dynamics. The data favors a model with two unspanned volatility factors that capture uncertainty about monetary policy and the term premium. Uncertainty contributes negatively to economic activity. Two dimensions of uncertainty react in opposite directions to a shock to the real economy, and the response of inflation to uncertainty shocks vary across different historical episodes.



Drew D. Creal
University of Chicago
Booth School of Business
5807 South Woodlawn Ave
Chicago, IL 60637
dcreal@chicagobooth.edu

Jing Cynthia Wu
University of Chicago
Booth School of Business
5807 South Woodlawn Avenue
Chicago, IL 60637
and NBER
cynthia.wu@chicagobooth.edu

# 1 Introduction

We investigate the relationship between uncertainty about monetary policy and its transmission mechanism, and economic fluctuations. The core question of interest is: does uncertainty about monetary policy have a real effect? An equally important question is: how do macroeconomic shocks influence interest rate uncertainty? While numerous studies have focused on monetary policy and its transmission mechanism, less attention has been placed on understanding uncertainty surrounding this transmission mechanism and their relation with the real economy. We study these questions by introducing a new term structure model with two novel features. First, we jointly model the first and second moments of macroeconomic variables and yields: uncertainty is extracted from their volatility, and it has a direct impact on the conditional means of these variables in a vector autoregression (VAR).[1]

Second, we decompose uncertainty of interest rates into two economic dimensions: the policy component, and the market transmission component captured by the term premium. Public commentary by policy-makers at central banks worldwide indicates that the term premium is one of the most important pieces of information extracted from the term structure of interest rates. Understanding the term premium and its uncertainty is crucial in making policy decisions and evaluating how successful monetary policy is in achieving its goals.

We contribute to the term structure literature by devising a no-arbitrage model with multiple unspanned stochastic volatility factors, i.e., the factors driving volatility are distinct from the factors driving yields. We show that our model can successfully fit the data for both the cross section of yields and their volatility, and the data suggests two volatility factors. We introduce a new rotation to the literature to capture the factor structure in an economically meaningful way. We decompose the long term interest rate into the expectation component and the term premium component. The former is agents' expectation about the future path of monetary policy, which the central bank can influence through policies like forward guidance. The latter relies on the market, and captures how monetary policy gets

[1] We define uncertainty as log volatility.

propagated from the short term interest rate to long term interest rates. The new rotation utilizes this decomposition and sets the three yield factors to be the short term interest rate, the expectation component of the long rate, and the term premium component. The two volatility factors are for the short rate and term premium, which can be conveniently interpreted as uncertainty about monetary policy and its transmission mechanism.

We document the relationship between interest rate uncertainty and economic fluctuations through impulse responses. Uncertainty is countercyclical, and precedes worse economic conditions and higher unemployment rates. This finding is consistent with the existing literature on uncertainty. What sets this paper apart from the literature is that our focus is on two aspects of interest rate uncertainty. The distinction between the two dimension lies in how they react to news about the real economy. A higher unemployment rate leads to higher uncertainty about term premia, which reflects the market's reaction to bad economic news. In contrast, uncertainty about monetary policy decreases in response to the same news. This is consistent with the Fed's proactive response to combat crisis historically.

One benefit of jointly modeling the first and second moments simultaneously is to allow the impulse responses to vary through time depending on the state of the economy. This is not possible for the models in the literature, unless they have time-varying autoregressive coefficients. Empirically, the response of inflation to uncertainty shocks varies through time. For example, in response to a positive shock to monetary policy uncertainty, inflation kept going up during the Great Inflation, when high inflation was considered bad for the economy. In contrast, inflation decreased in response to the same shock during Volcker's tenure. This is consistent with Volcker's reputation as an inflation hawk. It also barely responded during the Great Recession, when the concern is centered around deflation. These demonstrate the non-cyclical feature of inflation. A positive shock to term premia uncertainty leads to a positive reaction of inflation during Greenspan's Conundrum, and a negative reaction for the Volcker period. The former adds additional evidence of the importance of the term premium during that period. All of these economically meaningful distinctions can only be

observed through time-varying impulse responses. Standard impulse responses are close to zero, insignificant and potentially misleading, because they are averages of the positive and negative time-varying impulse responses.

Our historical decomposition further quantifies the two-way link between the real economy and interest rate uncertainty. Historically monetary policy uncertainty has contributed negatively to the inflation rate, which heightened at -0.7% after the Great Recession, placing further deflationary pressure during that period. Both monetary policy uncertainty and term premium uncertainty added positively to the unemployment rate historically. The contribution of monetary policy uncertainty peaked in the early 1980s at about 0.55%, while that of term premium uncertainty had peaks in the early 70s, early 80s and mid 2000s at the highest of 0.7%. The peak in the 2000's is associated with Greenspan's Conundrum, where our empirical evidence is consistent with the popular view that the term premium and its uncertainty increased. How monetary policy uncertainty and term premium uncertainty factor into unemployment differs since the Great Recession, with the former becoming negative potentially due to less uncertainty surrounding monetary policy, and the latter remaining positive with still significant uncertainty in the market. Consistent with our impulse responses, inflation contributed positively to both uncertainty measures in the 1980s, and negatively at the beginning and end of our sample period. The contributions of unemployment rate shocks to monetary policy uncertainty and term premium uncertainty take opposite signs. This is further evidence of the two dimensions of uncertainty.

Our paper contributes to the econometrics literature on estimation of vector autoregressions with stochastic volatility. When volatility enters the conditional mean of a vector autoregression, the popular Markov chain Monte Carlo (MCMC) algorithms for stochastic volatility models of Kim, Shephard, and Chib(1998) cannot be used. We develop a MCMC algorithm based on the particle Gibbs sampler which are efficient and can handle a wide range of models. We are the first to introduce this algorithm into the macro-finance literature.

The remainder of the paper is organized as follows. We describe our relationship to the

literatures in Subsection 1.1. Section 2 presents the new term structure model together with the new rotation. 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(2014) for a survey; Baker, Bloom, and Davis(2015), Jurado, Ludvigson, and Ng(2015), Bekaert, Hoerova, and Lo Duca(2013), and Aastveit, Natvik, and Sola(2013) for empirical evidence; and Ulrich(2012), Pástor and Veronesi(2012) and Pástor 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 macroeconomic variables and yields (the factors driving the volatility of inflation, unemployment, and interest rates) and it directly impacts the first moment of macroeconomic variables. In contrast, the uncertainty literature typically extracts an estimate of uncertainty in a data pre-processing step, often as the second moment of observed macroeconomic or financial time series. Researchers then use this estimate in a second step as an observable variable in a homoskedastic vector autoregression. (ii) This literature has so far focused on one dimension of uncertainty, whether it's policy uncertainty or macroeconomic uncertainty. We discuss two dimensions of interest rate uncertainty, and their distinct economic implications. (iii) Different from the rest of the literature, we focus on uncertainty about monetary policy and its transmission mechnism.

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). A recent paper by Mumtaz and Zanetti(2013) is closely related to ours in terms of how we specify the factor dynamics. Both papers allow the volatility factors to enter the conditional mean and have a first order impact on key macroeconomic aggregates. This is absent from most of the existing models in this literature, and we show its importance through impulse responses. The main difference between our paper and Mumtaz and Zanetti(2013) is that our paper introduces a factor structure for the volatilities, and ties these factors into a no-arbitrage term structure model. Our paper also introduces a new and more efficient Markov chain Monte Carlo algorithm known as a particle Gibbs sampler, Andrieu, Doucet, and Holenstein(2010), that can be used for a wide variety of multivariate time series models.

Finally, we contribute to the term structure literature by introducing a flexible way to simultaneously fit yields and their volatilities at different maturities. In the earlier literature, e.g. Dai and Singleton(2000) and Duffee(2002), volatility factors must simultaneously fit both the level of yields and their volatility. The factors from estimated models end up fitting the conditional mean of yields, and consequently they do not accurately estimate the conditional volatility. To break the tension, Collin-Dufresne and Goldstein(2002) proposed the class of unspanned stochastic volatility (USV) models which separate the dynamics of volatility from yield factors.[2] Creal and Wu(2015b) 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. Collin-Dufresne, Goldstein, and Jones(2009) point to the necessity of multiple unspanned volatility factors, and Joslin(2015) implemented a version with two volatility factors where only one is unspanned. Our model falls into the USV classification. The difference from the existing USV models is that we do not restrict the cross sectional fit of the model, and we

[2]Prior empirical work that studies whether volatility is priced using interest rate derivatives or high frequency data includes Bikbov and Chernov(2009), Andersen and Benzoni(2010), Joslin(2010), Mueller, Vedolin, and Yen(2011), and Christensen, Lopez, and Rudebusch(2014).

are the first to implement models with more than one unspanned volatility factor. In related work, Cieslak and Povala(2015) estimate a model with multiple spanned volatilities, where they use additional information from realized volatility to effectively place more weight in the objective function (such as the likelihood function) for the factors to fit the volatility.

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 has its own innovations and is not a deterministic function of past data. This is important because with our framework, we can use tools from the VAR literature such as impulse responses to study the influence of an uncertainty shock. This is not directly available in GARCH-M type models because there is no separate shock to volatility. Jo(2014) is similar to our paper in this spirit: while her focus is the uncertainty of oil prices shocks, we focus on uncertainty shocks from the term structure of interest rates.

# 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 – originating from the volatility of the yield curve and macroeconomic variables – 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.

## 2.1 Dynamics

The model has a $M \times 1$ vector of macroeconomic variables $m_t$, and a $G \times 1$ vector of conditionally Gaussian yield factors $g_t$ that drive bond prices. The $H \times 1$ vector of factors

$h_t$ determine the volatility of macroeconomic variables and yields, and we refer to them as uncertainty factors. The total number of factors is $F = M + G + H$.

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

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

The dynamics for the yield factors are

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

The diagonal time-varying volatility is a function of the uncertainty factors $h_t$

$$\begin{pmatrix} \text{diag}(D_{m,t}) \\ \text{diag}(D_{g,t}) \end{pmatrix} = \exp\left(\frac{\Gamma_0 + \Gamma_1 h_t}{2}\right). \quad (3)$$

The factors $h_t$ have dynamics

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

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 $\Sigma_{gm}, \Sigma_{hm}$ and $\Sigma_{hg}$.

We collect the state variables into the vector $f_t = (m'_t, g'_t, h'_t)'$ and write the system (1) - (4) as a vector autoregression

$$f_{t+1} = \mu_f + \Phi_f f_t + \Sigma_t \varepsilon_{t+1}. \quad (5)$$

We define $\bar{\mu}_f \equiv (I - \Phi_f)^{-1} \mu_f$ as the unconditional mean.

The identifying assumptions between the macroeconomic and yield factors are similar to

standard assumptions made in the VAR literature; i.e. macroeconomic 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(2015). We order the volatility factors after the macroeconomic and yield factors, so that uncertainty shocks are not contaminated by the first moment shocks. We also order interest rate uncertainty after macroeconomic uncertainty, so that our interest rate uncertainty shocks do not simply reflect macroeconomic uncertainty; i.e. they capture additional variation above and beyond what can be explained by macroeconomic uncertainty.

Of critical importance for our analysis, the uncertainty factors $h_t$ impact the macroeconomy through the conditional mean term $\Phi_{mh}h_t$ in (1), and are identified from the conditional variance of observed macroeconomic data and yields through $D_{mt}$ and $D_{gt}$. This unique combination unifies two literatures; the literature on VARs 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 and the macroeconomy and/or asset-prices (e.g. Baker, Bloom, and Davis(2015), Jurado, Ludvigson, and Ng(2015), Bekaert, Hoerova, and Lo Duca(2013), and Aastveit, Natvik, and Sola(2013)).

To ensure stability, the conditional mean of $h_{t+1}$ does not depend on the levels of $g_t$ or $m_t$. Otherwise, the system will be explosive even if the modulus of the eigenvalues of $\Phi_f$ in (5) are all less than one. To compensate for this restriction, we allow contemporaneous shocks of macroeconomic variables $\varepsilon_{m,t+1}$ and yields $\varepsilon_{g,t+1}$ to drive $h_{t+1}$. The timing assumption that today's shocks to the macroeconomy or yields determine their volatility next period makes intuitive sense.[3]

We follow the macroeconomics literature and use a log-normal process for the volatility

[3]In standard models where $\Sigma_{hm}$, $\Sigma_{hg}$, $\Phi_{mh}$, $\Phi_{gh}$ are zero, different timing conventions for the conditional volatility are observationally equivalent. In models with a leverage effect when $\Sigma_{hm}$, $\Sigma_{hg}$ are not zero, different timing leads to different models. Our timing is consistent with Omori, Chib, Shephard, and Nakajima(2007) and the discrete-time stochastic volatility models in the term structure literature; see, e.g. Bansal and Shaliastovich(2013) and Creal and Wu(2015b).

in (1) - (4). The matrices $\Gamma_0$ and $\Gamma_1$ 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 $M + G \neq H$.

## 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. Collin-Dufresne, Goldstein, and Jones(2009) and Creal and Wu(2015b). We also follow Joslin, Priebsch, and Singleton(2014) and assume the macro factors $m_t$ are unspanned. 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 $g_t$:

$$r_t = \delta_0 + \delta_{1,g}' g_t. \tag{6}$$

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,

$$P_t^{(n)} = \mathbb{E}_t^{\mathbb{Q}} \left[ \exp\left(-r_t\right) P_{t+1}^{n-1} \right], \tag{7}$$

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

$$g_{t+1} = \mu_g^{\mathbb{Q}} + \Phi_g^{\mathbb{Q}} g_t + \Sigma_g^{\mathbb{Q}} \varepsilon_{g,t+1}^{\mathbb{Q}}, \quad \varepsilon_{g,t+1}^{\mathbb{Q}} \sim \mathrm{N}\left(0, I\right). \tag{8}$$

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

variables

$$P_t^{(n)} = \exp\left(\bar{a}_n + \bar{b}_n' g_t\right). \tag{9}$$

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

$$\bar{a}_n = -\delta_0 + \bar{a}_{n-1} + \mu_g^{\mathbb{Q}\prime}\bar{b}_{n-1} + \frac{1}{2}\bar{b}_{n-1}'\Sigma_g^{\mathbb{Q}}\Sigma_g^{\mathbb{Q}\prime}\bar{b}_{n-1}, \tag{10}$$

$$\bar{b}_n = -\delta_{1,g} + \Phi_g^{\mathbb{Q}\prime}\bar{b}_{n-1}, \tag{11}$$

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

$$y_t^{(n)} = a_n + b_n' g_t \tag{12}$$

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 flexibly in no-arbitrage term structure models while keeping bond prices simple through the assumptions (6) - (8). In most non-Gaussian term structure models, 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, $h_t$ is not priced by construction. Consequently, our model does not impose any restrictions on the cross section of the yield curve like the restrictions that are imposed in the USV models in Collin-Dufresne, Goldstein, and Jones(2009) and Creal and Wu(2015b). 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.[4] We will show the no arbitrage condition –

[4] 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.

the equivalence of the two probability measures – by deriving the Radon-Nikodym derivative in Subsection 2.4.

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 by the functional forms that achieve analytical bond prices.

## 2.3 Rotation and identification

In order for the model specified in 2.1 and 2.2 to be identified, we need to impose identifying restrictions to prevent the latent factors $g_t$ and $h_t$ from shifting and rotating. The term premium has been a fundamental part in the literature on the term structure of interest rates, from both the view of policy makers and academic researchers; see, e.g. Wright(2011), Bauer, Rudebusch, and Wu(2012), Bauer, Rudebusch, and Wu(2014), and Creal and Wu(2015a). For a better economic interpretation of uncertainty, we propose a rotation new to the literature and let $g_t = \left(r_t \; er_t^{(n^*)} \; tp_t^{(n^*)}\right)'$ where $r_t = y_t^{(1)}$ is the short rate, $er_t^{(n^*)}$ is the average expected future short rate

$$er_t^{(n^*)} \equiv \frac{1}{n^*}\mathbb{E}_t\left[r_t + \ldots + r_{t+n^*-1}\right],$$

and $tp_t^{(n^*)}$ is the term premium $tp_t^{(n^*)} \equiv y_t^{(n^*)} - er_t^{(n^*)}$ for a pre-specified maturity $n^*$. The corresponding volatilities can be interpreted as uncertainty about current monetary policy, the future path of monetary policy, and the term premium. Proposition 1 provides conditions that guarantee this rotation.[5]

**Proposition 1.** *Eigen decompose* $\Phi_f = Q\Lambda Q^{-1}$, $\Phi_g^{\mathbb{Q}} = Q_g^{\mathbb{Q}}\Lambda_g^{\mathbb{Q}}\left(Q_g^{\mathbb{Q}}\right)^{-1}$, *where* $\Lambda$ *and* $\Lambda_g^{\mathbb{Q}}$ *are matrices of eigenvalues. Define* $\tilde{\Lambda} \equiv \frac{1}{n^*}\left(I - \Lambda^{n^*}\right)\left(I - \Lambda\right)^{-1}$, $\tilde{\Lambda}^{\mathbb{Q}} \equiv \frac{1}{n^*}\left(I - \left(\Lambda^{\mathbb{Q}}\right)^{n^*}\right)\left(I - \Lambda^{\mathbb{Q}}\right)^{-1}$, *The following conditions guarantee that the yield factors are* $g_t = \left(r_t \; er_t^{(n^*)} \; tp_t^{(n^*)}\right)'$.

