Term Structure of Interest Rate Volatility and Macroeconomic Uncertainty *

Drew D. Creal†
Chicago Booth

Jing Cynthia Wu‡
Chicago Booth

First draft: November 1, 2013
This draft: January 25, 2014

Abstract

We propose a new model of the yield curve to capture both the dynamics of their conditional mean and the term structure of interest rate volatilities. The new class of affine term structure models exhibits multiple unpriced stochastic volatility factors without imposing constraints on the conditional mean of yields. The common movement in the volatilities extracted from the model provides a new measure of economy-wide uncertainty, and we use it to study the impact uncertainty has on the macroeconomy. Towards the end of the Great Recession, uncertainty accelerated the zero lower bound for the short term interest rate, added to concerns over deflation, and contributed to higher unemployment rates.

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

---

*We thank Steve Davis, Jim Hamilton, Lars Hansen, Dale Rosenthal, Dora Xia, Lan Zhang and seminar and conference participants at Ohio State and UIC. 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.

†The University of Chicago Booth School of Business, 5807 South Woodlawn Avenue, Chicago, IL 60637, USA, dcreal@chicagobooth.edu

‡The University of Chicago Booth School of Business, 5807 South Woodlawn Avenue, Chicago, IL 60637, USA, Cynthia.Wu@chicagobooth.edu

1 Introduction

We propose a new model of the yield curve to capture both the dynamics of their conditional mean and the term structure of interest rate volatilities. The new class of term structure models exhibits multiple, unpriced stochastic volatility factors without imposing constraints on the conditional mean. We extract the common movement in the volatilities from the model and it provides a new measure of economy-wide uncertainty that captures policy uncertainty, inflation uncertainty, uncertainty over real activity, and financial market uncertainty. We use this new measure to assess the impact uncertainty has on the macroeconomy. We find uncertainty contractionary, and that it deepens recessions. For example, towards the end of the Great Recession, uncertainty accelerated the zero lower bound for the short term interest rate, added to concerns over deflation, and contributed to higher unemployment rates.

In the new discrete-time affine term structure model (ATSM), yields are priced by Gaussian factors. Consequently, bond prices are the same as standard Gaussian ATSMs, the benchmark model in the macro-finance literature on the term structure. Under the physical measure, the Gaussian factors follow a vector autoregression with stochastic volatility factors driving the covariance matrix. The stochastic volatility factors are not priced in the cross-section, providing these factors the freedom to fit the term structure of yield volatilities without compromising the fit of the conditional mean. Our model leverages the fact that in discrete-time the factors pricing bonds do not have to share the same covariance matrix under the physical and risk-neutral measures in order to satisfy no-arbitrage. Comparing several models with different numbers of factors, we find evidence that there are three economically meaningful factors in the term structure of interest rate volatilities.

We calculate the common movement in the volatilities, and it provides a new economy-wide measure of uncertainty. Our uncertainty measure displays strikingly high correlations with various measures of macroeconomic uncertainty that are popular in the literature, including the policy uncertainty index of Baker, Bloom, and Davis(2013), the CBOE Volatility Index (VIX), and forecast dispersions from the Survey of Professional Forecasters (SPF) for the Treasury Bill (T-bill), consumer price index (CPI) inflation, GDP deflator inflation, real GDP growth, and industrial production (IP) growth. These measures capture different aspects of macroeconomic uncertainty: policy uncertainty, monetary policy uncertainty, inflation uncertainty, uncertainty about real activities, and financial market uncertainty.

We use our uncertainty measure to assess the impact uncertainty has on the macroeconomy, and find it contractionary. Uncertainty increased inflation further by over 2% when it was historically high in the early 1980s. It accelerated the zero lower bound for the short term interest rate, and added to concerns over deflation after the Great Recession. It also contributed to higher unemployment when the economy was in deep recessions, in the 1980s and after the Great Recession for examples.

Literature This paper is related to two different parts of the macro-finance literature. First, we contribute to the term structure literature by proposing new affine term structure models that have the flexibility to capture both the conditional mean and variance. Second, we provide evidence on the importance of uncertainty shocks toward macroeconomic fluctuations.

Affine term structure models are instrumental in macroeconomics and finance for studying the interaction between monetary policy, financial markets, and the macroeconomy; for overviews, see Piazzesi(2010), Gürkaynak and Wright(2012), and Diebold and Rudebusch(2013). Much emphasis has been placed on fitting the cross section of yields. For example, in the class of non-Gaussian models studied by Dai and Singleton(2000) and Duffee(2002) among others, Creal and Wu(2013) showed that the latent factors serve as both yield and volatility factors but they are primarily chosen to fit the cross section of yields at the expense of fitting volatilities poorly. To address this limitation, Collin-Dufresne and Goldstein(2002) proposed the class of unspanned stochastic volatility (USV) which create greater flexibility in modeling yield volatility.[1] USV models improve the fit of volatility at the

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

cost of restricting the cross-sectional fit of yields. More importantly, adding multiple volatility factors becomes challenging. The existing literature stops at one factor, even though the data suggests that multiple factors are needed. This paper introduces a flexible model to fit the term structure of interest rate volatilities and yields at the same time. Furthermore, we show how to reliably estimate the models using Markov chain Monte Carlo (MCMC) and particle filtering algorithms.

Our paper also contributes to the recent and rapidly growing literature on uncertainty and its relationship with macroeconomic fluctuations; see, e.g. Bloom(2013) for a survey; Bekaert, Hoerova, and Lo Duca(2013), and Aastveit, Natvik, and Sola(2013) for empirical evidence; and Ulrich(2012), Pastor and Veronesi(2012) and Pastor and Veronesi(2013) for theoretical models. Baker, Bloom, and Davis(2013) recently proposed a policy uncertainty measure, and Jurado, Ludvigson, and Ng(2013) proposed a measure based on a large panel of macroeconomic variables. Our measure complements other measures, and is unique for the following reasons. First, the treasury market bridges together financial markets, monetary policy and the macroeconomy. Volatility from the bond market is therefore a good indicator of overall economic uncertainty. Second, data on yields is longer and more frequent than many available uncertainty measures. Our data has a monthly frequency and spans from 1952 to the present, which coincides with the availability of most macroeconomic time series. Finally, unlike many other measures, our data uses prices from a liquid market, which efficiently incorporate all the available information about the future economy.

The remainder of the paper is organized as follows. Section 2 presents some stylized facts first, and then accordingly proposes a new no arbitrage term structure model with multiple volatility factors. Section 3 uses MCMC and particle filtering algorithms to estimate and compare a collection of models with different numbers of volatility factors. In Section 4, we construct a new measure of economic uncertainty, compare it with measures from the literature, and study the economic implications of uncertainty. Section 5 concludes.