[5]Other rotations to take different linear combinations of yields as factors have been proposed in the literature, see Proposition 1 of Hamilton and Wu(2014) for example.

1. $b_{1,g} = \delta_{1,g} = \boldsymbol{e}_1$

2. $a_1 = \delta_0 = 0$

3. $\begin{pmatrix} 0 & \boldsymbol{e}_2' & 0 \\ 1\times M & 1\times G & 1\times H \end{pmatrix} Q = \begin{pmatrix} 0 & \boldsymbol{e}_1' & 0 \\ 1\times M & 1\times G & 1\times H \end{pmatrix} Q\tilde{\Lambda}$

4. $\bar{\mu}_{g,1} = \bar{\mu}_{g,2}$

5. $(\boldsymbol{e}_2' + \boldsymbol{e}_3')\, Q_g^{\mathbb{Q}} = \boldsymbol{e}_1' Q_g^{\mathbb{Q}} \tilde{\Lambda}_g^{\mathbb{Q}}$

6. $\bar{\mu}_{g,1}^{\mathbb{Q}} = \bar{\mu}_{g,2}^{\mathbb{Q}} + \bar{\mu}_{g,3}^{\mathbb{Q}} + Jensen's\ inequality$

*where $\boldsymbol{e}_i$ is the $i$-th column of the identity matrix $I_G$, $\bar{\mu}_{g,i}$ is the $i$-th element of the unconditional mean under $\mathbb{P}$, and $\bar{\mu}_{g,i}^{\mathbb{Q}}$ is the $i$-th element of the unconditional mean under $\mathbb{Q}$.*

**Proof:** See Appendix A.1.

The first and second conditions guarantee that the first element of the state vector is $r_t$ due to (6). Conditions three and four ensure that the second element of the state vector is the expected future short rate $er_t^{(n^*)}$. Condition 4 says that $r_t$ and $er_t^{(n^*)}$ have the same unconditional mean while condition three guarantees that the forecast function at horizon $n^*$ is internally consistent so that $er_t^{(n^*)}$ is the average of the expected future path of $r_t$. The last two restrictions are for $tp_t^{(n^*)}$. They ensure that the yield $y_t^{(n^*)}$ is the sum of $er_t^{(n^*)}$ and $tp_t^{(n^*)}$ by guaranteeing that the bond loadings at horizon $n^*$ are $a_{n^*} = 0$ and $b_{n^*} = (0, 1, 1)'$. Condition 6 says the unconditional mean of $r_t$ and $y_t^{(n^*)}$ under $\mathbb{Q}$ are the same up to a Jensen's inequality term. This fits the definition of $\mathbb{Q}$ as the risk neutral measure. In Appendix A, we prove the proposition and discuss further how we implement this rotation for our benchmark model in Section 4.

This general proposition can be implemented for different cases, whether the eigenvalues are all distinct and real, some eigenvalues are complex, or there exist repeated eigenvalues. The following Corollary details these cases respectively. In the corollary, for a matrix $A$, we use $A_{i,j}$ to denote the $(i, j)$ element.

**Corollary 1.** • ***Unique real eigenvalues:***

*For distinct real eigenvalues with diagonal matrices* $\Lambda$ *and* $\Lambda_g^{\mathbb{Q}}$*, conditions 3 and 5 are equivalent to*

- $Q_{M+2,k} = Q_{M+1,k}\tilde{\Lambda}_k$*, for* $k = 1, ..., N$*.*
- $Q^{\mathbb{Q}}_{g,2,k} + Q^{\mathbb{Q}}_{g,3,k} = Q^{\mathbb{Q}}_{g,1,k}\tilde{\Lambda}^{\mathbb{Q}}_{g,k}$*, for* $k = 1, ..., G$*.*

- ***Complex eigenvalues:*** *For matrices with pair(s) of complex eigenvalues, the corresponding block of* $\Lambda$ *or* $\Lambda_g^{\mathbb{Q}}$ *becomes* $\begin{pmatrix} \Lambda_{c1} & \lambda_{c2} \\ -\lambda_{c2} & \Lambda_{c1} \end{pmatrix}$*, and the corresponding part of* $\tilde{\Lambda}$ *or* $\tilde{\Lambda}_g^{\mathbb{Q}}$ *is* $\begin{pmatrix} \tilde{\Lambda}_{c1} & \tilde{\lambda}_{c2} \\ -\tilde{\lambda}_{c2} & \tilde{\Lambda}_{c1} \end{pmatrix}$*, the corresponding block for the right hand side for conditions 3 and 5 becomes* $Q_{.,c1}\tilde{\Lambda}_{c1} - Q_{.,c1+1}\tilde{\lambda}_{c2}$ *and* $Q_{.,c1+1}\tilde{\Lambda}_{c1} + Q_{.,c1}\tilde{\lambda}_{c2}$*.*

- ***Repeated eigenvalues:*** *For matrices with pair(s) of repeated eigenvalues, the corresponding block of* $\Lambda$ *or* $\Lambda_g^{\mathbb{Q}}$ *becomes* $\begin{pmatrix} \Lambda_{r1} & 1 \\ 0 & \Lambda_{r1} \end{pmatrix}$*, and the corresponding part of* $\tilde{\Lambda}$ *or* $\tilde{\Lambda}_g^{\mathbb{Q}}$ *is* $\begin{pmatrix} \tilde{\Lambda}_{r1} & \tilde{\lambda}_{r2} \\ 0 & \tilde{\Lambda}_{r1} \end{pmatrix}$*, the corresponding block for the right hand side for conditions 3 and 5 becomes* $Q_{.,r1}\tilde{\Lambda}_{r1}$ *and* $Q_{.,r1+1}\tilde{\Lambda}_{r1} + Q_{.,r1}\tilde{\lambda}_{r2}$*.*

## 2.4 Stochastic discount factor

The pricing equation in (7) can be equivalently written as

$$P_t^{(n)} = \mathbb{E}_t\left[\mathcal{M}_{t+1}P_{t+1}^{(n-1)}\right]. \tag{13}$$

The stochastic discount factor (SDF) $\mathcal{M}_{t+1}$ incorporates the risk premium and time discounting. In a micro-founded model, this depends on the ratio of marginal utility. For structural

models whose reduced form is an affine term structure model, like the one specified in this paper; see, e.g. Piazzesi and Schneider(2007) and Creal and Wu(2015a). To complete the model, we can write down general $\mathbb{Q}$ dynamics for the volatility factors $p^{\mathbb{Q}}\left(h_{t+1}|\mathcal{I}_t;\theta\right)$. The SDF is defined as

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

which makes (7) and (13) consistent. $\mathcal{I}_t$ denotes the information set up to and including time $t$, and $\theta$ is a vector of parameters. Although we can specify a process for $h_{t+1}$ under $\mathbb{Q}$, the parameters are not identified using bond prices alone. For example, if we specify $p^{\mathbb{Q}}\left(h_{t+1}|\mathcal{I}_t;\theta\right) = p\left(h_{t+1}|\mathcal{I}_t;\theta\right)$, then the pricing kernel does not depend on the $\mathbb{Q}$ dynamics of $h_{t+1}$:

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

# 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 a particle Markov chain Monte Carlo (MCMC) algorithm known as the particle Gibbs sampler, see Andrieu, Doucet, and Holenstein(2010) and, for a survey on particle filters, see Creal(2012). We are the first to introduce this algorithm into the macro-finance literature. The particle Gibbs sampler uses a particle filter within a standard Gibbs sampling algorithm to act as a proposal distribution for the latent variables whose full conditional distributions are intractable. We outline the basic ideas of the MCMC algorithm in this section and provide full details in Appendix B. Our paper contributes to the econometrics literature

on Bayesian estimation of term structure models and vector autoregressions with stochastic volatility; see, e.g. Cogley and Sargent(2005), and Primiceri(2005) for VARs and Chib and Ergashev(2009) and Bauer(2015) for Gaussian affine term structure models. The MCMC algorithms developed in this paper are efficient and can handle a wide range of VARMA models with stochastic volatility.

## 3.1 State space forms

Stack the yields $y_t^{(n)}$ from (12) in order of increasing maturity for $n = n_1, n_2, ..., n_N$, and assume that all yields are observed with Gaussian measurement errors:

$$y_t = A + Bg_t + \eta_t, \qquad \eta_t \sim \mathrm{N}(0, \Omega), \tag{14}$$

where $A = (a_{n_1}, \ldots, a_{n_N})'$ and $B = (b_{n_1}, ..., b_{n_N})'$. 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 particle 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 $\theta$. In practice, we use two different state space representations that either condition on the most recent draw of the yield factors $g_{1:T}$ or the volatility factors $h_{0:T}$.

**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 (14), the macroeconomic variables $m_t$ in (1), and the volatility factors $h_t$ in (4). Using this representation, we draw the latent yield factors $g_{1:T}$ in a large block using the Kalman filter and forward-filtering backward sampling algorithms; see, Durbin and Koopman(2002).

Importantly, most of the parameters that enter the dynamics of $g_t$ can be drawn without conditioning on the state variables $g_{1:T}$.

**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 consisting of the macroeconomic variables and yield factors in (1) and (2) and transition equation for $h_t$ in (4). The observation equations for $m_{t+1}$ and $g_{t+1}$ are non-linear in $h_t$. Given that $h_t$ enters the conditional mean of $m_{t+1}$ and $g_{t+1}$, the MCMC algorithms for stochastic volatility models developed by Kim, Shephard, and Chib(1998) that are widely used in macroeconometrics are not applicable. A contribution of this paper is developing efficient MCMC algorithms to handle models where volatility enters the conditional mean. Importantly, the volatility factors $h_{0:T}$ can still be drawn from their full conditional distribution $p\left(h_{0:T}|y_{1:T}, m_{1:T}, g_{1:T}, \theta\right)$ in large blocks using a particle Gibbs sampler, see Appendix B.2.

## 3.2 MCMC and particle filter

Our MCMC algorithm alternates between the two state space forms. We split the parameters into blocks $\theta = (\theta_g', \theta_r')'$, where we draw $\theta_g$ from the first state space form. We draw the parameters $\theta_r$ 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 $h_{0:T}$, write the model as a conditionally, linear Gaussian state space model.
   (a) Draw $\theta_g$ using the Kalman filter without conditioning on $g_{1:T}$.
   (b) Draw $g_{1:T}$ jointly using forward filtering and backward sampling; see, e.g. de Jong and Shephard(1995), Durbin and Koopman(2002).
2. Conditional on $g_{1:T}$, the model is a non-linear state space model.
   (a) Draw $h_{0:T}$ using the particle filter, see Appendix B for details.

3. Draw any remaining parameters $\theta_r$ 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 $p\left(g_{1:T}, h_{0:T}, \theta | y_{1:T}, m_{1:T}\right)$.

Another object of interest are the filtered estimates of the state variables and the value of the log-likelihood function. We calculate these using a particle filter known as the mixture Kalman filter; see Chen and Liu(2000).[6] 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 mixture Kalman filter 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. See Appendix B.3 for details.

# 4 Economic implications

Key questions of interest are: does uncertainty, specifically uncertainty about monetary policy and its transmission mechanism, have a real effect? And, how do macroeconomic shocks influence uncertainty? This section investigates these questions using the modeling and estimation tools described in the previous sections. Consistent with the existing literature on uncertainty, we also find that uncertainty contributes negatively to economic activity and is associated with higher unemployment. The novelty of our model is that we focus on two aspects of interest rate uncertainty: uncertainty about monetary policy itself, and uncertainty about the risk premium. This distinction allows us to explore different dimensions of uncertainty of economic interest: (1) monetary policy uncertainty and risk premium uncertainty

[6]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 macroeconomic plus yields model

| | | $\bar{\mu}_g^{\mathbb{Q}} \times 1200$ | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 11.267 | 8.874 | 2.354 | | | | |
| | | (0.629) | (0.624) | (0.442) | | | | |
| | | $\Phi_g^{\mathbb{Q}}$ | | | | | | |
| | | 0.794 | 0.269 | -0.003 | | | | |
| | | (0.022) | (0.025) | (0.022) | | | | |
| | | -0.002 | 0.985 | 0.076 | | | | |
| | | (0.010) | (0.016) | (0.009) | | | | |
| | | -0.017 | 0.031 | 0.948 | | | | |
| | | (0.010) | (0.017) | (0.009) | | | | |
| | | $\lambda_g^{\mathbb{Q}}$ | | | | | | |
| | | 0.996 | 0.948 | 0.7829 | | | | |
| | | (0.001) | (0.003) | (0.015) | | | | |
| $\bar{\mu}_m \times 1200$ | | $\bar{\mu}_g \times 1200$ | | | | | | |
| 2.197 | 6.149 | 2.130 | 2.130 | 1.133 | -17.337 | -16.891 | -16.386 | -18.695 |
| (1.176) | (0.658) | (0.353) | (0.353) | (0.295) | (0.393) | (0.369) | (0.544) | (0.269) |
| $\Phi_m$ | | $\Phi_{mg}$ | | | $\Phi_{mh}$ | | | |
| 0.982 | -0.011 | -0.015 | 0.022 | -0.009 | 0.032 | -0.009 | 0.003 | 0.010 |
| (0.009) | (0.018) | (0.020) | (0.022) | (0.017) | (0.029) | (0.061) | (0.017) | (0.029) |
| 0.006 | 0.949 | -0.005 | 0.0002 | 0.007 | 0.027 | 0.098 | -0.007 | 0.059 |
| (0.007) | (0.011) | (0.011) | (0.012) | (0.015) | (0.016) | (0.029) | (0.011) | (0.018) |
| $\Phi_{gm}$ | | $\Phi_g$ | | | $\Phi_{gh}$ | | | |
| 0.001 | -0.004 | 0.877 | 0.136 | 0.007 | 0.019 | -0.0002 | -0.005 | 0.001 |
| (0.005) | (0.007) | (0.015) | (0.019) | (0.010) | (0.013) | (0.020) | (0.012) | (0.010) |
| -0.0001 | 0.0004 | -0.020 | 1.018 | -0.001 | -0.002 | 0.0001 | 0.001 | -0.0001 |
| (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) | (0.002) | (0.001) | (0.001) |
| 0.001 | -0.007 | 0.002 | -0.001 | 0.984 | 0.007 | 0.018 | -0.002 | 0.015 |
| (0.003) | (0.005) | (0.004) | (0.005) | (0.007) | (0.006) | (0.013) | (0.004) | (0.009) |
| | | | | | $\Phi_h$ | | | |
| | | | | | 0.990 | -0.008 | 0.002 | -0.003 |
| | | | | | (0.005) | (0.009) | (0.003) | (0.005) |
| | | | | | -0.002 | 0.987 | 0.004 | -0.004 |
| | | | | | (0.004) | (0.010) | (0.003) | (0.005) |
| | | | | | -0.003 | -0.009 | 0.997 | -0.006 |
| | | | | | (0.009) | (0.019) | (0.008) | (0.010) |
| | | | | | 0.005 | -0.004 | 0.000 | 0.992 |
| | | | | | (0.006) | (0.006) | (0.003) | (0.005) |

| $\Gamma_0$ | | | | | $\Gamma_1 \times 1/1200$ | | | |
|---|---|---|---|---|---|---|---|---|
| 0 | | | | | 1 | 0 | 0 | 0 |
| — | | | | | — | — | — | |
| 0 | | | | | 0 | 1 | 0 | 0 |
| — | | | | | — | — | — | |
| 0 | | | | | 0 | 0 | 1 | 0 |
| — | | | | | — | — | — | |
| 0.408 | | | | | 0 | 0 | 0.492 | 0.481 |
| 1.546 | | | | | — | — | (0.059) | (0.072) |
| 0 | | | | | 0 | 0 | 0 | 1 |
| — | | | | | — | — | — | — |
| $\lambda_f$ | | | | | | | | |
| 0.995 | 0.989 | 0.977 | 0.950 | 0.900 | 0.998 | 0.995 | 0.990 | 0.984 |
| (0.002) | (0.004) | (0.006) | (0.009) | (0.016) | (0.001) | (0.002) | (0.004) | (0.005) |
| | | | | | | | | |
| $\Sigma_m^* \times 1200$ | | | | | | | | |
| 0.196 | | | | | | | | |
| (0.066) | | | | | | | | |
| -0.004 | 0.245 | | | | | | | |
| (0.004) | (0.082) | | | | | | | |
| $\Sigma_{mg}^* \times 1200$ | | $\Sigma_g^* \times 1200$ | | | | | | |
| 0.005 | -0.023 | 0.322 | 0 | 0 | | | | |
| (0.004) | (0.015) | (0.132) | | | | | | |
| 0.002 | -0.057 | 0.152 | 0.273 | 0 | | | | |
| (0.006) | (0.026) | (0.066) | (0.085) | | | | | |
| 0.008 | 0.029 | -0.038 | -0.090 | 0.099 | | | | |
| (0.006) | (0.017) | (0.024) | (0.035) | (0.030) | | | | |
| $\Sigma_{mh}^* \times 1200$ | | $\Sigma_{gh}^* \times 1200$ | | | $\Sigma_h^* \times 1200$ | | | |
| 0.008 | -0.008 | 0.003 | -0.002 | -0.006 | 0.142 | | | |
| (0.020) | (0.023) | (0.015) | (0.019) | (0.012) | (0.046) | | | |
| 0.012 | 0.012 | 0.001 | -0.013 | -0.006 | 0.033 | 0.102 | | |
| (0.016) | (0.022) | (0.011) | (0.015) | (0.009) | (0.032) | (0.035) | | |
| 0.069 | -0.049 | 0.045 | 0.0004 | -0.012 | 0.012 | 0.009 | 0.268 | |
| (0.037) | (0.037) | (0.034) | (0.033) | (0.022) | (0.027) | (0.026) | (0.087) | |
| 0.038 | 0.035 | 0.001 | -0.025 | -0.004 | 0.007 | 0.007 | 0.060 | 0.098 |
| (0.021) | (0.026) | (0.015) | (0.019) | (0.011) | (0.013) | (0.012) | (0.040) | (0.034) |

*Posterior mean and standard deviations for the benchmark macroeconomic plus yields model. The parameters $\lambda_g^{\mathbb{Q}}$ and $\lambda_f$ are the eigenvalues of $\Phi_g^{\mathbb{Q}}$ and $\Phi_f$, respectively. We report the sub-components of* $\Sigma_f^* = \Sigma_f \, diag\left(\exp\left(\frac{\Gamma_0 + \Gamma_1 \bar{\mu}_h}{2}\right)\right)$.

react in opposite directions as a consequence of a positive shock to the unemployment rate. (2) The response of inflation to uncertainty shocks vary across different historical episodes.

To answer these questions, impulse responses to a one time shock are reported in Subsection 4.1. We then aggregate the overall effect with a historical decompositions in Subsection 4.2. 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. Our sample is

monthly from June 1953 to December 2013.

Our model has $G = 3$ yield factors which are rotated to be $g_t = \left(r_t \; er_t^{(60)} \; tp_t^{(60)}\right)'$ as explained in Section 2.3. In the benchmark model, we use $H = 4$ volatility factors, and this choice is warranted by the specification analysis in Section 5. The volatility factors capture the volatility of inflation, unemployment, the short rate $r_t$ and the term premium $tp_t^{(60)}$. We interpret the last two factors as monetary policy uncertainty and term premium uncertainty. The details for implementation are in Appendix A.2. Posterior means and standard deviations for the model's parameters are reported in Table 1.

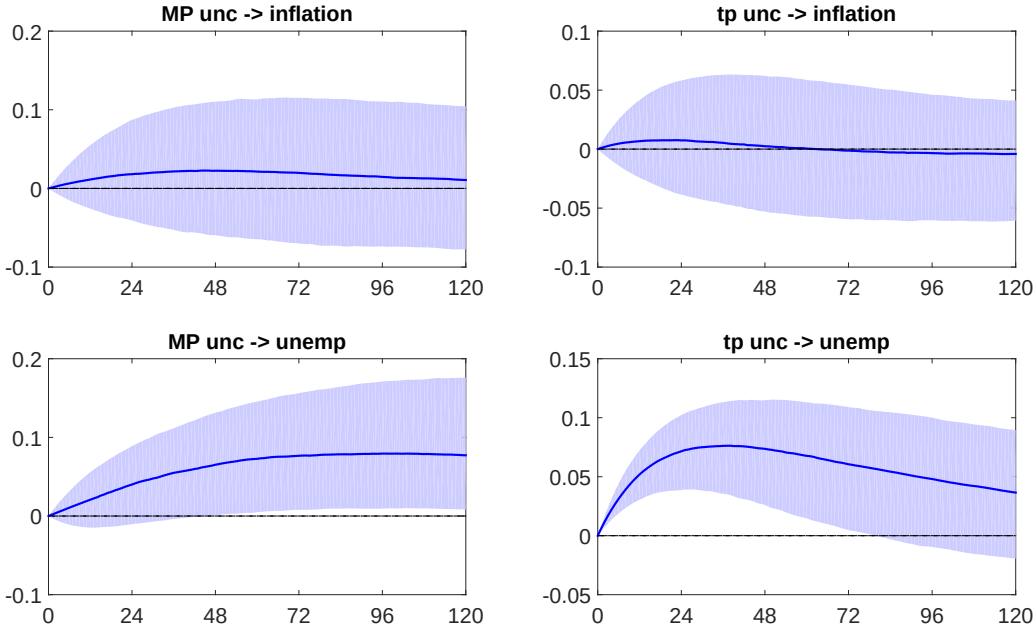
## 4.1 Impulse responses

**Responses of macroeconomic variables to uncertainty shocks** We first study the real effect of uncertainty shocks by plotting the impulse responses of macroeconomic variables to one standard deviation uncertainty shocks in Figure 1. We plot the median impulse responses in solid lines, with the [10%, 90%] highest posterior density intervals calculated from our MCMC draws in the shaded areas.

First, let us set intuition for how large a one standard deviation uncertainty shock is: a one standard deviation shock to short rate uncertainty is about 1/11 of the change in uncertainty leading up to the Great Recession. For term premium uncertainty, the relative magnitude is 1/12 . For further discussion on scale, time dynamics and the cyclical pattern of uncertainties, see Subsection 4.3.

Both the uncertainty of monetary policy and term premia have a negative impact on the real economy: higher uncertainty is associated with higher future unemployment rates. Both of them are statistically significant given by the [10%, 90%] highest posterior density intervals, and the sizes are similar as well. The impact of the monetary policy uncertainty shock peaks at 0.08%, and the effect dies out slowly. This is consistent with findings by Mumtaz and Zanetti(2013), for example. A higher uncertainty about the term premium, the component in the long term interest rate that is determined by the market rather than by

Figure 1: Responses of macroeconomic variables to uncertainty shocks




*Constant impulse responses to a one standard deviation shock to interest rate uncertainty. Left: monetary policy uncertainty ($h_{3,t}$); right: term premium uncertainty ($h_{4,t}$). top: inflation; bottom: unemployment rate. Units in y-axis: percentage points; Units in x-axis: months. The [10%, 90%] highest posterior density intervals are shaded. Sample: June 1953 - December 2013.*

the central bank, also leads to an increase of unemployment by as much as 0.08%.

The medium impulse response of inflation to uncertainty shocks are close to zero, and neither of them are statistically significant. This is related to the fact that this relationship changes over time, including signs. We will explore the time dependence of this relationship below.

**Time-varying (state dependent) impulse response** A unique feature of our hybrid model – jointly capturing the first and second moment effects of uncertainty – is the existence of time varying or state-dependent impulse responses where both the scale and shape vary across time.

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. In a VAR with heteroskedastic shocks – where uncertainty only effects the second moment, the impulse responses to a one unit shock have a different scale that varies through time because of the stochastic volatility. In either of these cases, the shape of the impulse response remains the same.

Our model can distinguish important historical episodes like the Great Recession (2007 - 2009) from the Great Inflation (1965 - 1982). 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 introducing many more state variables as time-varying parameters.

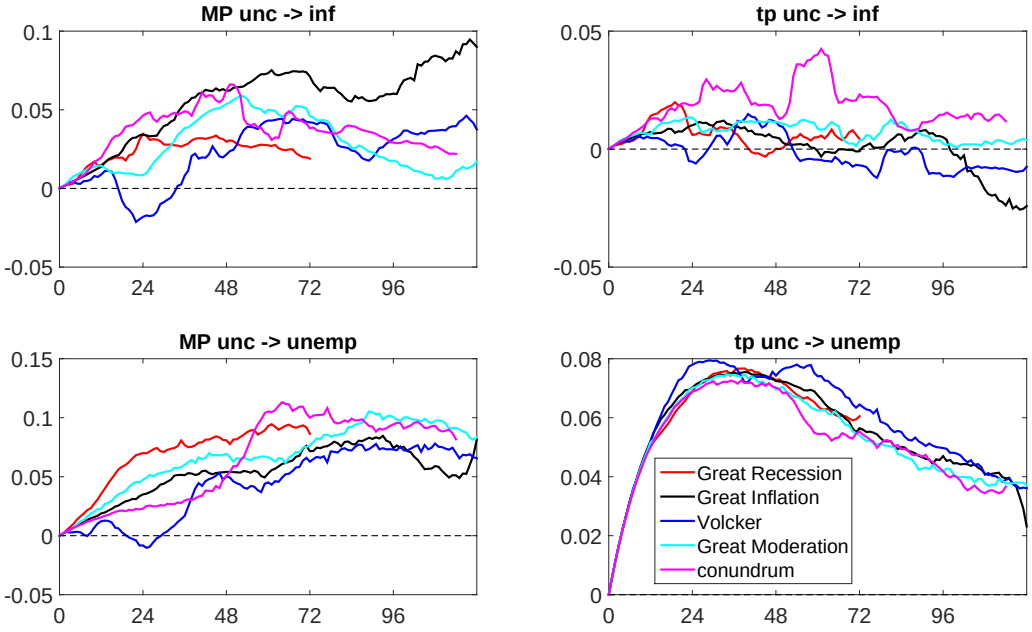
We plot the median impulse response[7] to a one standard deviation shock in Figure 2 for the following economically significant time periods: the Great Recession in red from December 2007, the Great Inflation in black from 1965, Volcker's tenure in blue from August 1979, the Great Moderation in turquoise from 1985, and Greenspan's conundrum in pink from June 2004. The basic intuition of the time varying impulse response is equivalent to the following counterfactual analysis: in the moving average representation of the vector autoregression, (i.e., representing the state of the economy in terms of an accumulation of past shocks), keep all the shocks the same except for the addition of a one standard deviation shock to a variable of interest at the beginning of the period we are investigating. See Appendix C for the calculations.

The responses of the unemployment rate to uncertainty, especially to term premium uncertainty across different periods in Figure 2, are close to each other and all clearly positive. This echoes the significant responses in the bottom panels of Figure 1.

In contrast, the variation across different periods for inflation (top row in Figure 2) is bigger, and the signs also differ across colors. This explains the small magnitude and insignificance in the top row of Figure 1. For example, a shock to monetary policy uncertainty

[7]Note: the uncertainty of the estimates are not reflected in the impulse responses in Figure 2, and we need to be cautious when interpreting the results.

Figure 2: Time-varying impulse response functions




*Impulse responses to a one standard deviation shock to interest rate uncertainty. Left: monetary policy uncertainty ($h_{3,t}$); right: term premium uncertainty ($h_{4,t}$). Top: inflation; bottom: unemployment rate. Units in y-axis: percentage points; Units in x-axis: months. 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.*

kept pushing up inflation during the Great Inflation, when inflation was high and associated with a bad state of the economy. In contrast, the blue line for the Volcker period demonstrated a downward pressure in the medium term. A plausible explanation is that Volcker is known to combat inflation aggressively. He acted to lower inflation in response to higher uncertainty in the market. Similarly, the red line also indicates a smaller upward pressure on inflation during the recent Great Recession, during which inflation was low and agents worry about deflation rather than hyper inflation. The Great Moderation is consistent with its reputation and the responses are not extreme. The variability of inflation's response to shocks can be explained by its non-cyclical nature; i.e., uncertainty is associated with worse economic condition but this can mean higher inflation for some periods and lower inflation for others.

The variation on the top right panel covers both positive and negative regions. The most positive reaction happened during the period known as Greenspan's Conundrum, when the then-chairman raised the benchmark overnight rate but failed to increase the rate with longer maturities. Researchers attributed this conundrum to variation in the term premium. Relatedly, we find term premium uncertainty had a bigger positive impact on inflation compared to other periods. The most negative response is during Volcker's tenure, again consistent with his reputation as an inflation hawk.

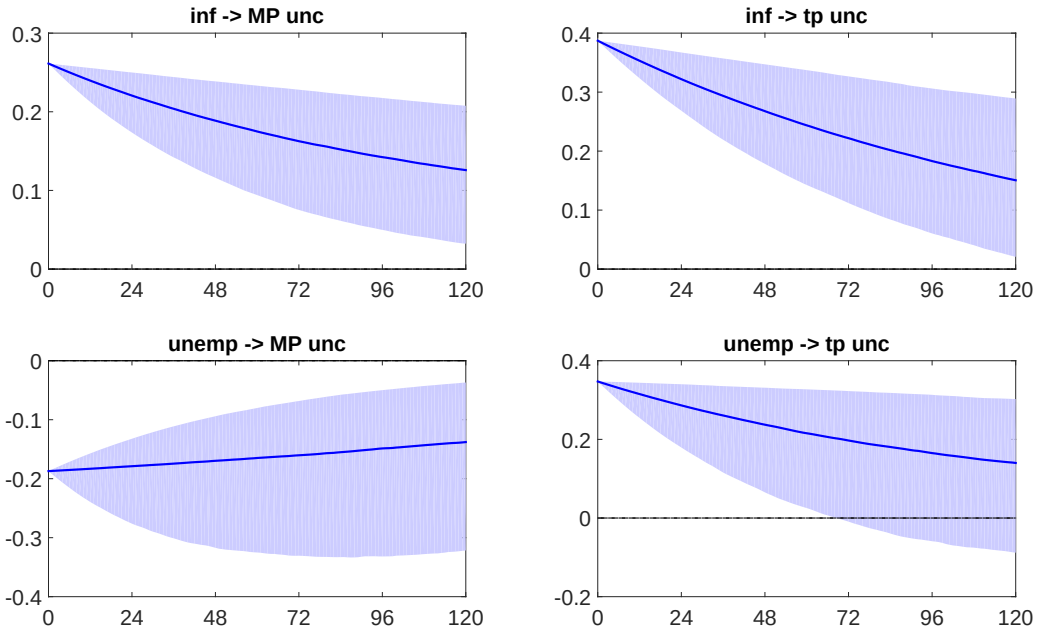
**Responses of uncertainty to macroeconomic shocks** An equally important question is how do macroeconomic shocks drive uncertainty? The constant impulse responses of uncertainty to macroeconomic shocks are captured in Figure 3. In response to a standard deviation shock to inflation, both monetary policy uncertainty and term premium uncertainty increase. The contemporaneous response is 0.26 standard deviations for the former, and almost 0.4 for the latter. Then, they die out through time.

Conversely, the two dimensions of uncertainty respond differently to shocks to the unemployment rate: monetary policy uncertainty responds negatively, whereas term premium uncertainty reacts positively. A higher unemployment rate injects more uncertainty to the market, hence term premium, the market determined component of interest rates. In contrast, the Fed has historically eased monetary policy when economic conditions worsen. That explains why we see a lower uncertainty about monetary policy following higher unemployment rate.

## 4.2 Historical decomposition

To better quantify the economic magnitude of the empirical link between macroeconomic variables and interest rate uncertainty, we resort to an alternative strategy through a historical decomposition, a strategy used, for example, by Wu and Xia(2015) to study unconventional monetary policy. A historical decomposition decomposes the historical dynamics of a

Figure 3: Responses of uncertainty to macroeconomic shocks



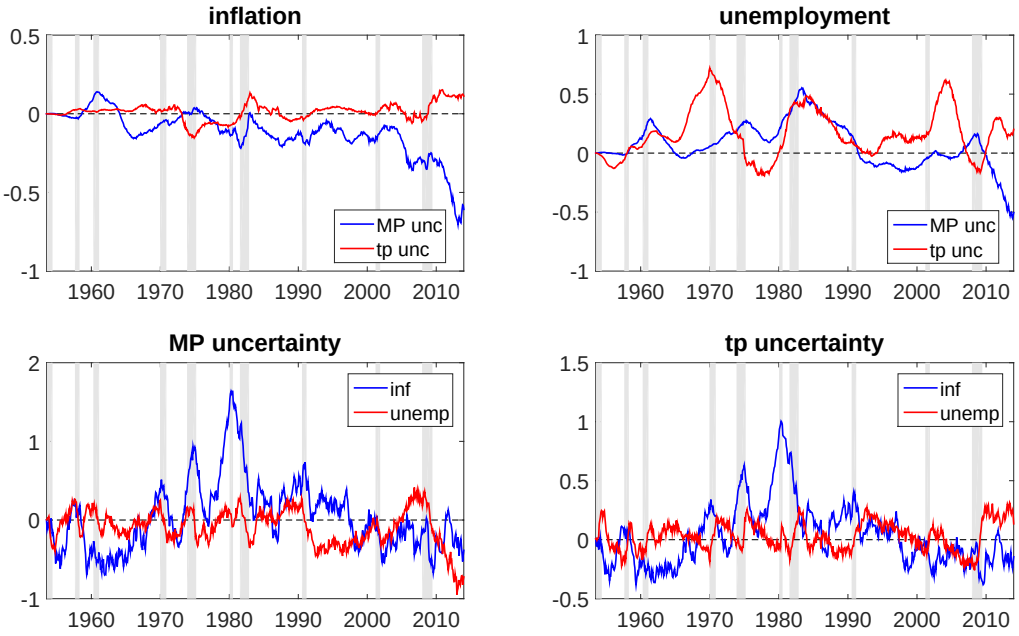

*Constant impulse responses to a one standard deviation shock to macroeconomic variables. The standard deviation is averaged across time. Left: monetary policy uncertainty ($h_{3,t}$); right: term premium uncertainty ($h_{4,t}$). Top: inflation; bottom: unemployment rate. Units in y-axis: standard deviation; Units in x-axis: months. The [10%, 90%] highest posterior density intervals are shaded. Sample: June 1953 - December 2013.*

variable into contributions of various shocks in the system. In our system, we have shocks to macroeconomic variables, shocks to yield factors, and shocks to uncertainty.