2 An ATSM with multiple unpriced volatility factors

2.1 Stylized facts

This section documents empirical facts about the term structure of interest rates volatilities. We focus on the following two questions: (i) can yield factors explain movements in yield volatilities? (ii) Is one factor enough to explain the cross-sectional variation of yield volatilities, as the literature suggests? Answers to these questions provide a guideline for the new model we propose in Section 2.2.

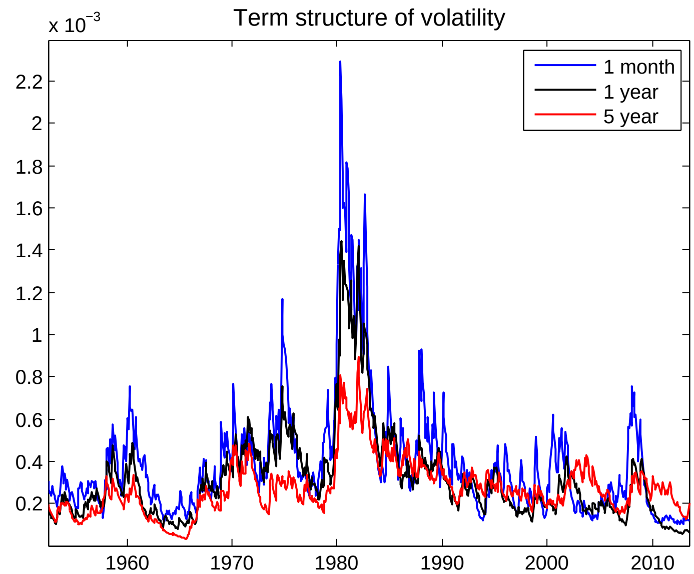
Data We use the Fama-Bliss zero-coupon yields available from the Center for Research in Securities Prices (CRSP) for maturities $n = (1, 3, 12, 24, 36, 48, 60)$ months. The data are available monthly from June 1952 to June 2013.

Factor Structure Figure 1 plots the conditional volatilities of yields for three different maturities spanning the short, middle, and long ends of the term structure. The conditional volatilities for yields of different maturities are estimated from univariate generalized autoregressive score models of Creal, Koopman, and Lucas(2013).[2] The figure illustrates the following features: (i) yields display a considerable amount of conditional heteroskedasticity; (ii) their volatilities co-move; (iii) the volatility of the five year bond behaves very differently from the volatility of shorter maturities. For example, the five year yield was much less volatile than the one month yield during the 1970s and 1980s, but it became more volatile after 1990, especially since 2009 when the short term interest rate got trapped at its zero lower bound. The evidence points to a factor structure for the cross section of yield volatilities and it suggests that more than one factor is necessary to model them.

A principal component analysis of yield volatilities provides further evidence of a factor structure. Table 1 contains the R^2s from yield volatilities regressed on their principal

[2]For each maturity, we estimate a univariate Student's t GAS(1,1) model with the best-fitting AR(p) model for the conditional mean. The Student's t GAS model is similar to a GARCH model but is more robust to outliers, e.g. Creal, Koopman, and Lucas(2011) and Creal, Koopman, and Lucas(2013).

Figure 1: Estimated conditional volatility for the 1 month, 1 year, and 5 year yields.



Estimated conditional volatility for the 1 month, 1 year, and 5 year yields. The estimates of conditional volatility are from univariate generalized autoregressive score models.

components. Although the first principal component explains a significant amount of the variation in the cross section, especially for middle maturities, the second and third principal components still capture meaningful variation at the shorter and longer ends of the term structure. The three factors together explain 96% to 100% of the cross sectional variation.

As a comparison, we also report in Table 2 the R^2s from a regression of yields on their first three principal components. This table replicates the standard finding from the term structure literature that three factors explain almost all the cross sectional variation in yields. A comparison of Tables 1 and 2 makes it clear that more than one volatility factor is needed for a term structure model to fit the cross section of volatilities. Arguably, three factors may not be enough to summarize the information in the yield volatilities as well as the first three principal components explain the cross section of yields.

Unpriced volatility factors If we replace the independent variables in the regression in Table 1 with the first three principal components of yields, the resulting R^2s are around 0.3.

Table 1: R^2s of yield volatilities on their principal components

	1m	3m	1y	2y	3y	4y	5y
1 PC	0.890	0.927	0.972	0.930	0.867	0.794	0.756
2 PCs	0.962	0.954	0.973	0.984	0.989	0.981	0.951
3 PCs	1.000	0.995	0.981	0.984	0.991	0.986	0.964

Table 2: R^2s from a regression of yields on their principal components

	1m	3m	1y	2y	3y	4y	5y
1 PC	0.957	0.972	0.992	0.997	0.990	0.980	0.969
2 PCs	0.995	0.998	0.995	0.998	0.999	0.999	0.998
3 PCs	0.999	0.998	0.999	0.999	0.999	1.000	0.999

This confirms a recently well documented idea that yield volatilities are not priced in the bond market. The bond market is not complete because agents cannot hedge all their risks by trading only in bonds. See Collin-Dufresne and Goldstein(2002), Heidari and Wu(2003), Li and Zhao(2006), Trolle and Schwartz(2009), Andersen and Benzoni(2010) and Creal and Wu(2013) for evidence from reduced form and structural models.

In summary, the stylized facts documented in the previous section point to two important features in the data for building a term structure model. They are that: (i) the volatility factors are not priced; and (ii) there are multiple volatility factors. In the next section, we introduce a new class of discrete-time ATSMs to capture both features.

2.2 A new affine term structure model

2.2.1 Risk-neutral dynamics and bond prices

The state vector $x_t = (g_t', h_t')'$ consists of two types of factors. There are G Gaussian factors g_t that determine the conditional mean and H non-Gaussian factors h_t that determine the volatility of yields. To separate the conditional mean from the conditional volatilities as observed in the data, we only allow the Gaussian factors to be priced. In an affine model, this means that the bond loadings on the non-Gaussian factors h_t must be zero for all maturities.

Our approach to building affine models with unpriced volatility factors is new to the literature. It requires that the short term interest rate only depends on the Gaussian factors

$$r_t = \delta_0 + \delta_1' g_t, \tag{1}$$

and, under the risk-neutral measure $\mathbb{Q}$, the distribution of g_{t+1} does not depend on the volatility factors; i.e. $p(g_{t+1}|h_{t+1}, g_t, h_t; \theta, \mathbb{Q}) = p(g_{t+1}|g_t; \theta, \mathbb{Q})$. We specify g_t as a Gaussian vector autoregression

$$g_{t+1} = \mu_g^{\mathbb{Q}} + \Phi_g^{\mathbb{Q}} g_t + \varepsilon_{g,t+1}^{\mathbb{Q}}, \qquad \varepsilon_{g,t+1}^{\mathbb{Q}} \sim \mathrm{N}\left(0, \Sigma_g^{\mathbb{Q}}\right). \tag{2}$$

where $\Sigma_g^{\mathbb{Q}}$ is the covariance matrix under $\mathbb{Q}$.