First, we study the relationship between shocks to monetary policy and term premium uncertainty on inflation and unemployment. We show them in the top panels of Figure 4. Throughout the six decades in our sample, the contribution of monetary policy uncertainty to inflation is overall negative. The biggest impact happened recently since the Great Recession. Without monetary policy uncertainty, inflation would have been about 0.7% higher at the end of the sample. The size is substantial, as during this period, the economy is battling deflationary pressure. On the other hand, the impact of term premium uncertainty on inflation switched between positive and negative, and the size is small overall.

The upper right panel captures the relationship between an uncertainty shock and the unemployment rate. Overall, uncertainty shocks contributed positively to the unemployment

Figure 4: Historical decomposition




*Top: monetary policy uncertainty shocks' (blue) and term premium uncertainty shocks' (red) contributions to inflation (left) and unemployment rate (right). Bottom: inflation shocks' (blue) and unemployment rate shocks' (red) contributions to monetary policy uncertainty (left) and term premium uncertainty (right).*

rate historically, or negatively to the economy, although the importance of monetary policy uncertainty and term premium uncertainty alternates. For example, the contribution of monetary policy uncertainty peaked in the early 1980s at about 0.55%. Its contribution became negative towards the end of the sample, this might imply a lower uncertainty about monetary policy from better implementation and understanding of unconventional monetary policy. Or it could come from the mere fact that the short term interest rate is stuck at zero. In contrast, the contribution of term premium uncertainty was still positive and about 0.2 - 0.3% for the same period. The unemployment rate would have been lower if there were no uncertainty shocks to the term premium. The average contribution term premium uncertainty has on the unemployment rate across time is about 0.17%, with three peaks up to 0.7% in the early 70s, early 80s and mid 2000s. The last one is associated with Greenspan's Conundrum. This was often attributed to an increase in term premia, which is consistent with our empirical evidence that higher term premium uncertainty drove higher

unemployment rate.

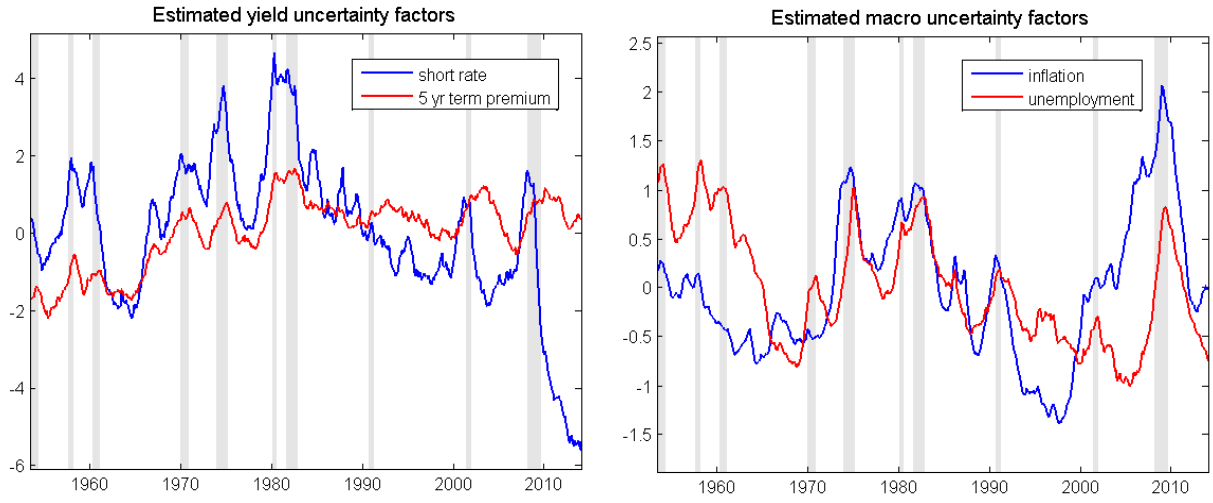
Second, we now focus on how macroeconomic variables drive uncertainty. Inflation imposes a positive contribution to both uncertainty measures, especially during the early 1980s, when inflation was at its peak. During lower inflation periods early and late in the sample, its contribution became negative. These findings are consistent with the impulse responses in the top panels of Figure 3. Shocks to the unemployment rate contribute mostly negatively to monetary policy uncertainty, especially for the prolonged periods in the 1990s to mid 2000s and after the Great Recession. In contrast, its contribution to term premium uncertainty was more positive than negative. The two red lines in the bottom panels often move in opposite directions, with a correlation as high as -0.88. This contrast again is consistent with Figure 3, and is a further evidence supporting two dimensions of uncertainty.

## 4.3 Estimates of uncertainty

We plot the monetary policy and term premium uncertainty factors in the left panel of Figure 5, together with the uncertainty factors of macroeconomic variables on the right. Both interest rate uncertainties increased in the first half of the sample, and peaked during the two recessions in the early 1980s. The short rate uncertainty displayed a decreasing trend since, although it increased again right before the two recessions in the 2000s. It finally settled down at its lowest at the end of the sample. Whereas the term premium uncertainty remained relatively stable for the second half of the sample. Inflation uncertainty was around average at the beginning of the sample, then it peaked twice at the second recession in the 1970s and two recessions in early 1980s. Then it kept going down until late 1990s. Since then, it went up, and peaked in the Great Recession at historical high before dropping to average.

**Magnitude of uncertainty** The left panel of Figure 5 provides a visual inspection of the magnitude of a one standard deviation shock as discussed in Subsection 4.1. One standard

Figure 5: Volatility factors




*Left: short rate and 5 yr term premia volatility factors. Right: inflation and unemployment volatility factors. The factors $h_t$ have been multiplied by 1200 and demeaned.*

deviation of the short rate uncertainty shock is about 0.3. The change in uncertainty leading up to the Great Recession is about 3.4 (11 times the shock size), and the change is about 4.5 before the 1980 recession (15 standard deviations). One standard deviation for the term premium uncertainty shock is about 0.14. Its hike before the Great Recession is about 12 times this magnitude, comparable to the movement for the short rate uncertainty relative to its respective standard deviation.

**Uncertainty and recession** Although the four uncertainty measures have their own dynamics, Figure 5 shows that all of them are counter-cyclical: they increase drastically before almost every recession and remain high throughout recessions; when the economy recovers, they all drop.

To illustrate this statistically, we use the following simple regression:

$$h_{jt} = \alpha + \beta \mathbb{1}_{recession,t} + u_{jt}, \tag{15}$$

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 2.0 for monetary policy uncertainty meaning that uncertainty is 2.0 units higher during recessions as opposed to expansions, and the difference is 0.4 for term premium uncertainty, 0.7 for both inflation and unemployment uncertainty. All coefficients are statistically significant with $p$-values numerically at zero.

# 5 Model comparison

## 5.1 Model specifications

**Identification and other model restrictions** In addition to the rotation restrictions imposed in Proposition 1, we impose further restrictions to achieve identification: (i) $\Sigma_m$, $\Sigma_g$, $\Sigma_h$ are lower triangular; (ii) $\Gamma_1$ is a $(G+M)\times H$ matrix with $H$ rows corresponding to $h_t$ having a scaled identity matrix $1200 * I_H$.[8] (iii) when $H > 0$, the diagonal elements of $\Sigma_m$ and $\Sigma_g$ are fixed at 1; (iv) We set $H$ elements corresponding to $h_t$ in $\Gamma_0$ to zero.

We 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 $\Sigma_h \to 0$. Finally, we estimate $\Omega$ in (14) to be a diagonal matrix. A demonstration of implementing the benchmark model is in Appendix A.2.

**Models** To understand the factor structure in volatility, on top of the macro model studied in Section 4, we compare yield only models with $M = 0$ and $H = 0, 1, 2, 3$ volatility factors. We label these models $\mathbb{H}_H$. Yields in the $\mathbb{H}_1$ model share one common volatility factor. Our choice of $\Gamma_1$ implies that the volatility factors in the $\mathbb{H}_2$ model capture the short rate and term premium volatilities. The $\mathbb{H}_3$ model adds another degree of freedom for the expected future short rate volatility.

[8] With one yield volatility factor, we normalize it to be the volatility of $r_t$. With two yield volatility factors, we normalize them to be the volatility of $r_t$ and $tp_t^{(n^*)}$. With three yield volatility factors, $\Gamma_1$ is the scaled identity matrix.

**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 B.4. The parameter estimates for the $\mathbb{H}_0$ 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 $\Phi_g$ 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 37425.4 for the $\mathbb{H}_0$ model to 37993.9 for the $\mathbb{H}_1$ model. The addition of a second volatility factor that captures movements in the 5 yr term premium adds another 100 points to the likelihood of the $\mathbb{H}_2$ model to 38095.9. Adding a third volatility factor increases the likelihood by less than 20 points to 38104.7. As the number of volatility factors increases, the number of parameters also increases. The BIC penalizes the log-likelihood for these added parameters. It selects the $\mathbb{H}_2$ model with two volatility factors as the best model for its overall fit.

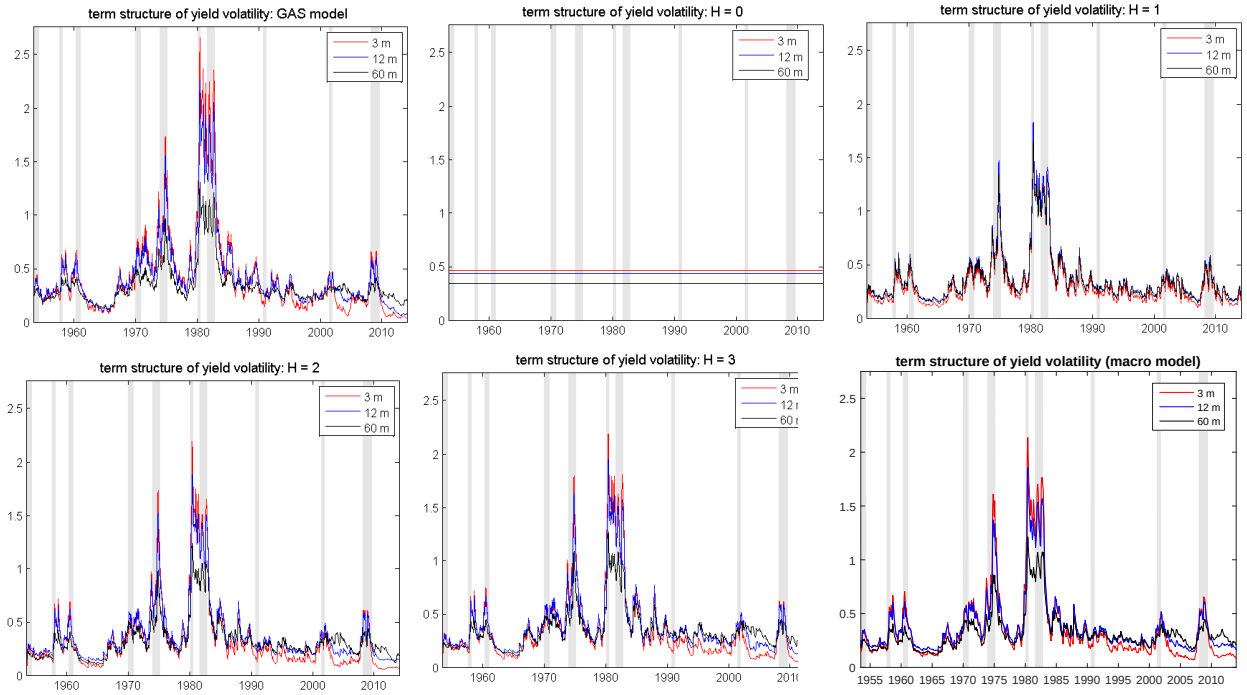
In Figure 6, we compare the estimated volatilities from the three yields only models with a reduced form description of data. The top left panel plots the conditional volatility of the $3, 12$, and 60 month yields from the generalized autoregressive score volatility model from Creal, Koopman, and Lucas(2011) and Creal, Koopman, and Lucas(2013) to capture

Table 2: Estimates for yields-only term structure models

| $\mathbb{H}_0$ | | | $\mathbb{H}_1$ | | | | $\mathbb{H}_2$ | | | | | $\mathbb{H}_3$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | log-like: | 37425.4 | | | log-like: | 37993.9 | | | | log-like: | 38095.9 | | | | | log-like: | 38104.7 |
| | BIC: | -74663.0 | | | BIC: | -75723.4 | | | | BIC: | -75842.0 | | | | | BIC: | -75766.0 |
| $\bar{\mu}_g^{\mathbb{Q}} \times 1200$ | | | $\bar{\mu}_g^{\mathbb{Q}} \times 1200$ | | | | $\bar{\mu}_g^{\mathbb{Q}} \times 1200$ | | | | | $\bar{\mu}_g^{\mathbb{Q}} \times 1200$ | | | | | |
| 11.651 | 9.260 | 2.326 | 11.931 | 9.338 | 2.534 | | 11.376 | 9.204 | 2.130 | | | 11.256 | 9.465 | 1.759 | | | |
| (0.622) | (0.669) | (0.513) | (0.729) | (0.718) | (0.533) | | (0.66) | (0.647) | (0.436) | | | (0.655) | (0.712) | (0.520) | | | |
| $\Phi_g^{\mathbb{Q}}$ | | | $\Phi_g^{\mathbb{Q}}$ | | | $\Gamma_0$ | $\Phi_g^{\mathbb{Q}}$ | | | $\Gamma_0$ | | $\Phi_g^{\mathbb{Q}}$ | | | $\Gamma_0$ | | |
| 0.660 | 0.457 | 0.044 | 0.804 | 0.255 | 0.019 | 0 | 0.795 | 0.261 | -0.022 | 0 | | 0.805 | 0.237 | -0.021 | 0 | | |
| (0.049) | (0.070) | (0.036) | (0.027) | (0.032) | (0.022) | — | (0.027) | (0.032) | (0.021) | — | | (0.028) | (0.031) | (0.022) | — | | |
| -0.071 | 1.082 | 0.053 | 0.004 | 0.975 | 0.078 | -1.032 | -0.004 | 0.989 | 0.080 | 0.763 | | -0.002 | 0.989 | 0.087 | 0 | | |
| (0.035) | (0.047) | (0.008) | (0.012) | (0.020) | (0.012 | (0.836) | (0.011) | (0.018) | (0.012) | (1.126) | | (0.011) | (0.017) | (0.014) | — | | |
| 0.056 | -0.071 | 0.969 | -0.023 | 0.042 | 0.945 | -3.64 | -0.015 | 0.027 | 0.944 | 0 | | -0.0173 | 0.027 | 0.938 | 0 | | |
| (0.036) | (0.048) | (0.009) | (0.013) | (0.021) | (0.011) | (0.892) | (0.012) | (0.019) | (0.011) | — | | (0.012) | (0.0181) | (0.014) | — | | |
| $\lambda_g^{\mathbb{Q}}$ | | | $\lambda_g^{\mathbb{Q}}$ | | | | $\lambda_g^{\mathbb{Q}}$ | | | | | $\lambda_g^{\mathbb{Q}}$ | | | | | |
| 0.996 | 0.950 | 0.766 | 0.996 | 0.950 | 0.781 | | 0.996 | 0.948 | 0.784 | | | 0.996 | 0.948 | 0.789 | | | |
| (0.0005) | (0.003) | (0.018) | (0.0005) | (0.003) | (0.015) | | (0.0005) | (0.003) | (0.015) | | | (0.0005) | (0.003) | (0.015) | | | |
| $\bar{\mu}_g \times 1200$ | | | $\bar{\mu}_f \times 1200$ | | | | $\bar{\mu}_f \times 1200$ | | | | | $\bar{\mu}_f \times 1200$ | | | | | |
| 3.989 | 3.989 | 1.185 | 2.137 | 2.137 | 0.698 | -16.513 | 1.841 | 1.841 | 0.822 | -16.767 | -18.145 | 1.989 | 1.989 | 0.717 | -16.592 | -17.026 | -18.025 |
| (0.741) | (0.741) | (0.191) | (0.256) | (0.256) | (0.220) | (0.266) | (0.272) | (0.272) | (0.257) | (0.401) | (0.304) | (0.297) | (0.297) | (0.260) | (0.446) | (0.401) | (0.337) |
| $\Phi_g$ | | | $\Phi_g$ | | | $\Phi_{gh}$ | $\Phi_g$ | | | $\Phi_{gh}$ | | $\Phi_g$ | | | $\Phi_{gh}$ | | |
| 0.868 | 0.160 | -0.001 | 0.885 | 0.128 | 0.001 | 0.001 | 0.872 | 0.141 | 0.002 | -0.008 | 0.005 | 0.859 | 0.157 | -0.003 | 0.007 | -0.037 | 0.029 |
| (0.019) | (0.028) | (0.004) | (0.015) | (0.018) | (0.003) | (0.008) | (0.017) | (0.020) | (0.005) | (0.009) | (0.007) | (0.020) | (0.023) | (0.008) | (0.014) | (0.025) | (0.017) |
| -0.019 | 1.016 | 0.00004 | -0.020 | 1.018 | -0.0001 | -0.0001 | -0.020 | 1.018 | -0.0001 | 0.001 | -0.001 | -0.019 | 1.017 | 0.0003 | -0.001 | 0.004 | -0.003 |
| (0.001) | (0.002) | (0.0003) | (0.001) | (0.001) | (0.0003) | (0.001) | (0.001) | (0.001) | (0.0005) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) | (0.002) |
| 0.069 | -0.077 | 0.966 | 0.006 | -0.002 | 0.974 | 0.005 | 0.011 | -0.008 | 0.974 | 0.002 | 0.004 | 0.014 | -0.013 | 0.978 | -0.0001 | 0.005 | 0.001 |
| (0.022) | (0.028) | (0.008) | (0.010) | (0.011) | (0.007) | (0.005) | (0.012) | (0.014) | (0.010) | (0.003) | (0.007) | (0.012) | (0.014) | (0.008) | (0.004) | (0.008) | (0.006) |
| | | | | | | $\Phi_h$ | | | | $\Phi_h$ | | | | | $\Phi_h$ | | |
| | | | | | | 0.982 | | | | 0.987 | $-5.22e^{-7}$ | | | | 0.987 | -0.001 | 0.0003 |
| | | | | | | (0.008) | | | | (0.005) | (0.0001) | | | | (0.005) | (0.007) | (0.005) |
| | | | | | | | | | | 0.00002 | 0.993 | | | | -0.001 | 0.989 | 0.003 |
| | | | | | | | | | | (0.0001) | (0.004) | | | | (0.003) | (0.007) | (0.005) |
| | | | | | | | | | | | | | | | -0.0002 | 0.002 | 0.992 |
| | | | | | | | | | | | | | | | (0.002) | (0.004) | (0.004) |
| $\lambda_g$ | | | $\lambda_f$ | | | | $\lambda_f$ | | | | | $\lambda_f$ | | | | | |
| 0.991 | 0.965 | 0.893 | 0.994 | 0.974 | 0.909 | 0.982 | 0.995 | 0.975 | 0.894 | 0.993 | 0.987 | 0.995 | 0.977 | 0.882 | 0.994 | 0.990 | 0.984 |
| (0.003) | (0.008) | (0.019) | (0.003) | (0.007) | (0.016) | (0.008) | (0.002) | (0.009) | (0.018) | (0.004) | (0.005) | (0.002) | (0.008) | (0.021) | (0.003) | (0.004) | (0.005) |
| $\Sigma_g \times 1200$ | | | $\Sigma_g^* \times 1200$ | | | $\Gamma_1 \times 1/1200$ | $\Sigma_g^* \times 1200$ | | | $\Gamma_1 \times 1/1200$ | | $\Sigma_g^* \times 1200$ | | | $\Gamma_1 \times 1/1200$ | | |
| 0.480 | 0 | 0 | 0.314 | 0 | 0 | 1 | 0.261 | 0 | 0 | 1 | 0 | 0.287 | 0 | 0 | 1 | 0 | 0 |
| (0.015) | — | — | (0.043) | | | — | (0.091) | | | — | — | (0.104) | | | — | — | — |
| 0.309 | 0.213 | 0 | 0.162 | 0.379 | 0 | 0.916 | 0.132 | 0.303 | 0 | 0.515 | 0.472 | 0.156 | 0.230 | 0 | 0 | 1 | 0 |
| (0.038) | (0.034) | — | (0.035) | (0.063) | | (0.048) | (0.048) | (0.093) | | (0.041) | (0.051) | (0.061) | (0.079) | | — | — | — |
| -0.126 | 0.007 | 0.185 | -0.037 | -0.015 | 0.215 | 0.825 | -0.037 | -0.101 | 0.130 | 0 | 1 | -0.046 | -0.084 | 0.139 | 0 | 0 | 1 |
| (0.037) | (0.036) | (0.018) | (0.025) | (0.047) | (0.026) | (0.054) | (0.021) | (0.043) | (0.041) | — | — | (0.026) | (0.039) | (0.045) | — | — | — |
| | | | $\Sigma_{gh}^* \times 1200$ | | | $\Sigma_h^* \times 1200$ | $\Sigma_{gh}^* \times 1200$ | | | $\Sigma_h^* \times 1200$ | | $\Sigma_{gh}^* \times 1200$ | | | $\Sigma_h^* \times 1200$ | | |
| | | | 0.051 | -0.015 | -0.029 | 0.348 | 0.086 | 0.058 | -0.047 | 0.371 | 0 | 0.092 | 0.042 | -0.041 | 0.367 | | |
| | | | (0.030) | (0.033) | (0.035) | (0.044) | (0.043) | (0.044) | (0.040) | (0.112) | — | (0.047) | (0.040) | (0.040) | (0.110) | | |
| | | | | | | | -0.005 | -0.052 | -0.007 | 0.035 | 0.181 | 0.045 | 0.036 | -0.029 | 0.205 | 0.159 | |
| | | | | | | | (0.024) | (0.033) | (0.021) | (0.046) | (0.057) | (0.032) | (0.033) | (0.029) | (0.069) | (0.048) | |
| | | | | | | | | | | | | -0.001 | -0.063 | -0.006 | 0.027 | 0.020 | 0.175 |
| | | | | | | | | | | | | (0.03) | (0.035) | (0.022) | (0.049) | (0.042) | (0.055) |

*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. For the models with $H = 1, 2, 3$, we report the sub-components of $\Sigma_f^* = \Sigma_f \, diag\left(\exp\left(\frac{\Gamma_0 + \Gamma_1 \bar{\mu}_h}{2}\right)\right)$.*

Figure 6: Estimated (filtered) conditional volatility of yields from six models




*Estimated conditional volatility of* $3, 12, 60$ *month yields from six different models. Top left: univariate generalized autoregressive score model; Top middle:* $\mathbb{H}_0$ *model; Top right:* $\mathbb{H}_1$ *model; Bottom left:* $\mathbb{H}_2$ *model; Bottom middle:* $\mathbb{H}_3$ *model; Bottom right: main macro model.*

this feature of the data.[9] This graph illustrates a factor structure for yield volatilities of 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 sloped downward as the volatility of short-term interest rates was higher than long-term rates. 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.

[9] 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.

The remaining panels in Figure 6 plot the (filtered) 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(2015b). 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. In the $\mathbb{H}_2$ model, the correlation between the two volatility factors is only 0.082. The $\mathbb{H}_3$ model adds a third volatility factor which is highly correlated with the short rate volatility factor at 0.90. 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 – and their volatilities into the $\mathbb{H}_2$ model selected above. The conditional volatilities from this model are plotted in the bottom right panel of Figure 6. 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.

## 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 four columns report the ratios of pricing errors for

Table 3: Pricing errors relative to the Gaussian model

| | $\mathbb{H}_0$ | $\mathbb{H}_1$ | $\mathbb{H}_2$ | $\mathbb{H}_3$ | macro |
|---|---|---|---|---|---|
| 1m | 0.3235 | 1.0893 | 1.0980 | 1.0983 | 1.0980 |
| 3m | 0.1074 | 0.5531 | 0.4873 | 0.4991 | 0.4879 |
| 12m | 0.1213 | 0.9616 | 0.9732 | 0.9923 | 0.9632 |
| 24m | 0.0716 | 0.9788 | 1.0246 | 1.0168 | 1.0193 |
| 36m | 0.0689 | 1.0001 | 1.0019 | 0.9966 | 1.0001 |
| 48m | 0.0693 | 0.9325 | 0.9854 | 0.9893 | 0.9743 |
| 60m | 0.0852 | 1.0549 | 0.9858 | 0.9721 | 1.0068 |

*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.*

the yields-only models with $H = 1, 2, 3$ and the macroeconomic model relative to the $\mathbb{H}_0$ model for the same maturity. Relative to the Gaussian $\mathbb{H}_0$ model, the biggest improvement happens for the 3 month yield, with measurement error dropping about half across models. For other maturities, the fit improves more often than not. Unlike the standard USV model which imposes restrictions on the cross-section of yields, our new models actually improve it.

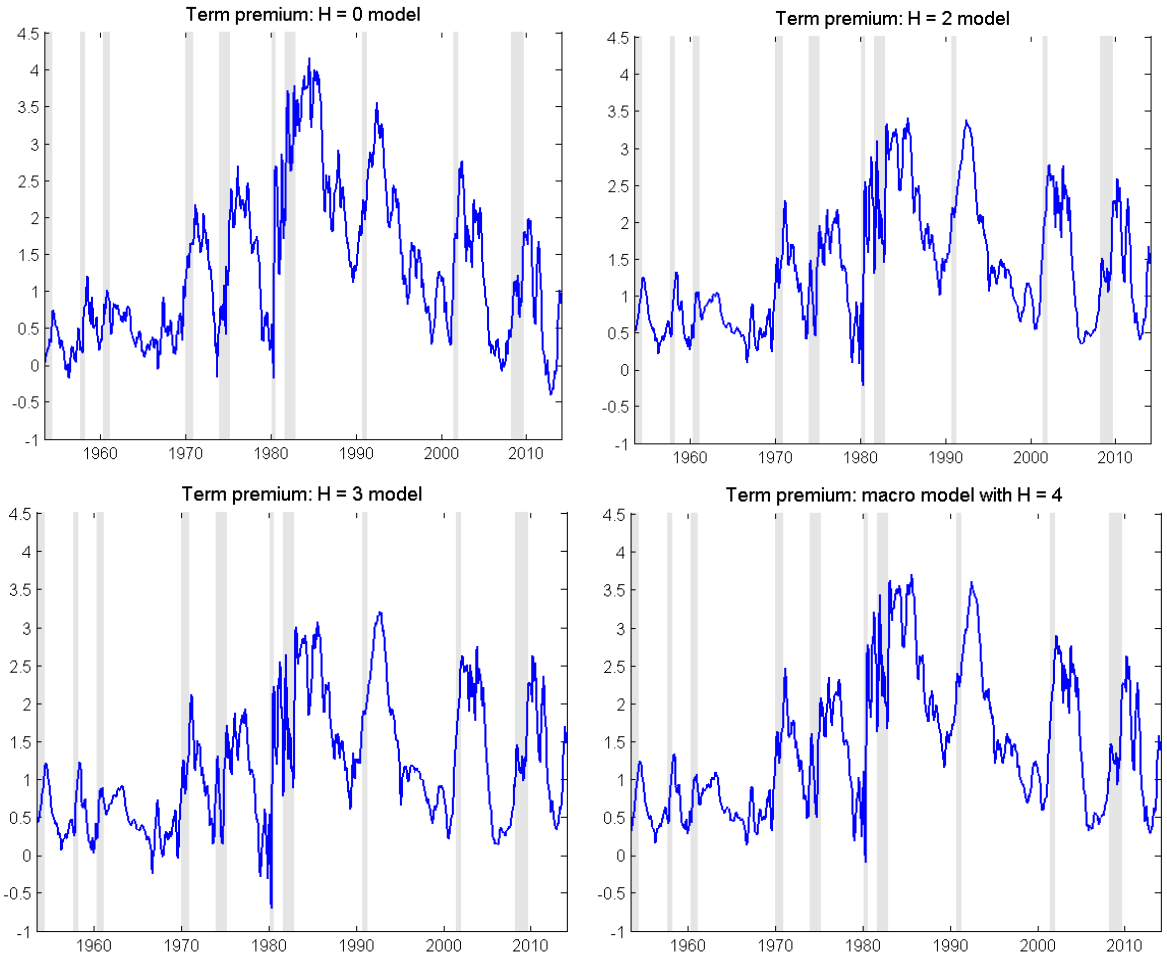
### 5.4 Term premia

Term premia and its uncertainty are one of the key aspects of our paper. Figure 7 provides estimates of 5 yr term premia for four models we estimated: $\mathbb{H}_0, \mathbb{H}_2, \mathbb{H}_3$ and the macro model. We do not plot the $\mathbb{H}_1(4)$ model to save space; it looks similar to the rest. Term premium estimates share common business cycle pattern, size and time dynamics. This validates our modeling strategy and empirical results.

## 6 Conclusion

We developed a new macro finance affine term structure model with stochastic volatilities to study the empirical importance of interest rate uncertainty. In our model, the volatility

Figure 7: Estimated 5 yr term premia from four models




*Estimated 5 yr term premia. Top left:* $\mathbb{H}_0(3)$ *model; Top right:* $\mathbb{H}_2(5)$ *model; Bottom left:* $\mathbb{H}_3(6)$ *model; Bottom right: main macro model.*

factor serves two roles: it is the volatility of the yield curve and macroeconomic variables, and it also measures uncertainty which directly interacts with macroeconomic 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.

We find that uncertainty contributes negatively to the real economy, which is consistent with what researchers find in the uncertainty literature. Unique conclusions drawn from our two dimensions of uncertainty (monetary policy uncertainty and term premium uncertainty)

include: (1) they react in opposite directions as a consequence of a positive shock to the unemployment rate. (2) The response of inflation to uncertainty shocks vary across different historical episodes.

# 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.

Andrieu, Christophe, Arnaud Doucet, and Roman Holenstein (2010) "Particle Markov chain Monte Carlo methods (with discussion)" *Journal of the Royal Statistical Society, Series B* 72, 1–33.

Baker, Scott R., Nicholas Bloom, and Steven J. Davis (2015) "Measuring economic policy uncertainty." University of Chicago, Booth School of Business, Working paper.

Bansal, Ravi, and Ivan Shaliastovich (2013) "A long-run risks explanation of predictability puzzles in bond and currency markets" *The Review of Financial Studies* 26, 1–33.

Bauer, Michael D. (2015) "Restrictions on Risk Prices in Dynamic Term Structure Models" Federal Reserve Bank of San Francisco Working Paper.

Bauer, Michael D., Glenn D. Rudebusch, and Jing Cynthia Wu (2012) "Correcting estimation bias in dynamic term structure models." *Journal of Business and Economic Statistics* 30, 454–467.

Bauer, Michael D., Glenn D. Rudebusch, and Jing Cynthia Wu (2014) "Term premia and inflation uncertainty: empirical evidence from an international panel dataset: comment." *American Economic Review* 1, 323–337.

Bekaert, Geert, Marie Hoerova, and Marco Lo Duca (2013) "Risk, uncertainty, and monetary policy." *Journal of Monetary Economics* 60, 771–788.

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 (2014) "Fluctuations in uncertainty." *Journal of Economic Perspectives* 28, 153–176.

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 (2015) "Information in the term structure of yield curve volatility." *The Journal of Finance* forthcoming.

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.

Collin-Dufresne, Pierre, Robert S. Goldstein, and Charles Jones (2009) "Can the volatility of interest rates be extracted from the cross section of bond yields? An investigation of unspanned stochastic volatility." *Journal of Financial Economics* 94, 47–66.

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 (2015a) "Bond risk premia in consumption-based models." Working paper, University of Chicago, Booth School of Business.

Creal, Drew D., and Jing Cynthia Wu (2015b) "Estimation of affine term structure models with spanned or unspanned stochastic volatility." *Journal of Econometrics* 185, 60–81.

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 US treasury market." Working paper, University of North Carolina, Department of Economics.

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 (2014) "The effect of oil price uncertainty on global real economic activity." *Journal of Money, Credit, and Banking* 46, 1113–1135.

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

Joslin, Scott (2015) "Can unspanned stochastic volatility models explain the cross section of bond volatilities?" *Management Science* forthcoming.

Joslin, Scott, Marcel Priebsch, and Kenneth J. Singleton (2014) "Risk premiums in dynamic term structure models with unspanned macro risks" *The Journal of Finance* 69, 1197–1233.

Jurado, Kyle, Sydney C. Ludvigson, and Serena Ng (2015) "Measuring uncertainty." *American Economic Review* 105, 1177–1216.

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.

Mumtaz, Haroon, and Francesco Zanetti (2013) "The impact of the volatility of monetary policy shocks." *Journal of Money, Credit, and Banking* 45, 535–558.

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.

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

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

Piazzesi, Monika, and Martin Schneider (2007) "Equilibrium yield curves." in *NBER Macroeconomics Annual 2006*, edited by Daron Acemoglu, Kenneth Rogoff, and Michael Woodford MIT Press, Cambridge, MA pages 389–442.

Primiceri, Giorgio E. (2005) "Time varying structural vector autoregressions and mon-

etary 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.

Whiteley, Nick (2010) "Discussion on Particle Markov chain Monte Carlo Methods." *Journal of the Royal Statistical Society, Series B* 72, 306–307.

Wright, Jonathan H. (2011) "Term premia and inflation uncertainty: empirical evidence from an international panel dataset." *American Economic Review* 101(4), 1514–1534.

Wu, Jing Cynthia, and Dora Fan Xia (2015) "Measuring the macroeconomic impact of monetary policy at the zero lower bound." *Journal of Money, Credit, and Banking* forthcoming.