As volatility factors are not priced in the bond market, yields do not contain information about their risk-neutral distribution $p(h_{t+1}|g_t, h_t; \theta, \mathbb{Q})$. Consequently, we do not need to specify this distribution explicitly, except that it has to satisfy the no-arbitrage restriction. i.e. events that occur with positive probability under $\mathbb{P}$ also occur with positive probability under $\mathbb{Q}$.

The $\mathbb{Q}$ measure is defined such that a risk neutral investor prices the n period bond at time t as the expected price of the same bond next period discounted by the short rate

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

As a result, zero-coupon bond prices are an exponentially affine function of the Gaussian state variables

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

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

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

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

$$y_t^n = a_n + b_n' g_t \tag{3}$$

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

2.2.2 Dynamics under the physical measure

Next, we specify the stochastic process for the state vector $x_t = (g_t', h_t')'$ under $\mathbb{P}$. The conditionally Gaussian state variables g_{t+1} follow a vector autoregression with conditional heteroskedasticity

$$g_{t+1} = \mu_g + \Phi_g g_t + \varepsilon_{g,t+1}, \qquad \varepsilon_{g,t+1} \sim \mathrm{N}\left(0, \Sigma_{g,t}\right). \tag{4}$$

where the covariance matrix $\Sigma_{g,t}$ is a function (to be specified below) of the volatility factors h_t. Unlike previous no-arbitrage term structure models with stochastic volatility, the covariance matrix of the shocks to the pricing factors in this model are not the same under $\mathbb{P}$ and $\mathbb{Q}$. In a discrete-time model, it is not necessary for these to be equivalent in order to satisfy no-arbitrage. No arbitrage is guaranteed as long as the probability measures are equivalent under $\mathbb{P}$ and $\mathbb{Q}$ such that we have a well-defined Radon-Nikodym derivative, see Section 2.2.3 below.

In our model, a researcher can parameterize the covariance matrix $\Sigma_{g,t}$ and the dynamics of the volatility factors h_{t+1} flexibly, as long as the covariance matrix is positive definite. For example, the dynamics of the volatility factors could be a linear combination of a multivariate, discrete-time Cox, Ingersoll, and Ross(1985) process as in Creal and Wu(2013). To strike a balance between flexibility and tractability, we model the volatility factors as a (conditionally) log-normal process

$$\begin{aligned}
\Sigma_{g,t} &= L_g D_t L_g' \qquad\qquad D_t = \operatorname{diag}\left(\exp\left(\Gamma h_t\right)\right) \\
h_{t+1} &= \mu_h + \Phi_h h_t + \varepsilon_{h,t+1}, \qquad \varepsilon_{h,t+1} \sim \mathrm{N}\left(0, \Sigma_h\right) \qquad (5)
\end{aligned}$$

where D_t is a diagonal matrix, and L_g is a lower triangular matrix. The $G \times H$ matrix Γ permits a factor structure within the covariance matrix and allows us to estimate models where $G \neq H$. This specification has several advantages. It is tractable to estimate by Markov chain Monte Carlo methods. And, it allows a substantial amount of flexibility in the covariance matrix.[3]

We also impose an additional restriction that the long-run mean of the covariance matrix is the same under $\mathbb{P}$ and $\mathbb{Q}$ measures; i.e. $\Sigma_g^{\mathbb{Q}} = L_g \operatorname{diag}\left(\exp\left[\Gamma\left(I_H - \Phi_h\right)^{-1} \mu_h\right)\right] L_g'$. This restriction prevents there from being additional free parameters within the bond loadings that are not present in homoskedastic Gaussian ATSMs; the benchmark models in the macro-finance literature. It also ensures that our model nests Gaussian ATSMs when the variance of the volatility shocks goes to zero $\Sigma_h = 0$ in (5).

2.2.3 Implied stochastic discount factor

The stochastic discount factor (SDF) can be written in general as

$$M_{t+1} = \frac{\exp\left(-r_t\right) p\left(g_{t+1} | h_{t+1}, g_t, h_t; \theta, \mathbb{Q}\right) p\left(h_{t+1} | g_t, h_t; \theta, \mathbb{Q}\right)}{p\left(g_{t+1} | h_{t+1}, g_t, h_t; \theta, \mathbb{P}\right) p\left(h_{t+1} | g_t, h_t; \theta, \mathbb{P}\right)}.$$

[3]The covariance matrix $\Sigma_{g,t}$ is a special case of the LDL decomposition used by Primiceri(2005).

Without the dynamics of h_{t+1} under $\mathbb{Q}$, the SDF is not fully specified. Nevertheless, we can write the SDF for the conditionally Gaussian portion of our model as

$$\begin{aligned} M_{g,t+1} &= \frac{\exp\left(-r_t\right) p\left(g_{t+1}|h_{t+1}, g_t, h_t; \theta, \mathbb{Q}\right)}{p\left(g_{t+1}|h_{t+1}, g_t, h_t; \theta, \mathbb{P}\right)} \\ &= \exp\left(-r_t - \lambda'_{g,t+1}\varepsilon_{g,t+1} - \frac{1}{2}\lambda'_{g,t+1}\lambda_{g,t+1} - \iota'\left[\Phi_h h_t + \varepsilon_{h,t+1}\right]\right) \end{aligned}$$

where the price of risk is $\lambda_{g,t+1} = \varepsilon^{\mathbb{Q}}_{g,t+1} - \varepsilon_{g,t+1}$. The first term uses the short rate r_t to discount over time. The second term is the risk premium, which measures the price of risk per quantity of risk. The third term is the usual Jensen's inequality. These terms are the same as a standard Gaussian ATSM with no stochastic volatility. The last term is new to this model. It comes from the ratio of determinants between the Gaussian densities under both measures.

2.2.4 Relation to models in the literature

Gaussian ATSMs Our risk neutral dynamics for the Gaussian state variables and the resulting expressions for bond prices are identical to those found in Gaussian ATSMs; see, e.g. Ang and Piazzesi(2003) and Hamilton and Wu(2012a). The Gaussian ATSM is a special case of our model when the variance of the shocks to volatility goes to zero, $\Sigma_h = 0$ in (5). The covariance matrices under $\mathbb{P}$ and $\mathbb{Q}$ are then equal $\Sigma_{g,t} = \Sigma^{\mathbb{Q}}_g$ in each time period.