# Appendix A Rotation of the state vector

## Appendix A.1 Proof for Proposition 1

*Proof.* The short rate, expected future short rate, and term premium are defined as

$$\begin{aligned}
r_t &= y_t^{(1)} = a_1 + b_{1,g}' g_t = a_1 + b_1' f_t \\
er_t^{(n^*)} &\equiv \frac{1}{n^*}\mathrm{E}_t\left[r_t + \ldots + r_{t+n^*-1}\right] \\
&= a_1 + \frac{b_1'}{n^*}\left[(n^*-1)I + (n^*-2)\Phi_f + \ldots + \Phi_f^{n^*-2}\right](I-\Phi_f)\,\bar{\mu}_f \\
&\quad + \frac{b_1'}{n^*}\left[I + \Phi_f + \Phi_f^2 + \ldots + \Phi_f^{n^*-1}\right] f_t \\
&\equiv c_{n^*} + d_{n^*}' f_t \\
tp_t^{(n^*)} &\equiv y_t^{(n^*)} - er_t^{(n^*)} \\
&= a_{n^*} - c_{n^*} + \left(b_{n^*} - d_{n^*}\right)' f_t
\end{aligned}$$

where $b_{n^*}' = \begin{pmatrix} 0 & b_{n^*,g}' & 0 \end{pmatrix}$, $c_{n^*} = a_1 + \frac{b_1'}{n^*}\left[(n^*-1)I + (n^*-2)\Phi_f + \ldots + \Phi_f^{n^*-2}\right](I-\Phi_f)\,\bar{\mu}_f$, $d_{n^*}' = \frac{b_1'}{n^*}\left[I + \Phi_f + \Phi_f^2 + \ldots + \Phi_f^{n^*-1}\right]$, while $a_{n^*}$ and $b_{n^*}$ are defined in (10) and (11).

From the definition of the loadings, we find

$$\begin{aligned}
d_{n^*}' &= \frac{b_1'}{n^*}\left[I + \Phi_f + \Phi_f^2 + \ldots + \Phi_f^{n^*-1}\right] \\
&= \frac{1}{n^*}\begin{pmatrix} 0 & \mathbf{e}_1' & 0 \end{pmatrix} Q\left[I + \Lambda + \Lambda^2 + \ldots + \Lambda^{n^*-1}\right] Q^{-1} \\
&= \frac{1}{n^*}\begin{pmatrix} 0 & \mathbf{e}_1' & 0 \end{pmatrix} Q\left(I - \Lambda^{n^*}\right)(I-\Lambda)^{-1} Q^{-1} \\
&= \begin{pmatrix} \underset{1\times M}{0} & \underset{1\times G}{\mathbf{e}_2'} & \underset{1\times H}{0} \end{pmatrix}
\end{aligned}$$

and hence $d_{n^*,g}' = \mathbf{e}_2'$, and

$$\begin{aligned}
b_{n^*,g}' &= \frac{1}{n^*} b_{1,g}'\left[I + \Phi_g^{\mathbb{Q}} + ... + \left(\Phi_g^{\mathbb{Q}}\right)^{n^*-1}\right] \\
&= \frac{1}{n^*}\mathbf{e}_1' Q_g^{\mathbb{Q}}\left[I + \Lambda_g^{\mathbb{Q}} + ... + \left(\Lambda_g^{\mathbb{Q}}\right)^{n^*-1}\right]\left(Q_g^{\mathbb{Q}}\right)^{-1} \\
&= \frac{1}{n^*}\mathbf{e}_1' Q_g^{\mathbb{Q}}\left(I - \left(\Lambda_g^{\mathbb{Q}}\right)^{n^*}\right)\left(I - \Lambda_g^{\mathbb{Q}}\right)^{-1}\left(Q_g^{\mathbb{Q}}\right)^{-1} \\
&= \mathbf{e}_2' + \mathbf{e}_3'
\end{aligned}$$

For the loading $c_{n^*}$ on the expected future short rate, we start with

$$c_{n^*} = a_1 + \frac{b_1'}{n^*}\left[(n^*-1)I + (n^*-2)\Phi_f + \ldots + \Phi_f^{n^*-2}\right](I-\Phi_f)\,\bar{\mu}_f$$

Let $K$ be the coefficient in front of $\bar{\mu}_f$ in the second term. Then, recognizing that this contains an arithmetico-geometric sequence, we can derive

$$\begin{aligned}
K &= \frac{b_1'}{n^*}\left[(n^*-1)I + (n^*-2)\Phi_f + \ldots + \Phi_f^{n^*-2}\right](I-\Phi_f) \\
&= \frac{b_1'}{n^*}\left[(n^*-1)QQ^{-1} + (n^*-2)Q\Lambda Q^{-1} + \ldots + \left(Q\Lambda Q^{-1}\right)^{n^*-2}\right]\left(I - Q\Lambda Q^{-1}\right) \\
&= \frac{b_1'}{n^*}Q\left[(n^*-1)I + (n^*-2)\Lambda + \ldots + \Lambda^{n^*-2}\right](I-\Lambda)\,Q^{-1} \\
&= \frac{b_1'}{n^*}Q\left[\left((n^*-1)I - \Lambda^{n^*-1}\right)(I-\Lambda) - \Lambda + \Lambda^{n^*-1}\right](I-\Lambda)^{-2}(I-\Lambda)\,Q^{-1} \\
&= \frac{b_1'}{n^*}Q\left[(n^*-1)I - n^*\Lambda + \Lambda^{n^*}\right](I-\Lambda)^{-2}(I-\Lambda)\,Q^{-1} \\
&= \frac{b_1'}{n^*}Q\left[n^*(I-\Lambda) - (I-\Lambda)^{n^*}\right](I-\Lambda)^{-1}Q^{-1} \\
&= b_1'Q\left[I - \frac{1}{n^*}(I-\Lambda)^{n^*}(I-\Lambda)^{-1}\right]Q^{-1} \\
&= b_1'Q\left[I - \tilde{\Lambda}\right]Q^{-1} \\
&= b_1' - b_1'Q\tilde{\Lambda}Q^{-1}
\end{aligned}$$

Then, we use conditions 1 and 3 of the proposition which give

$$K = \underset{1\times M}{0}\quad \underset{1\times G}{\mathbf{e}_1'}\quad \underset{1\times H}{0} - \underset{1\times M}{0}\quad \underset{1\times G}{\mathbf{e}_2'}\quad \underset{1\times H}{0}$$

Therefore, conditions 2 and 4 imply

$$c_{n^*} = a_1 + K\bar{\mu}_f = 0$$

We start with the definition of $a_{n^*}$ given by

$$\begin{aligned}
a_{n^*} &= \delta_0 - \frac{1}{2n^*}\left[b_{1,g}'\Sigma_g^{\mathbb{Q}}\Sigma_g^{\mathbb{Q},\prime}b_{1,g} + 2^2 b_{2,g}'\Sigma_g^{\mathbb{Q}}\Sigma_g^{\mathbb{Q},\prime}b_{2,g} + \ldots + (n^*-1)^2 b_{n^*-1,g}'\Sigma_g^{\mathbb{Q}}\Sigma_g^{\mathbb{Q},\prime}b_{n^*-1,g}\right] \\
&\quad + \frac{1}{n^*}\left[b_{1,g} + 2b_{2,g} + \ldots + (n^*-1)\,b_{n^*-1,g}\right]\left(I - \Phi_g^{\mathbb{Q}}\right)\bar{\mu}_g^{\mathbb{Q}}
\end{aligned}$$

First, we can write $b'_{n,g}$ for any $n$ as

$$\begin{aligned} b'_{n,g} &= \frac{1}{n}\delta'_{1,g}\left[I + \Phi_g^{\mathbb{Q}} + \ldots + \left(\Phi_g^{\mathbb{Q}}\right)^{n-1}\right] \\ &= \frac{1}{n}\delta'_{1,g}Q^{\mathbb{Q}}\left[I + \Lambda_g^{\mathbb{Q}} + \ldots + \left(\Lambda_g^{\mathbb{Q}}\right)^{n-1}\right](Q^{\mathbb{Q}})^{-1} \\ &= \frac{1}{n}\delta'_{1,g}Q^{\mathbb{Q}}\left[I - \left(\Lambda_g^{\mathbb{Q}}\right)^{n}\right]\left[I - \Lambda_g^{\mathbb{Q}}\right]^{-1}(Q^{\mathbb{Q}})^{-1} \end{aligned}$$

Let $V = \frac{1}{n^*}\left[b_{1,g} + 2b_{2,g} + \ldots + (n^*-1)\,b_{n^*-1,g}\right]\left(I - \Phi_g^{\mathbb{Q}}\right)$. We write this as

$$\begin{aligned} V &= \frac{1}{n^*}\left[b_{1,g} + 2b_{2,g} + \ldots + (n^*-1)\,b_{n^*-1,g}\right]\left(I - \Phi_g^{\mathbb{Q}}\right) \\ &= \frac{1}{n^*}\delta'_{1,g}Q_g^{\mathbb{Q}}\left[\left(I - \Lambda_g^{\mathbb{Q}}\right) + \left(I - \left(\Lambda_g^{\mathbb{Q}}\right)^2\right) + \ldots + \left(I - \left(\Lambda_g^{\mathbb{Q}}\right)^{n^*-1}\right)\right]\left(I - \Lambda_g^{\mathbb{Q}}\right)^{-1}\left(I - \Lambda_g^{\mathbb{Q}}\right)\left(Q_g^{\mathbb{Q}}\right)^{-1} \\ &= \frac{1}{n^*}\delta'_{1,g}Q_g^{\mathbb{Q}}\left[(n^*-1)I - \left(\Lambda_g^{\mathbb{Q}} + \left(\Lambda_g^{\mathbb{Q}}\right)^2 + \ldots + \left(\Lambda_g^{\mathbb{Q}}\right)^{n^*-1}\right)\right]\left(Q_g^{\mathbb{Q}}\right)^{-1} \\ &= \frac{1}{n^*}\delta'_{1,g}Q_g^{\mathbb{Q}}\left[(n^*-1)I - \Lambda_g^{\mathbb{Q}}\left(I - \left(\Lambda^{\mathbb{Q}}\right)^{n^*-1}\right)\left(I - \Lambda^{\mathbb{Q}}\right)^{-1}\right]\left(Q_g^{\mathbb{Q}}\right)^{-1} \\ &= \frac{1}{n^*}\delta'_{1,g}Q_g^{\mathbb{Q}}\left[(n^*-1)I - \left(\Lambda_g^{\mathbb{Q}} - I + I - \left(\Lambda^{\mathbb{Q}}\right)^{n^*}\right)\left(I - \Lambda^{\mathbb{Q}}\right)^{-1}\right]\left(Q_g^{\mathbb{Q}}\right)^{-1} \\ &= \frac{1}{n^*}\delta'_{1,g}Q_g^{\mathbb{Q}}\left[(n^*-1)I - \left(\Lambda_g^{\mathbb{Q}} - I\right)\left(I - \Lambda_g^{\mathbb{Q}}\right)^{-1} - \left(I - \left(\Lambda^{\mathbb{Q}}\right)^{n^*}\right)\left(I - \Lambda^{\mathbb{Q}}\right)^{-1}\right]\left(Q_g^{\mathbb{Q}}\right)^{-1} \\ &= \frac{1}{n^*}\delta'_{1,g}Q_g^{\mathbb{Q}}\left[n^*I - n^*\tilde{\Lambda}^{\mathbb{Q}}\right]\left(Q_g^{\mathbb{Q}}\right)^{-1} \\ &= \mathrm{e}'_1 Q_g^{\mathbb{Q}}\left[I - \tilde{\Lambda}^{\mathbb{Q}}\right]\left(Q_g^{\mathbb{Q}}\right)^{-1} \\ &= \mathrm{e}'_1 - \mathrm{e}'_1 Q_g^{\mathbb{Q}}\tilde{\Lambda}^{\mathbb{Q}}\left(Q_g^{\mathbb{Q}}\right)^{-1} \\ &= \mathrm{e}'_1 - (\mathrm{e}'_2 + \mathrm{e}'_3) \end{aligned}$$

where the last step uses condition 5. Hence, conditions 2 and 6 imply

$$\begin{aligned} a_{n^*} &= -\frac{1}{2n^*}\left[b'_{1,g}\Sigma_g^{\mathbb{Q}}\Sigma_g^{\mathbb{Q},\prime}b_{1,g} + 2^2 b'_{2,g}\Sigma_g^{\mathbb{Q}}\Sigma_g^{\mathbb{Q},\prime}b_{2,g} + \ldots + (n^*-1)^2\, b'_{n^*-1,g}\Sigma_g^{\mathbb{Q}}\Sigma_g^{\mathbb{Q},\prime}b_{n^*-1,g}\right] \\ &\quad +(\mathrm{e}'_1 - (\mathrm{e}'_2 + \mathrm{e}'_3))\bar{\mu}_g^{\mathbb{Q}} = 0 \end{aligned}$$

where the Jensen's inequality term referred to in the proposition is the expression in the first line.

Collecting each of the terms, we find

$$\begin{pmatrix} r_t \\ er_t^{(n^*)} \\ tp_t^{(n^*)} \end{pmatrix} = \begin{pmatrix} a_1 \\ c_{n^*} \\ a_{n^*} - c_{n^*} \end{pmatrix} + \begin{pmatrix} b'_{1,g} \\ d'_{n^*,g} \\ b'_{n^*,g} - d'_{n^*,g} \end{pmatrix} g_t$$
$$= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \mathbf{e}'_1 \\ \mathbf{e}'_2 \\ \mathbf{e}'_3 \end{pmatrix} g_t = g_t$$

□

## Appendix A.2 Implementation for our benchmark model

In this section, we discuss how we parameterize: (i) the vectors $\bar{\mu}_f$ and $\bar{\mu}_g^{\mathbb{Q}}$; (ii) the matrices $\Phi_f$ and $\Phi_g^{\mathbb{Q}}$; (iii) the matrices $\Gamma_0$ and $\Gamma_1$ in (3). In our empirical work, we set $n^* = 60$, and for our benchmark model in Section 4, $M = 2$, $G = 3$ and $H = 4$. The state factor for yields is ordered by $g_t = \left(r_t \; er_t^{(n^*)} \; tp_t^{(n^*)}\right)'$.

The vector $\bar{\mu}_f = \left(\bar{\mu}'_m, \bar{\mu}'_g, \bar{\mu}'_h\right)'$ and only $\bar{\mu}_g$ is restricted. The $G \times 1$ vector $\bar{\mu}_g$ only has $G - 1$ free parameters given by an unrestricted vector $\bar{\mu}_g^u$. We can therefore write $\bar{\mu}_g = M_1 \bar{\mu}_g^u$, where for the main model $M_1$ is given by

$$M_1 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Similarly, the $G \times 1$ vector $\bar{\mu}_g^{\mathbb{Q}}$ only has $G - 1$ free parameters given by an unrestricted vector $\bar{\mu}_g^{\mathbb{Q},u}$. We can write their relationship as $\bar{\mu}_g^{\mathbb{Q}} = M_0^{\mathbb{Q}} + M_1^{\mathbb{Q}} \bar{\mu}_g^{\mathbb{Q},u}$ where, for the main model, the vector and matrix are given by

$$M_0^{\mathbb{Q}} = \begin{pmatrix} J.I. \\ 0 \\ 0 \end{pmatrix} \qquad M_1^{\mathbb{Q}} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

The top element $J.I.$ denotes the Jensen's inequality term from Appendix A.1.

For estimation, we assume the eigenvalues are real and distinct. Following Proposition 1, the matrix $\Phi_f$ has $F^2 - F - (M + G)^2$ free parameters and the matrix $\Phi_g^{\mathbb{Q}}$ has $G^2 - G$ free parameters. We impose these

restrictions on the matrices of eigenvectors $Q$ and $Q_g^{\mathbb{Q}}$ in the benchmark case as follows

$$
\begin{aligned}
Q &= \begin{pmatrix}
1 & q_{12} & q_{13} & q_{14} & q_{15} & q_{16} & q_{17} & q_{18} & q_{19} \\
q_{21} & 1 & q_{23} & q_{24} & q_{25} & q_{26} & q_{27} & q_{28} & q_{29} \\
q_{31} & q_{32} & 1 & 1/\tilde{\lambda}_4 & q_{35} & q_{36} & q_{37} & q_{38} & q_{39} \\
q_{31}\tilde{\lambda}_1 & q_{32}\tilde{\lambda}_2 & \tilde{\lambda}_3 & 1 & q_{35}\tilde{\lambda}_5 & q_{36}\tilde{\lambda}_6 & q_{37}\tilde{\lambda}_7 & q_{38}\tilde{\lambda}_8 & q_{39}\tilde{\lambda}_9 \\
q_{51} & q_{52} & q_{53} & q_{54} & 1 & q_{56} & q_{57} & q_{58} & q_{59} \\
0 & 0 & 0 & 0 & 0 & 1 & q_{67} & q_{68} & q_{69} \\
0 & 0 & 0 & 0 & 0 & q_{76} & 1 & q_{78} & q_{79} \\
0 & 0 & 0 & 0 & 0 & q_{86} & q_{87} & 1 & q_{89} \\
0 & 0 & 0 & 0 & 0 & q_{96} & q_{97} & q_{98} & 1
\end{pmatrix} \\
Q_g^{\mathbb{Q}} &= \begin{pmatrix}
1 & 1 & 1 \\
\tilde{\lambda}_1^{\mathbb{Q}} - q_{31}^{\mathbb{Q}} & \tilde{\lambda}_2^{\mathbb{Q}} - q_{32}^{\mathbb{Q}} & \tilde{\lambda}_3^{\mathbb{Q}} - q_{33}^{\mathbb{Q}} \\
q_{31}^{\mathbb{Q}} & q_{32}^{\mathbb{Q}} & q_{33}^{\mathbb{Q}}
\end{pmatrix}
\end{aligned}
\tag{A.1}
$$

where $\tilde{\lambda}_i$ and $\tilde{\lambda}_i^{\mathbb{Q}}$ are the diagonal elements of $\tilde{\Lambda}$ and $\tilde{\Lambda}_g^{\mathbb{Q}}$. Given a matrix, its eigenvectors are identified up to their scale. Therefore, each column of $Q$ requires one restriction for identification. We set the diagonal elements $q_{ii} = 1$. The normalization assumption for $Q_g^{\mathbb{Q}}$ is $q_{1i}^{\mathbb{Q}} = 1$, which then facilitates us to impose restrictions directly on $\Phi_g^{\mathbb{Q}}$ for actual implementation below. There are no free parameters in the 4th row of $Q$ and the second row of $Q_g^{\mathbb{Q}}$ due to Proposition 1. Finally, we note that the eigenvectors in the bottom left $H \times (M+G)$ block of $Q$ are all equal to zero due to the assumption that the level of $m_t$ and $g_t$ do not enter the conditional mean of $h_{t+1}$ in (4).

We allow for a factor structure in the volatility in (3) through the vector $\Gamma_0$ and matrix $\Gamma_1$. In our benchmark model, we set

$$
\Gamma_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \gamma_{0,4} \\ 0 \end{pmatrix}, \qquad \Gamma_1 = \begin{pmatrix}
1200 & 0 & 0 & 0 \\
0 & 1200 & 0 & 0 \\
0 & 0 & 1200 & 0 \\
0 & 0 & \gamma_{1,43} & \gamma_{1,44} \\
0 & 0 & 0 & 1200
\end{pmatrix}
$$

where $\gamma_{0,4}, \gamma_{1,43}, \gamma_{1,44}$ are estimated parameters. For estimation, we scale the elements of $\Gamma_1$ by 1200 so that $h_t$ has approximately the same scale as $m_t$ and $g_t$.

**Restrictions on $\Phi_g^{\mathbb{Q}}$** As we have a better understanding of the autoregressive matrix $\Phi_g^{\mathbb{Q}}$ itself rather than its eigenvector matrix $Q_g^{\mathbb{Q}}$, we demonstrate here how to impose restrictions on $\Phi_g^{\mathbb{Q}}$ to achieve the restrictions described in (A.1). We parameterize $\Phi_g^{\mathbb{Q}}$ in terms of 3 eigenvalues and 3 elements in the first row of $\Phi_g^{Q}$. Then, we can solve the three elements of $Q_g^{\mathbb{Q}}$ in the third row as follows.

$$q_{3,i}^{\mathbb{Q}} = \frac{\lambda_i^{\mathbb{Q}} - \phi_{11}^{\mathbb{Q}} - \phi_{12}^{\mathbb{Q}}\tilde{\lambda}_i^{\mathbb{Q}}}{\phi_{13}^{\mathbb{Q}} - \phi_{12}^{\mathbb{Q}}}$$

and the second row of $Q_g^{\mathbb{Q}}$ is

$$q_{2,i}^{\mathbb{Q}} = \tilde{\lambda}_i^{\mathbb{Q}} - q_{3,i}^{\mathbb{Q}}$$

for $i = 1, 2, 3$, $\lambda_i^{\mathbb{Q}}$ is the $i^{th}$ eigenvalue in $\Lambda^{\mathbb{Q}}$, and $\phi_{i,j}^{\mathbb{Q}}$ is the $(i, j)$ component of $\Phi_g^{\mathbb{Q}}$ . Knowing $Q_g^{\mathbb{Q}}$, we then can solve for the remaining values of $\Phi_g^{\mathbb{Q}} = Q_g^{\mathbb{Q}}\Lambda_g^{\mathbb{Q}}Q_g^{\mathbb{Q},-1}$.

# 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 $t + k$. 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 $\theta$; (ii) drawing the volatilities $h_{0:T}$ conditional on $g_{1:T}$ and $\theta$; (iii) and then drawing the parameters of the model $\theta$. The MCMC algorithm is designed to minimize the amount that we condition on the latent variables $g_{1:T}$ by using the Kalman filter to marginalize over them. And, it draws the latent variables $h_{0:T}$ in large blocks using the particle Gibbs sampler, see Andrieu, Doucet, and Holenstein(2010). We will use two different state space representations as described in Subsection 3.1.

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

For linear, Gaussian models, we use the following state space form

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

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

with $x_1 \sim \mathrm{N}(\underline{x}_1, P_1)$. The intercept $A$ in (14) is a linear function of $\bar{\mu}_g^{\mathbb{Q}}$. We write it as $A = A_0 + A_1 \times \bar{\mu}_g^{\mathbb{Q}}$ where $\bar{\mu}_g^{\mathbb{Q}} = M_0^{\mathbb{Q}} + M_1^{\mathbb{Q}} \bar{\mu}_g^{\mathbb{Q},u}$, and $\bar{\mu}_g^{\mathbb{Q},u}$ is the vector of unrestricted parameters. The vector $M_0$ and matrix $M_1$ are discussed in Appendix A.2. The vector $A_0$ and matrix $A_1$ are determined by the bond loading recursions.

We place the unconditional means of the macro variables $\bar{\mu}_m$, the unconditional means of the yield factors $\bar{\mu}_g^u$, and the unconditional mean of the yield factors under $\mathbb{Q}$ given by $\bar{\mu}_g^{\mathbb{Q},u}$ in the state vector. Note that $\bar{\mu}_f^u = \left(\bar{\mu}_m', \bar{\mu}_g^{u,\prime}\right)'$. We draw them jointly with the yield factors $\bar{g}_{1:T}$ using simulation smoothing algorithms (forward-filtering backward sampling). We 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

$$Y_t = \begin{pmatrix} m_t \\ y_t \\ h_t \end{pmatrix}, \quad Z_t = \begin{pmatrix} I & 0 & 0 & 0 & 0 & 0 \\ 0 & B & 0 & 0 & 0 & A_1 M_1^{\mathbb{Q}} \\ 0 & 0 & I & 0 & 0 & 0 \end{pmatrix}, \quad d_t = \begin{pmatrix} 0 \\ A_0 + A_1 M_0^{\mathbb{Q}} \\ 0 \end{pmatrix} \quad H_t = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \Omega & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$x_t = \begin{pmatrix} m_t \\ g_t \\ h_t \\ \bar{\mu}_m \\ \bar{\mu}_g^u \\ \bar{\mu}_g^{\mathbb{Q},u} \end{pmatrix}, \quad C_t = I,$$

$$c_t = \begin{pmatrix} \bar{\Phi}_{mh}\bar{\mu}_h \\ \bar{\Phi}_{mg}\bar{\mu}_h \\ \bar{\Phi}_{hh}\bar{\mu}_h \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad T_t = \begin{pmatrix} \Phi_m & \Phi_{mg} & \Phi_{mh} & \bar{\Phi}_{mm} & \bar{\Phi}_{mg} & 0 \\ \Phi_{gm} & \Phi_g & \Phi_{gh} & \bar{\Phi}_{mg} & \bar{\Phi}_{gg} & 0 \\ 0 & 0 & \Phi_h & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & 0 & I \end{pmatrix}, \quad R_t = \begin{pmatrix} \Sigma_m D_{m,t} & 0 & 0 \\ \Sigma_{gm} D_{m,t} & \Sigma_g D_{g,t} & 0 \\ \Sigma_{hm} & \Sigma_{hg} & \Sigma_h \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

where the matrices $(\bar{\Phi}_{mm}, \bar{\Phi}_{mg}, \bar{\Phi}_{mg}, \bar{\Phi}_{gg}, \bar{\Phi}_{mh}, \bar{\Phi}_{mg}, \bar{\Phi}_{hh})$ are defined from the relation

$$\bar{\mu}_f = L_1 \bar{\mu}_f^u, \qquad \begin{pmatrix} \bar{\Phi}_{mm} & \bar{\Phi}_{mg} & \bar{\Phi}_{mh} \\ \bar{\Phi}_{gm} & \bar{\Phi}_{gg} & \bar{\Phi}_{gh} \\ 0 & 0 & \bar{\Phi}_{hh} \end{pmatrix} = (I - \Phi_f) L_1.$$

where $L_1$ is a selection matrix of zeros and ones that imposes the restriction on $\bar{\mu}_g$ discussed in Appendix A.2. The priors for the parameters are $\bar{\mu}_m \sim \mathrm{N}\left(\underline{\bar{\mu}_m}, V_{\bar{\mu}_m}\right)$, $\bar{\mu}_g^u \sim \mathrm{N}\left(\underline{\bar{\mu}_g^u}, V_{\bar{\mu}_g^u}\right)$, and $\bar{\mu}_g^{\mathbb{Q},u} \sim \mathrm{N}\left(\underline{\bar{\mu}_g^{\mathbb{Q},u}}, V_{\bar{\mu}_g^{\mathbb{Q},u}}\right)$. The initial conditions for $x_1 \sim \mathrm{N}\left(\underline{x}_1, P_1\right)$ are

$$\underline{x}_1 = \begin{pmatrix} \Phi_{mh} h_0 \\ \Phi_{gh} h_0 \\ h_0 \\ \underline{\bar{\mu}}_m \\ \underline{\bar{\mu}}_g^u \\ \underline{\bar{\mu}}_g^{\mathbb{Q},u} \end{pmatrix}$$

$$C_1 = \begin{pmatrix} \Sigma_m D_{m,0}^2 \Sigma_m' & \Sigma_m D_{m,0}^2 \Sigma_{gm}' & \Sigma_m D_{m,0} \Sigma_{hm}' & 0 & 0 & 0 \\ \Sigma_{gm} D_{m,0}^2 \Sigma_m' & \Sigma_{gm} D_{m,0}^2 \Sigma_{gm}' + \Sigma_g D_{g,0}^2 \Sigma_g' & \Sigma_g D_{g,0} D_{m,0} \Sigma_{hm}' + \Sigma_g D_{g,0} \Sigma_{gh}' & 0 & 0 & 0 \\ \Sigma_{hm} D_{m,0} \Sigma_m' & \Sigma_{hm} D_{m,0} D_{g,0} \Sigma_g' + \Sigma_{hg} D_{g,0} \Sigma_g' & \Sigma_{hm} \Sigma_{hm}' + \Sigma_{hg} \Sigma_{hg}' + \Sigma_h \Sigma_h' & 0 & 0 & 0 \\ 0 & 0 & 0 & V_{\bar{\mu}_m} & 0 & 0 \\ 0 & 0 & 0 & 0 & V_{\bar{\mu}_g^u} & 0 \\ 0 & 0 & 0 & 0 & 0 & V_{\bar{\mu}_g^{\mathbb{Q},u}} \end{pmatrix}.$$

with $P_1 = R_1 C_1 R_1'$ and $R_1 = I$.

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

Conditional on the draw of $g_{1:T}$, the observation equations are

$$m_{t+1} = \mu_m + \Phi_m m_t + \Phi_{mg} g_t + \Phi_{mh} h_t + \Sigma_m D_{m,t} \varepsilon_{m,t+1} \tag{B.3}$$

$$g_{t+1} = \mu_g + \Phi_{gm} m_t + \Phi_g g_t + \Phi_{gh} h_t + \Sigma_{gm} D_{m,t} \varepsilon_{m,t+1} + \Sigma_g D_{g,t} \varepsilon_{g,t+1}. \tag{B.4}$$

The transition equation is

$$h_{t+1} = \mu_h + \Phi_h h_t + \Sigma_{hm} \varepsilon_{m,t+1} + \Sigma_{hg} \varepsilon_{g,t+1} + \Sigma_h \varepsilon_{h,t+1}. \tag{B.5}$$

We use this state space representation to draw $h_{0:T}$ using the particle Gibbs sampler.

## Appendix B.2 Drawing the volatilities by particle Gibbs sampling

The full conditional distribution of the volatilities $p\left(h_{0:T}|g_{1:T}, m_{1:T}; \theta\right)$ is a non-standard distribution. We use a particle Gibbs sampler to draw paths of the state from this distribution in large blocks, which improves the mixing of the Markov chain. The particle Gibbs sampler runs a particle filter at each iteration of the MCMC algorithm to build a discrete approximation of the continuous distribution $p\left(h_{0:T}|g_{1:T}, m_{1:T}; \theta\right)$. At each date $t$, the marginal filtering distribution of the state variable is $p\left(h_t|g_{1:T}, m_{1:T}; \theta\right)$. The particle filter approximates this marginal through a collection of $J$ particles $\left\{h_t^{(j)}, \hat{w}_t^{(j)}\right\}_{j=1}^{J}$ where $h_t^{(j)}$ is a point on the support of the distribution and $\hat{w}_t^{(j)}$ is the probability mass at that point. Collecting the particles for all dates $t = 1, ..., T$, the particle filter approximates the joint distribution $\left\{h_{0:T}^{(j)}, \hat{w}_{0:T}^{(j)}\right\}_{j=1}^{J} \approx p\left(h_{0:T}|g_{1:T}, m_{1:T}; \theta\right)$. The PG sampler draws one path of the state variables from this discrete approximation. As the number of particles $M$ goes to infinity, the PG sampler draws from the exact full conditional distribution.

The particle Gibbs sampler requires a small modification of a standard particle filter. The particle Gibbs sampler requires that the pre-existing path $h_{0:T}^{(*)} = \left(h_0^{(*)}, h_1^{(*)}, \ldots, h_T^{(*)}\right)$ that was sampled at the last iteration have positive probability of being sampled again. This means that the path $h_{0:T}^{(*)}$ must survive the resampling step of the particle filtering algorithm. Instead of implementing a standard resampling algorithm, Andrieu, Doucet, and Holenstein(2010) describe a conditional resampling algorithm that needs to be implemented. Other than this, the particle filter is standard and proceeds as follows.

At time $t = 0$.

- Set $h_0^{(1)} = h_0^{(*)}$. For $j = 2, \ldots, J$, draw $h_0^{(j)} \sim p\left(h_0; \theta\right)$ and set $\hat{w}_0^{(j)}$.

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

- Set $h_t^{(1)} = h_t^{(*)}$. For $j = 2, \ldots, J$, draw from a proposal distribution $h_t^{(j)} \sim q\left(h_t|m_t, g_t, h_{t-1}^{(j)}; \theta\right)$
- For $j = 1, \ldots, J$ calculate the importance weights

$$w_t^{(j)} \propto w_{t-1}^{(j)} \frac{p\left(g_t, m_t|m_{t-1}, g_{t-1}, h_t^{(j)}, h_{t-1}^{(j)}; \theta\right) p\left(h_t^{(j)}|h_{t-1}^{(j)}; \theta\right)}{q\left(h_t^{(j)}|m_t, g_t, h_{t-1}^{(j)}; \theta\right)}$$

- For $j = 1, \ldots, J$ normalize the weights $\hat{w}_t^{(j)} = \frac{w_t^{(j)}}{\sum_{k=1}^{J} w_t^{(k)}}$
- Conditionally resample the particles $\left\{h_t^{(j)}\right\}_{j=1}^{J}$ with probabilities $\left\{\hat{w}_t^{(j)}\right\}_{j=1}^{J}$. In this step, the first particle $h_t^{(1)}$ always gets resampled and may be randomly duplicated.

At each time step of the algorithm, we store the particles and their weights $\left\{h_{0:T}^{(j)}, \hat{w}_{0:T}^{(j)}\right\}_{j=1}^{J}$. We draw a path of the state variables from this discrete distribution according to an algorithm proposed by Whiteley(2010).

At time $t = T$, draw $h_T^{(*)} = h_T^{(j)}$ with probability $\hat{w}_T^{(j)}$. Then, for $t = T-1, \ldots, 0$, we draw recursively backwards

- For $j = 1, \ldots, J$, calculate the backwards weights $w_{t|T} \propto \hat{w}_t^{(j)} p(h_{t+1}^{(*)} | h_t^{(j)}; \theta)$.
- For $j = 1, \ldots, J$ normalize the weights $\hat{w}_{t|T}^{(j)} = \frac{w_{t|T}^{(j)}}{\sum_{j=1}^{J} w_{t|T}^{(j)}}$.
- Draw $h_t^{(*)} = h_t^{(j)}$ with probability $\hat{w}_{t|T}^{(j)}$.

The algorithm produces a draw $h_{0:T}^{(*)} = \left( h_0^{(*)}, \ldots, h_T^{(*)} \right)$ from the full conditional distribution.