Unspanned stochastic volatility (USV) models If we follow Creal and Wu(2013) and model the volatility factors as a discrete-time, multivariate Cox, Ingersoll, and Ross(1985) process and parameterize the covariance matrix $\Sigma_{g,t}$ as a linear function of h_t, our model nests the discrete-time version of the USV model of Collin-Dufresne and Goldstein(2002). There are two key differences between our model and USV models. First, USV models impose non-trivial restrictions on the Gaussian bond loadings in order to set the non-Gaussian loadings to zero. These restrictions limit the flexibility of the functions being fit through the cross-section of yields. Second, almost all existing USV models have only one unspanned volatility factor. From the evidence provided in Section 2.1, a model will require multiple factors to fit volatility. In contrast, the flexibility of our model makes it straightforward to introduce as many volatility factors as needed to describe the data.

Models with hidden Gaussian factors Our model is related to the Gaussian ATSMs with hidden factors introduced by Duffee(2011). In his model, some Gaussian factors are hidden from the cross-section of yields. Specifically, the short rate r_t does not load on the hidden factors and, under $\mathbb{Q}$, the Gaussian factors that price bonds evolve independently of the factors that are hidden. In our model, the volatility factors are also hidden in this sense.

3 Estimation

3.1 State space form

Stacking the yields y_t^n from (3) in order for N different maturities $n_1, n_2, ..., n_N$ gives $Y_t = A + Bg_t$, where $A = (a_{n_1}, \dots, a_{n_N})'$, $B = (b'_{n_1}, ..., b'_{n_N})'$. We assume that all yields are observed with Gaussian measurement errors which results in the following measurement equation

$$Y_t \quad = \quad A + Bg_t + \eta_t, \qquad \eta_t \sim \mathrm{N}(0, \Omega), \tag{6}$$

where we take Ω to be a diagonal matrix. The transition equations for the Gaussian and non-Gaussian state variables are (4) and (5).

3.2 Bayesian estimation

The ATSM with stochastic volatility in (4)-(6) is a non-linear, non-Gaussian state space model whose log-likelihood is not known in closed-form. Therefore, we estimate the model using Bayesian methods. We use Markov chain Monte Carlo (MCMC) to estimate the parameters of the model and a particle filter to calculate the log-likelihood and filtered estimates of the state variables. Particle filters are simulation based algorithms for handling non-linear, non-Gaussian state space models; see Creal(2012) for a survey.

Our MCMC algorithm leverages the fact that the term structure model can be written as a linear, Gaussian state space model, conditional on knowing the non-Gaussian state variables $h_{0:T-1} = (h_0, \ldots, h_{T-1})$. This is important both statistically and computationally as it allows us to use the Kalman filter for many steps of the MCMC algorithm. The MCMC algorithm, discussed in more detail below, requires drawing the static parameters of the model as well as two sets of state variables $g_{1:T} = (g_1, \ldots, g_T)$ and $h_{0:T-1} = (h_0, \ldots, h_{T-1})$. The Kalman filter makes these steps straightforward computationally and it allows us to: (i) draw the parameters of the model while minimizing the amount we condition on the state variables; (ii) draw the state variables in large blocks using simulation smoothing (forward filtering, backward sampling) algorithms. Both of these practices lead to improved performance of MCMC by inducing better mixing Markov chains; see, e.g. Liu, Wong, and Kong(1994). In practice, we have found that marginalizing out the Gaussian factors while drawing the static parameters of the model is particularly important for providing reliable estimates.

Full details of the MCMC algorithm are provided in Appendix A. We briefly describe some of the steps that highlight why it is efficient. Conditional on knowing the volatility factors $h_{0:T-1}$, we draw the parameters and Gaussian state variables $(\Phi_g^{\mathbb{Q}}, \delta_0, \mu_g, g_{1:T})$ jointly from their full conditional distribution. This is performed in two steps. First, we draw $\Phi_g^{\mathbb{Q}}$ while marginalizing out $(\delta_0, \mu_g, g_{1:T})$. Then, conditional on the draw of $\Phi_g^{\mathbb{Q}}$, we draw $(\delta_0, \mu_g, g_{1:T})$ using a simulation smoothing algorithm; see, de Jong and Shephard(1995) and Durbin and Koopman(2002). We then use the Kalman filter to marginalize these parameters $(\delta_0, \mu_g, g_{1:T})$ out of the Metropolis-Hastings acceptance ratio, when we draw the $\mathbb{P}$ parameters in the Gaussian factors and the covariance matrix of the measurement errors.[4] Finally, conditional on the draw of $g_{1:T}$, we draw the volatilities $h_{0:T-1}$ in one block using the MCMC algorithm

[4] A similar approach has been taken by Chib and Ergashev(2009) for Gaussian models with macroeconomic factors but no stochastic volatility.

of Kim, Shephard, and Chib(1998). Most of the parameters of the volatility process can also be drawn without conditioning on the volatilities.

The particle filter that we implement is called a mixture Kalman filter (MKF); see Chen and Liu(2000). It also utilizes the conditionally, linear Gaussian state space form to calculate statistically efficient estimates of the log-likelihood and state variables. The intuition behind the MKF is straightforward. If we knew the volatilities $h_{0:T-1}$, then the Kalman filter would calculate the filtered estimates of the Gaussian state variables exactly. In practice, we do not know the value of the volatilities. Therefore, the MKF calculates a weighted average of Kalman filter estimates 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 is always integrating out the Gaussian state variables exactly once we condition on any one path of the volatilities. The MKF has recently been applied in economics by Creal, Koopman, and Zivot(2010) and Shephard(2013).

3.3 Identifying restrictions

We classify the new models as $\mathbb{CW}(G, H)$ where G is the number of Gaussian factors and H is the number of factors that enter the covariance matrix $\Sigma_{g,t}$. To prevent the latent Gaussian factors from rotating we impose the following identifying restrictions; see Hamilton and Wu(2012b): (i) $\mu_g^{\mathbb{Q}} = 0$; (ii) $\Phi_g^{\mathbb{Q}}$ in ordered Jordan form; (iii) $\delta_1 = \iota$ is a column vector of ones.

For the non-Gaussian part, we impose the diagonal elements of L_g to equal $1e^{-4}$, which is done for convenience to keep the scale of the volatility factors close to one. For the $\mathbb{CW}(3, 1)$ model, Γ is a 3×1 vector with the second row restricted to one. For the $\mathbb{CW}(3, 2)$ model, Γ is a 3×2 matrix with the upper two rows imposed to the identity matrix.

3.4 Model comparison

The reduced form results in Table 1 indicate that there is a term structure for the volatility of interest rates. Our new model aims to fit the cross section of yields and their volatilities at the same time. With it, we are able to answer the following question: how many factors do we need to capture the term structure of yield volatilities? We will answer this question statistically and economically in this section.

We estimate models with $H = 0, 1, 2, 3$ volatility factors, and assess their performance. We focus on models with three Gaussian factors $G = 3$, because the consensus in the literature is that three Gaussian factors can capture almost all of the cross sectional variation of yields. It is also consistent with our reduced form evidence in Table 2.

Table 3 presents estimates from four models with different numbers of volatility factors. As a result the $\mathbb{Q}$ parameters are extremely similar across different models. For example $\Phi_g^{\mathbb{Q}}$ are almost identical across all four models.

We use the Bayesian information criterion for model selection.

$$BIC = -2 \times \log \text{likelihood} + T \log (\# \text{ of parameters})$$

where the BIC has been evaluated at the posterior mean. The BIC ranks the models from best to worst in the following order: $\mathbb{CW}(3,3)$, $\mathbb{CW}(3,2)$, $\mathbb{CW}(3,1)$, and $\mathbb{CW}(3,0)$. The BIC implied model probabilities for these models are 100%, 0, 0, and 0. According to BIC, the models $\mathbb{CW}(3,3)$ fits the data the best. From here on, we will base our analysis on the model preferred by the BIC: $\mathbb{CW}(3,3)$.

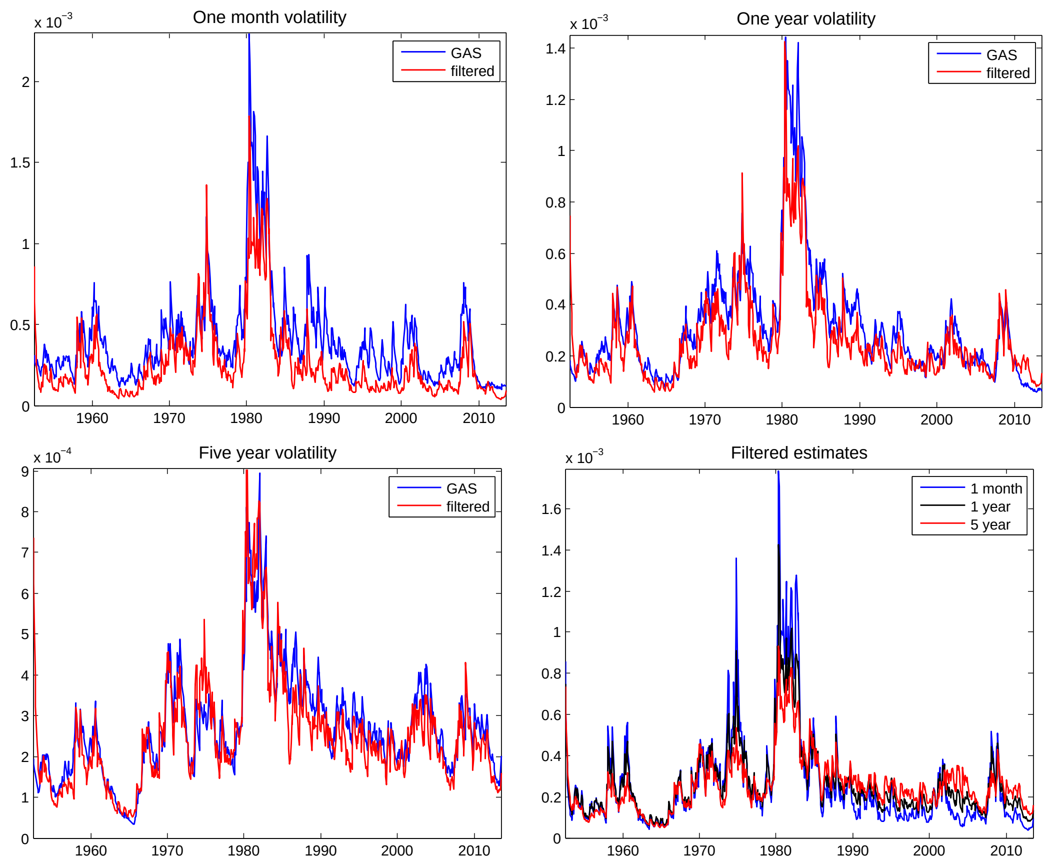
Figure 2 provides a comparison between the filtered estimates of the volatility from our preferred model and the reduced form GAS estimates of Creal, Koopman, and Lucas(2013) from Section 2. The correlations between the filtered estimates and GAS estimates for maturities of 1 month, 1 year and 5 years are 0.9019, 0.9318, and 0.9042. In the bottom right panel of Figure 2, we plot the term structure of volatilities from the $\mathbb{CW}(3,3)$ model

Table 3: Estimates for the $\mathbb{C}$W $(3, H)$ models for $H = 0, 1, 2, 3$.