In practice, when the dimension of $H$ is large, we draw an individual path of the volatilities conditional on the other paths. Let $h_{i,t}$ for $i = 1, \ldots, H$ denote the set of stochastic volatility state variables. Specifically, we draw path $h_{i,0:T}$ conditional on the remaining paths $h_{k,0:T}$ for all $k \neq i$. This remains a valid particle Gibbs sampler. In the paper, we use $J = 300$ particles and choose the transition density $p(h_t | h_{t-1}; \theta)$ as the proposal $q\left(h_t | m_t, g_t, h_{t-1}; \theta\right)$.

### Appendix B.2.1 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.2.2, we will repeatedly apply the independence Metropolis Hastings (IMH) algorithm along the lines of Chib and Greenberg(1994) and Chib and Ergashev(2009), in a combination with the Kalman filter. Here is how it works. Let $Y_t = \left(m_t' \; y_t' \; h_t'\right)'$ . Suppose we separate the parameter vector $\theta = (\psi, \psi^-)$ and we want to draw a subset of the parameters $\psi$ conditional on the remaining parameters $\psi^-$.

- 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 the Kalman filter. Let $\hat{\psi}$ be the posterior mode and $H_{\psi}^{-1}$ 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 $H_{\psi}^{-1}$, and 5 degrees of freedom.
- The proposal $\psi^*$ is accepted with probability $\alpha = \frac{p\left(Y_{1:T} | \psi^*, \psi^-\right) p(\psi^*) q\left(\psi^{(j-1)}\right)}{p\left(Y_{1:T} | \psi^{(j-1)}, \psi^-\right) p\left(\psi^{(j-1)}\right) q(\psi^*)}$.

### Appendix B.2.2 MCMC algorithm

We use the notation $\theta^{(-)}$ to denote all the remaining parameters in $\theta$ other than the parameters being drawn in that step. Let $q_f$ and $\phi_g^{\mathbb{Q}}$ denote the free parameters in the matrix of eigenvectors $Q$ and $\Phi_g^{\mathbb{Q}}$, respectively. Our MCMC algorithm proceeds as follows:

1. **Draw** $\lambda_f, q_f, q_g^{\mathbb{Q}}, \Sigma_g$: Conditional on $h_{0:T}$ and the remaining parameters of the model, write the model in state space form as in Appendix B.1.1. Draw parameters listed below using the IMH algorithm as explained in Appendix B.2.1.

   - Draw the elements of $\Omega$. For any diagonal elements in $\Omega$, we use a proposal distribution that draws them in logarithms.
   - Draw the free parameters in $\Sigma_g$: note the diagonal elements of $\Sigma_g$ are fixed.
   - Draw $\lambda_f, q_f, \phi_g^{\mathbb{Q}}$.[10]
   - Draw the free parameters in $\bar{\mu}_h, \Gamma_0, \Gamma_1$.

2. **Draw** $\left(g_{1:T}, \bar{\mu}_m, \bar{\mu}_g^u, \bar{\mu}_g^{\mathbb{Q},u}, \lambda_g^{\mathbb{Q}}\right)$ **jointly in one block.**

   - **Draw** $\lambda_g^{\mathbb{Q}}$ **from** $p\left(\lambda_g^{\mathbb{Q}} | Y_{1:T}, h_{0:T}, \theta^{(-)}\right)$**:** 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 $\lambda_g^{\mathbb{Q}}$ using the IMH algorithm as explained above.
   - **Draw** $\left(g_{1:T}, \bar{\mu}_m, \bar{\mu}_g^u, \bar{\mu}_g^{\mathbb{Q},u}\right)$ **jointly from** $p\left(g_{1:T}, \bar{\mu}_m, \bar{\mu}_g^u, \bar{\mu}_g^{\mathbb{Q},u} | Y_{1:T}, h_{0:T}, \theta^{(-)}, \lambda_g^{\mathbb{Q}}\right)$**:** Conditional on $\lambda_g^{\mathbb{Q}}$, draw $(g_{1:T}, \bar{\mu}_m, \bar{\mu}_g^u, \bar{\mu}_g^{\mathbb{Q},u})$ using the simulation smoother of Durbin and Koopman(2002).

3. **Draw** $h_{0:T}$**:** Draw the paths of the volatilities using the particle Gibbs sampler as explained in Appendix B.2.

4. **Draw** $(\Sigma_m, \Sigma_{gm}, \Sigma_{hm}, \Sigma_{hg}, \Sigma_h)$**:** Conditional on $g_{1:T}$ and $h_{0:T}$, the full conditional distribution for these parameters are known in closed form. The matrices $\Sigma_{gm}, \Sigma_{hm}, \Sigma_{hg}$ can be drawn recursively from the regression models (B.4) and (B.5) once we treat the errors $\varepsilon_{mt}$ and $\varepsilon_{gt}$ as observable. The full conditional distribution of the matrix $\Sigma_h \Sigma_h'$ is an inverse Wishart distribution.

## Appendix B.3 Particle filter for the log-likelihood and filtered estimates of state variables

The particle filter of Appendix B.2 assumes we observe the yield factors $g_{1:T}$. To calculate the likelihood of the model, we must integrate over both $g_{1:T}$ and $h_{0:T}$ simultaneously. The particle filter we implement for this purpose is the mixture Kalman filter of Chen and Liu(2000).

[10] Note that in a standard vector autoregression with stochastic volatility that does not require rotating the yield state vector as $g_t = (r_t \; er_t \; tp_t)'$, the matrix of autoregressive parameters $\Phi$ can be drawn using Gibbs sampling as in a standard Bayesian VAR; see, e.g. Del Negro and Schorfheide(2011).

### Appendix B.3.1 State space form

In order to implement the particle filter, it is easier to write the model in terms of the marginal dynamics of $h_t$ and the dynamics of the conditionally Gaussian state variables $x_t = (m_t' \;\; g_t')'$. In this sub-section, we write $Y_t = (m_t', y_t')'$. Using properties of the multivariate normal distribution, the marginal distribution $p(h_{t+1}|h_t;\theta)$ can be written as

$$h_{t+1} = \mu_h + \Phi_h h_t + \varepsilon^*_{h,t+1} \qquad \varepsilon^*_{h,t+1} \sim \mathrm{N}(0, S_h) \tag{B.6}$$

$$S_h = \Sigma_h \Sigma_h' + \Sigma_{hm}\Sigma_{hm}' + \Sigma_{hg}\Sigma_{hg}' \tag{B.7}$$

The conditional distribution of $p(x_{t+1}|h_{t+1}, x_t, h_t, \theta)$ is

$$\begin{aligned} x_{t+1} &= \begin{pmatrix} \mu_m \\ \mu_g \end{pmatrix} + \begin{pmatrix} \Phi_m & \Phi_{mg} \\ \Phi_{gm} & \Phi_g \end{pmatrix} x_t + \begin{pmatrix} \Phi_{mh} \\ \Phi_{gh} \end{pmatrix} h_t + \begin{pmatrix} \Sigma_m D_{m,t}\Sigma_{hm}' & \Sigma_g D_{g,t}\Sigma_{hg}' \end{pmatrix} S_h^{-1}\varepsilon^*_{h,t+1} \\ &\quad + \varepsilon^*_{x,t+1} \qquad \varepsilon^*_{x,t+1} \sim \mathrm{N}(0, S_{x,t}) \\ S_{x,t} &= \begin{pmatrix} \Sigma_m D^2_{m,t}\Sigma_m' & \Sigma_{gm} D^2_{m,t}\Sigma_m' \\ \Sigma_m D^2_{m,t}\Sigma_{gm}' & \Sigma_{gm} D^2_{m,t}\Sigma_{gm}' + \Sigma_g D^2_{g,t}\Sigma_g' \end{pmatrix} \\ &\quad - \begin{pmatrix} \Sigma_m D_{m,t}\Sigma_{hm}' & \Sigma_g D_{g,t}\Sigma_{hg}' \end{pmatrix} S_h^{-1} \begin{pmatrix} \Sigma_{hg} D_{g,t}\Sigma_g' & \Sigma_{hm} D_{m,t}\Sigma_m' \end{pmatrix} \end{aligned}$$

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

$$Y_t = \begin{pmatrix} m_t \\ y_t \end{pmatrix}, \qquad Z_t = \begin{pmatrix} I & 0 \\ 0 & B \end{pmatrix}, \qquad d_t = \begin{pmatrix} 0 \\ A \end{pmatrix}, \qquad H_t = \begin{pmatrix} 0 & 0 \\ 0 & \Omega \end{pmatrix},$$

$$x_t = \begin{pmatrix} m_t \\ g_t \end{pmatrix}, \qquad T_t = \begin{pmatrix} \Phi_m & \Phi_{mg} \\ \Phi_{gm} & \Phi_g \end{pmatrix}, \qquad C_t = S_{x,t},$$

$$c_t = \begin{pmatrix} \mu_m + \Phi_{mh} h_t \\ \mu_g + \Phi_{gh} h_t \end{pmatrix} + \begin{pmatrix} \Sigma_m D_{m,t}\Sigma_{hm}' & \Sigma_g D_{g,t}\Sigma_{hg}' \end{pmatrix} S_h^{-1}\varepsilon^*_{h,t+1},$$

and where $R_t = I$. Note that $c_t$ is a function of $h_{t+1}$.

### Appendix B.3.2 Mixture Kalman filter

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.

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 $h_0^{(i)} \sim p\left(h_0; \theta\right)$.
- Set $x_{0|0}^{(i)} = \begin{pmatrix} \bar{\mu}_m \\ \bar{\mu}_g \end{pmatrix}$, $P_{0|0}^{(i)} = \begin{pmatrix} \Sigma_m \left(D_{m,0}^2\right)^{(i)} \Sigma_m' & \Sigma_m \left(D_{m,0} D_{g,0}\right)^{(i)} \Sigma_{gm}' \\ \Sigma_{gm} \left(D_{m,0} D_{g,0}\right)^{(i)} \Sigma_m' & \Sigma_{gm} \left(D_{m,0}^2\right)^{(i)} \Sigma_{gm}' + \Sigma_g \left(D_{g,0}^2\right)^{(i)} \Sigma_g' \end{pmatrix}$,
- Set $\ell_0 = 0$.

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

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

- Draw from the transition density: $h_t^{(i)} \sim p(h_t|h_{t-1}^{(i)}; \theta)$ given by B.6.
- Calculate $c_{t-1}^{(i)}$ and $C_{t-1}^{(i)}$ using $\left(h_{t-1}^{(i)}, h_t^{(i)}\right)$.
- Run the Kalman filter:

$$\begin{aligned}
x_{t|t-1}^{(i)} &= T x_{t-1|t-1}^{(i)} + c_{t-1}^{(i)} \\
P_{t|t-1}^{(i)} &= T P_{t-1|t-1}^{(i)} T' + R C_{t-1}^{(i)} R' \\
v_t^{(i)} &= Y_t - Z x_{t|t-1}^{(i)} - d \\
F_t^{(i)} &= Z P_{t|t-1}^{(i)} Z' + H \\
K_t^{(i)} &= P_{t|t-1}^{(i)} Z' \left(F_t^{(i)}\right)^{-1} \\
x_{t|t}^{(i)} &= x_{t|t-1}^{(i)} + K_t^{(i)} v_t^{(i)} \\
P_{t|t}^{(i)} &= P_{t|t-1}^{(i)} - K_t^{(i)} Z_t P_{t|t-1}^{(i)}
\end{aligned}$$

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

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 $\left\{x_{t|t}^{(i)}, P_{t|t}^{(i)}, h_t^{(i)}\right\}_{i=1}^J$ 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.4 Prior distributions

We use proper priors for all parameters of the model. Throughout this discussion, a normal distribution is defined as $x \sim \mathrm{N}(\mu_x, V_x)$, where $\mu_x$ 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 $\iota_k$ denote a $k \times 1$ vector of ones.

- The matrix $\Phi_g^{\mathbb{Q}}$ has a total of 6 free parameters, see Appendix A.2
  - We set the 3 ordered eigenvalues as $\lambda_{g,1}^{\mathbb{Q}} \sim \mathrm{N}(0.99, 0.0001)$, $\lambda_{g,2}^{\mathbb{Q}} \sim \mathrm{N}(0.95, 0.0010)$, $\lambda_{g,3}^{\mathbb{Q}} \sim \mathrm{N}(0.7, 0.0025)$. We reject any draws that re-order these eigenvalues.
  - The top row of $\Phi_g^{\mathbb{Q}}$ is: $\phi_{g,11}^{\mathbb{Q}} \sim \mathrm{N}(0.9, 0.02)$, $\phi_{g,12}^{\mathbb{Q}} \sim \mathrm{N}(0.1, 0.02)$, $\left|\phi_{g,13}^{\mathbb{Q}}\right| \sim \mathrm{Gamma}(2, 0.1)$
- $\bar{\mu}_m \sim \mathrm{N}\left((3.75, 6)'/1200, I_M \times (0.8485/1200)^2\right)$.
- $\bar{\mu}_g^u \sim \mathrm{N}\left((4.2, 1.3)'/1200, I_G \times (0.8485/1200)^2\right)$.
- The matrix $\Phi_f$ is decomposed into its eigenvalues and eigenvectors, see Appendix A.2.
  - The matrix has 9 eigenvalues. The first $M+G=5$ must be ordered and the second $H=4$ must be ordered. We use conditional beta distributions which guarantee an ordering. For the macro and yield factors: $\lambda_{f,1} \sim \mathrm{beta}(600, 5)$, $\lambda_{f,2}/\lambda_{f,1} \sim \mathrm{beta}(600, 5)$, $\lambda_{f,3}/\lambda_{f,2} \sim \mathrm{beta}(300, 5)$, $\lambda_{f,4}/\lambda_{f,3} \sim \mathrm{beta}(50, 4)$, $\lambda_{f,5}/\lambda_{f,4} \sim \mathrm{beta}(50, 4)$. And, for the volatility factors: $\lambda_{f,6} \sim \mathrm{beta}(500, 3.5)$, $\lambda_{f,7}/\lambda_{f,6} \sim \mathrm{beta}(500, 3.5)$, $\lambda_{f,8}/\lambda_{f,7} \sim \mathrm{beta}(500, 3.5)$, $\lambda_{f,9}/\lambda_{f,8} \sim \mathrm{beta}(500, 3.5)$
  - The free eigenvectors $q_{ij}$ are $q_{ij} \sim \mathrm{N}(0, 6000)$.
- We place a prior on the covariance matrix $\Sigma_f^* \Sigma_f^{*\prime} \sim \mathrm{InvWish}(\nu, S^*)$ where

$$\Sigma_f^* = \begin{pmatrix} \Sigma_m \bar{D}_m & 0 & 0 \\ \Sigma_{mg} \bar{D}_m & \Sigma_g \bar{D}_g & 0 \\ \Sigma_{hm} & \Sigma_{hg} & \Sigma_h \end{pmatrix} \qquad \begin{pmatrix} \mathrm{diag}(\bar{D}_m) \\ \mathrm{diag}(\bar{D}_g) \end{pmatrix} = \exp\left(\frac{\Gamma_0 + \Gamma_1 \bar{\mu}_h}{2}\right)$$

This is a prior over the free parameters in $\Sigma_m, \Sigma_g, \Sigma_h, \Sigma_{mg}, \Sigma_{hm}, \Sigma_{hg}, \bar{\mu}_h, \Gamma_0$ and is conditional on the matrix $\Gamma_1$. We set $\nu = G + H + M + 5$ and, for our macro model, we set $S^*$ equal to

$$S^* = \begin{pmatrix} S_m^* & 0 & 0 \\ 0 & S_g^* & 0 \\ 0 & 0 & S_h^* \end{pmatrix} \quad S_m^* = I_M * 0.4e^{-6} \quad S_g^* = I_G * 0.1e^{-5} \quad S_h^* = I_H * 0.4e^{-7}$$

- For the yield-only models with $H = 1, 2$ factors, we estimate parameters in the matrix $\Gamma_1$. 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 $g_t$.
- Diagonal elements of $\Omega$ are $\sigma^2_{\omega,i} \sim \mathrm{InvGamma}(a, b)$ where $a = N_y$ and $b = 5.0000e-08$ where $N_y = 7$ yields.

# Appendix C Impulse response functions

We summarize (1) - (4) by

$$f_t = \mu_f + \Phi_f f_{t-1} + \Sigma_{t-1}\varepsilon_t.$$

where

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

For each draw $\left\{\theta^{(k)}, f_{1:T}^{(k)}\right\}$ in the MCMC algorithm where $f_t^{(k)} = \left(m_t', g_t^{(k)\prime}, h_t^{(k)\prime}\right)'$, 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 C.1 Time-varying (state dependent) impulse responses

Let $\tilde{f}_t^{(k)} = \left(\tilde{m}_t^{(k)\prime}, \tilde{g}_t^{(k)\prime}, \tilde{h}_t^{(k)\prime}\right)'$ 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 $\tilde{f}_{s-1}^{(k)} = f_{s-1}^{(k)}$, for a $\tau$ period impulse response, we iterate forward on the dynamics of the vector autoregression

$$\tilde{f}_t^{(k)} \quad = \quad \mu_f^{(k)} + \Phi_f^{(k)} \tilde{f}_{t-1}^{(k)} + \Sigma_{t-1}^{(k)} \tilde{\varepsilon}_t^{(k)}, \qquad t = s, \ldots, s+\tau.$$

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

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

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