$\mathbb{C}$W $(3,0)$		log-like =	37805.10	$\mathbb{C}$W $(3,1)$		log-like =	38193.95	$\mathbb{C}$W $(3,2)$		log-like =	38267.74	$\mathbb{C}$W $(3,3)$		log-like =	38297.73
δ_1			δ_0	δ_1			δ_0	δ_1			δ_0	δ_1			δ_0
1	1	1	0.010	1	1	1	0.010	1	1	1	0.008e-03	1	1	1	0.009e-04
—	—	—	(6.12e-04)	—	—	—	(6.66e-04)	—	—	—	(5.61e-04)	—	—	—	(6.60e-04)
$\mu_g^{\mathbb{Q}}$				$\mu_g^{\mathbb{Q}}$				$\mu_g^{\mathbb{Q}}$				$\mu_g^{\mathbb{Q}}$			
0	0	0		0	0	0		0	0	0		0	0	0	
—	—	—		—	—	—		—	—	—		—	—	—	
$\Phi_g^{\mathbb{Q}}$				$\Phi_g^{\mathbb{Q}}$				$\Phi_g^{\mathbb{Q}}$				$\Phi_g^{\mathbb{Q}}$			
0.996	0.950	0.754		0.996	0.951	0.740		0.995	0.955	0.734		0.996	0.951	0.761	
(0.0004)	(0.003)	(0.017)		(0.0004)	(0.002)	(0.013)		(0.001)	(0.003)	(0.014)		(0.001)	(0.003)	(0.015)	
$\tilde{\mu}_g \times 1200$				$\tilde{\mu}_g \times 1200$				$\tilde{\mu}_g \times 1200$				$\tilde{\mu}_g \times 1200$			
-11.20	0.856	-0.397		-9.358	1.530	-0.414		-8.705	1.588	-0.280		-9.093	1.267	-0.307	
(1.370)	(0.821)	(0.329)		(0.886)	(0.493)	(0.160)		(0.750)	(0.718)	(0.189)		(0.831)	(0.597)	(0.166)	
Φ_g				Φ_g				Φ_g				Φ_g			
1.009	0.047	0.092		0.962	-0.014	-0.064		0.950	-0.022	-0.078		0.953	-0.020	-0.072	
(0.005)	(0.011)	(0.023)		(0.007)	(0.011)	(0.022)		(0.008)	(0.012)	(0.024)		(0.007)	(0.012)	(0.023)	
0.008	0.935	-0.070		-0.010	0.954	-0.066		0.012	0.969	-0.047		-0.001	0.951	-0.066	
(0.012)	(0.016)	(0.034)		(0.007)	(0.012)	(0.026)		(0.012)	(0.017)	(0.038)		(0.012)	(0.019)	(0.037)	
0.007	0.013	0.873		0.040	0.035	1.012		0.033	0.033	1.027		0.038	0.039	1.034	
(0.007)	(0.010)	(0.027)		(0.007)	(0.012)	(0.024)		(0.007)	(0.010)	(0.019)		(0.007)	(0.011)	(0.022)	
$10^6\Sigma_g$				$10^4 L_g$				$10^4 L_g$				$10^4 L_g$			
0.114	-0.036	-0.048		1	0	0		1	0	0		1	0	0	
(0.104)	(0.131)	(0.074)		—	—	—		—	—	—		—	—	—	
-0.036	(0.247	-0.089		-0.392	1	0		-0.644	1	0		-0.602	1	0	
(0.131)	(0.233)	(0.202)		(0.074)	—	—		(0.067)	—	—		(0.076)	—	—	
-0.048	-0.089	0.163		-0.517	-0.659	1		-0.373	-0.625	1		-0.402	-0.715	1	
(0.074)	(0.202)	(0.196)		(0.075)	(0.047)	—		(0.065)	(0.044)	—		(0.072)	(0.044)	—	
				μ_h				μ_h				μ_h			
				0.123				0.048	0.038			0.050	0.257	-0.124	
				(0.046)				(0.035)	(0.055)			(0.093)	(0.115)	(0.117)	
				Φ_h		Γ		Φ_h		Γ		Φ_h			
				0.934		1.074		0.969	0.004	1	0	0.966	0.002	0.005	
				(0.022)		(0.047)		(0.019)	(0.012)	—	—	(0.020)	(0.057)	(0.045)	
						1		0.015	0.946	0	1	0.008	0.827	0.101	
						—		(0.030)	(0.023)	—	—	(0.034)	(0.064)	(0.050)	
						0.883				-0.120	0.819	0.007	0.079	0.891	
				Σ_h		(0.033)				(0.062)	(0.053)	(0.033)	(0.070)	(0.057)	
				0.240				Σ_h				Σ_h			
				(0.069)				0.109	0.107			0.112	0.087	0.098	
								(0.027)	(0.038)			(0.029)	(0.032)	(0.038)	
								0.107	0.364			0.087	0.257	0.218	
								(0.038)	(0.084)			(0.032)	(0.062)	(0.060)	
												0.098	0.218	0.352	
												(0.038)	(0.060)	(0.085)	

Posterior mean and standard deviations. We report the value of the log-likelihood function evaluated at the posterior mean.

Figure 2: Estimated volatility of 1m, 1yr, and 5 yr yields from the $\mathbb{CW}(3,3)$ model.



Comparison of estimated conditional volatilities for different maturities from the GAS(1,1) model versus filtered estimates from the $\mathbb{CW}(3,3)$ *model. Top left: 1 month yield; Top right: 1 year yield; Bottom left: 5 year yield; Bottom right: filtered estimates from 1 month, 1 year, 5 year plotted together.*

for three maturities covering short, middle, and long term yields. We can see the estimated conditional volatilities are, by and large, higher for shorter maturities prior to 1985 but the term structure of volatility changes its behavior after this period, with long term yields having higher volatility. This change in the term structure of volatility coincides with a change in the behavior of monetary policy. Prior to the early 1980's, monetary policy was conducted largely by money supply rules and after this period it has focused on influencing the short-term interest rate to induce changes in longer rates.

Statistical model comparison criteria suggest two or three volatility factors. Principal component analysis offers another perspective on how many volatility factors are needed.

Table 4: Correlation between volatilities and their principal components

	σ_{1t}	σ_{2t}	σ_{3t}
pc_{1t}	0.861	0.999	0.994
pc_{2t}	0.509	-0.036	-0.104
pc_{3t}	-0.012	0.035	-0.030

Correlations between the volatilities $(\sigma_{1t}, \sigma_{2t}, \sigma_{3t})$ *of the Gaussian yield factors* g_t *and their principal components* $(pc_{1t}, pc_{2t}, pc_{3t})$.

An advantage of principal components is that they are not subject to rotation of the latent factors. We use our preferred model $\mathbb{CW}(3,3)$. Define the 3×1 vector of factor volatilities σ_t as the standard deviations of Σ_{gt} in (4). We use the filtered estimates calculated from the particle filter because they use information only up to time t. From the estimated values of σ_t, we compute the three principal components labeled pc_t. Their correlation with the factor volatilities σ_t is reported in Table 3.4.

The first principal component captures the broad pattern across the volatilities of all three factors, as it is highly correlated with level, slope, and curvature volatilities. If anything, it captures more of slope and curvature volatilities than level volatility. The second and third principal component picks up what is left in the volatility of the level factor.

Table 3.4 reports the R^2s from regressions of model-implied yield volatilities on the principal components. One factor is able to explain the medium term yield volatility extremely well. The explanatory power declines for longer and shorter maturities. This is an indication that we need multiple factors to explain the term structure of interest volatilities. Adding a second principal component increases the R^2s significantly. With three factors, we are able to capture almost 100% of the variations in yield volatilities. To conclude, both statistically and economically, three factors could capture most of the variation in the term structure of volatility.

Table 5: R^2 of yield volatilities on PCs

	1m	3m	1y	2y	3y	4y	5y
1PC	0.964	0.973	0.999	0.989	0.959	0.921	0.887
2PCs	0.994	0.997	1.000	0.999	0.999	0.999	1.000
3PCs	1.000	1.000	1.000	1.000	0.999	0.999	1.000

R^2 from regressions of the principal components of factor volatilities ($pc_{1t}, pc_{2t}, pc_{3t}$) on the volatilities of yields implied by the model.

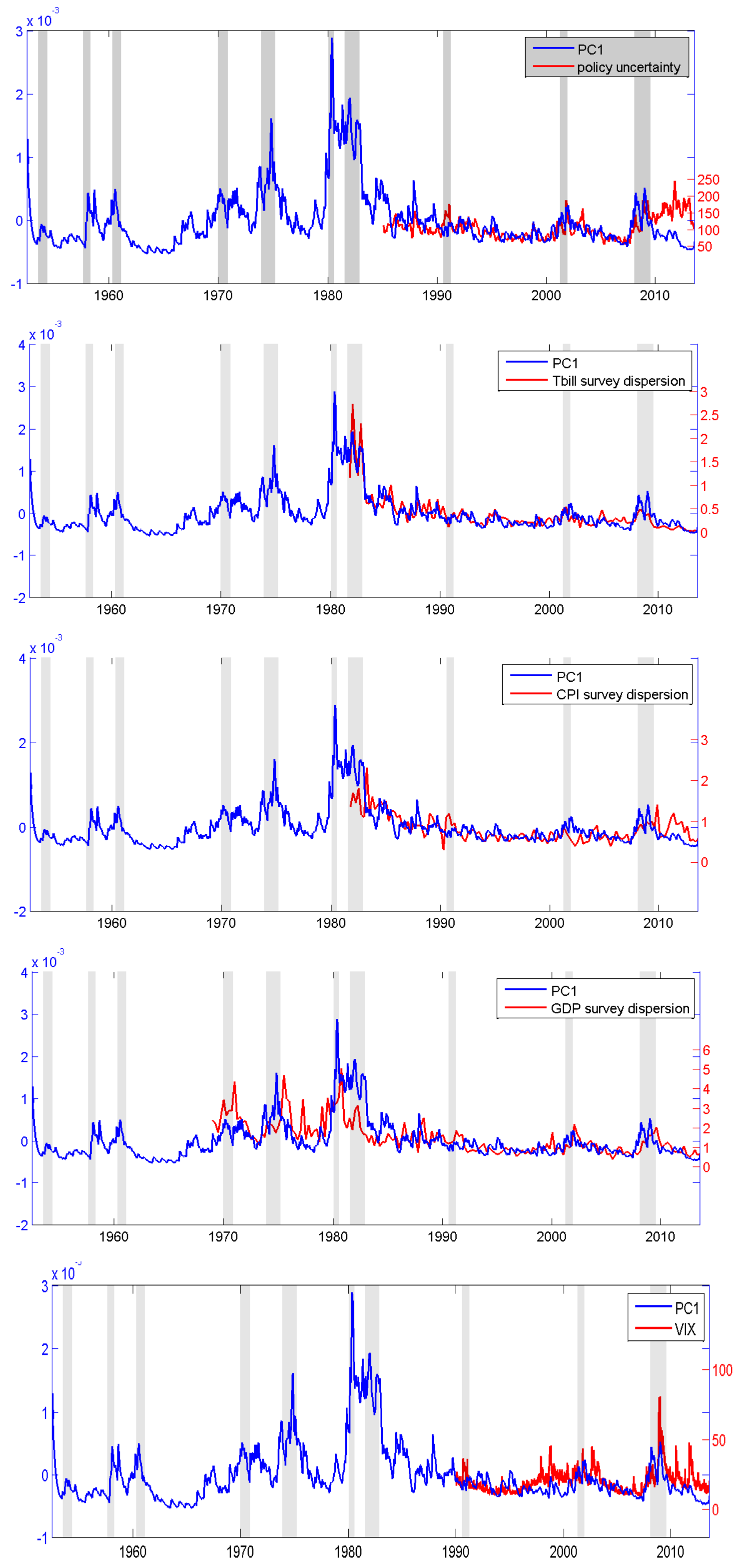
4 Economic uncertainty

What aspect of the economy do the dynamics of the yield curve volatilities capture? How do they impact the rest of the economy? This section uses these questions as a device to demonstrate the meaning, usefulness and importance of the information contained in yield volatilities. The Treasury market is the bridge between financial markets, monetary policy and the macroeconomy. Volatility from this market is therefore a good indicator of economic uncertainty. In this section, we propose a new uncertainty measure based on the first principal component of the factor volatilities as it captures the yield volatilities across all maturities. We first compare it with popular uncertainty measures, constructed from very different methodologies in the literature to see what aspect of economic uncertainty it captures. Then, we study its impact on the macroeconomy with a vector autoregression (VAR), similar to Stock and Watson(2001).

4.1 A new measure of uncertainty

We start by comparing our uncertainty measure with the policy uncertainty measure of Baker, Bloom, and Davis(2013), a prominent and popular news-based measure. We plot both measures in the top panel of Figure 3. Our measure (in blue) starts from 1952, whereas their measure (in red) goes back to 1985. Our uncertainty measure co-moved with their index of policy uncertainty from 1985 to 2008. The correlation was 0.53. Since 2009, the short term interest rate has entered into its zero lower bound regime. The yield curve has

Figure 3: Our uncertainty measure compared with various uncertainty measures.



Blue line (scale on the left): the first principal component of the factor volatilities. Red lines (scale on the left): in the top panel, policy uncertainty index of Baker, Bloom, and Davis(2013); in the next three panels, SPF dispersions for T bill, CPI inflation and GDP, where forecast horizons for T bill and GDP are 1 quarter ahead, and for CPI is 4 quarters ahead; in the bottom panel, VIX from CBOE. Shaded areas are NBER recessions dates.

Table 6: Correlations between our uncertainty measure and survey dispersions

	0Q	1Q	2Q	3Q	4Q
T-bill	0.84	0.90	0.87	0.82	0.85
T-bond	0.07	0.26	0.22	0.16	0.19
CPI inflation	0.36	0.55	0.53	0.55	0.64
GDP deflator inflation	0.50	0.48	0.56	0.54	-
IP growth	0.49	0.57	0.55	0.47	-
real GDP growth	0.54	0.60	0.57	0.57	-

Starting dates are for T-bill, CPI inflation: 1981Q3; T-bond: 1992Q1; IP growth, GDP growth, GDP deflator inflation: 1968Q4. Survey horizons: 0-4 quarters. For the longer sample series, earlier observations for 4 quarter ahead forecast are missing. Dispersion measure: 75th percentile minus 25th percentile of the forecasts. All growths are measured as quarter over quarter growth.

been lower overall, so have their volatilities, especially since the economy moved out of the Great Recession. On the other hand, the measure of policy uncertainty has increased. This is likely due to the fact that the policy uncertainty index is a news-based measure and this period has seen a marked increase in media coverage over the debt ceiling, passing of the government budget, and political brinkmanship over Obamacare.[5]

Besides policy uncertainty, we also compare our uncertainty measure with forecast dispersions from the Survey of Professional Forecasters (SPF) of the Federal Reserve Bank of Philadelphia. We use cross-sectional survey dispersions of forecasts for the 3-month Treasury Bill, 10-year Treasury Bond (T-bond), CPI inflation, GDP deflator inflation, IP and real GDP growth. SPF surveys are given at a quarterly frequency for forecasting horizons of 1,2,3, and 4 quarters. The sample starts from 1981 for T-bill and CPI inflation; from 1992 for T-bond, and from 1968 for IP growth, GDP growth and GDP deflator inflation. Each one of them has a shorter and less frequent sample than our measure.

We plot in the second panel of Figure 3 our uncertainty measure with the survey dispersion for forecasting the T-bill at a 1 quarter horizon. The latter serves as a measure of monetary policy uncertainty. The survey dispersion and our uncertainty measure co-move almost perfectly. The correlations between our measure and survey dispersions with different

[5]Bloom(2013) acknowledges this as well.

Table 7: Correlations between PCs and survey dispersions of T-bond

	0Q	1Q	2Q	3Q	4Q
pc_{1t}	0.09	0.26	0.17	0.13	0.20
pc_{2t}	0.34	0.30	0.16	0.26	0.36
pc_{3t}	0.10	0.30	0.25	0.31	0.35

Correlations between the first three principal components of volatility ($pc_{1t}, pc_{2t}, pc_{3t}$) and survey dispersions of the T-bond for horizons of (0,1,2,3,4) quarters.

forecasting horizons are calculated in Table 6. For the 1 quarter ahead forecast, their correlation is 90% and it is equally high across all forecast horizons. In contrast, from Table 7, T-bond survey dispersion is more correlated with the second and third principal components.

Next, we compare our uncertainty measure with survey dispersion over forecasting CPI inflation in the third panel of Figure 3. Again, the two measures move together, with correlations being as high as 64% in Table 6. A similar pattern prevails if we replace CPI inflation with GDP deflator inflation.

More interestingly, our uncertainty measure not only captures economic uncertainty in nominal terms, but also in real terms. The fourth panel of Figure 3 plots the survey dispersion of forecasting real GDP growth together with our uncertainty measure. They both peak during recessions and fall during expansions. The correlations are up to 60%, and the numbers stay high across all different forecasting horizons. The numbers are very similar if we replace real GDP growth with industrial production, see Table 6.

In the last panel of Figure 3, we also compare our uncertainty measure with VIX, a popular measure capturing uncertainty in equity markets. It measures the implied volatility (volatility under the risk-neutral measure) of S&P 500 index options. The VIX index is available from 1990, whereas our measure starts in 1952. Our measure and the VIX display some comovement with a correlation of 42.59%.

Another feature we observe from Figure 3 is that all recessions (shaded areas) but the one in 1991 coincide with spikes of our uncertainty measure. In fact all three principal components are associated with the business cycle. We confirm this relationship by regressing the three principal components of volatilities on a recession dummy. The p-values on the recession dummies are all zero.

We conclude the comparison between our uncertainty measure and other measures by regressing the first principal component of factor volatilities on a recession dummy, and survey dispersions for forecasting real GDP growth and GDP deflator inflation. We choose these measures because they have a longer sample (from 1968Q4 to 2013Q3). The results from the regression (with p-values in parenthesis) are

$$1000 \times \hat{pc}_1 = -\underset{(0.000)}{0.47} + \underset{(0.000)}{0.41} \times \mathbb{1}_{recession} + \underset{(0.006)}{0.16} \times dispersion_{inflation} + \underset{(0.000)}{0.23} \times dispersion_{growth},$$

All the coefficients are significant. Recessions are systematically associated with higher uncertainty. Also, our uncertainty measure captures both nominal and real macroeconomic uncertainty.

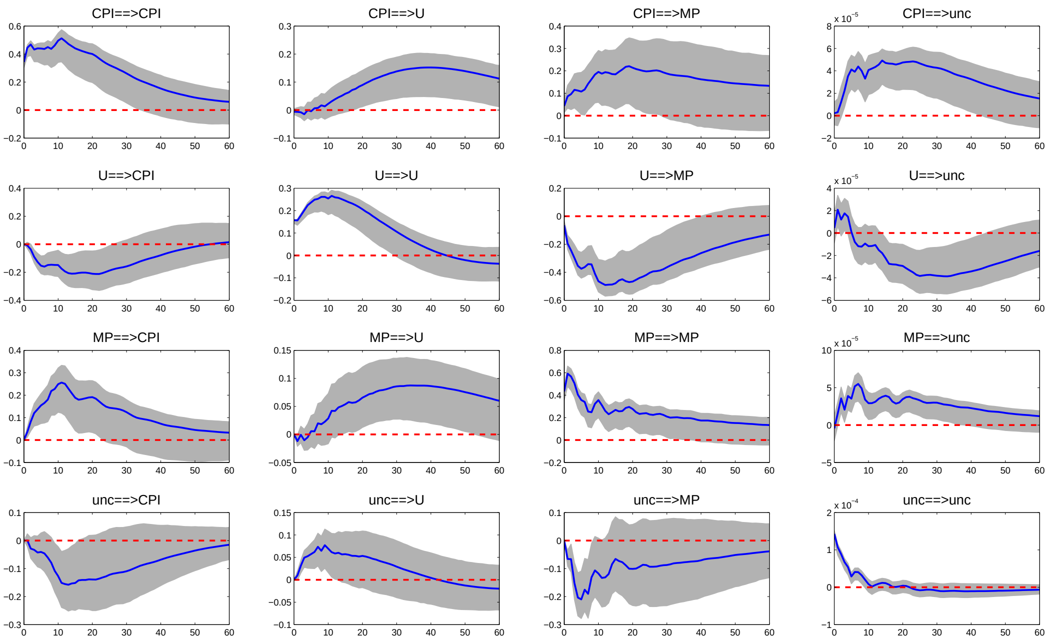
In summary, our uncertainty measure constructed from the Treasury market moves together with other popular uncertainty measures capturing different aspects of economic uncertainty: policy uncertainty, monetary policy uncertainty, inflation uncertainty, uncertainty about real activity, and financial market uncertainty. It complements the rapidly growing literature on uncertainty; see, e.g. Baker, Bloom, and Davis(2013), Jurado, Ludvigson, and Ng(2013). Notably, our measure is constructed directly from market prices which are forward looking and, as shown above, it does a good job at capturing economy-wide uncertainty. Next, we use this measure to study the role uncertainty plays in macroeconomic fluctuations.

4.2 Macroeconomic implications

In the previous section, we compared our uncertainty measure with other popular measures. Now, our goal is to understand how uncertainty impacts the rest of the economy. We address this question using a vector autoregression (VAR) as in Stock and Watson(2001).

We consider a four variable VAR with 12 lags with the variables ordered as: CPI inflation,

Figure 4: Impulse response functions



Impulse response functions from a four variable VAR(12) with CPI inflation, unemployment, the monetary policy rate, and uncertainty. Across the columns, we trace the impact each shock has on the observed data. Error bands are 90% confidence intervals.

unemployment rate, monetary policy rate, and our uncertainty measure.[6] All macro variables are measured monthly and are taken from the FRED database. CPI inflation is measured as year-over-year percentage change. The monetary policy measure we use is proposed by Wu and Xia(2013), who extended the effective federal funds rate with a latent measure of the monetary policy stance (the shadow rate) when the economy was at the zero lower bound. We use the standard recursiveness assumption for the VAR system.

Impulse responses Impulse responses to one standard deviation shocks are in Figure 4. The shaded regions are 90% confidence intervals constructed by the bootstrap. We find that uncertainty shocks are contractionary. A one standard deviation shock to our uncertainty

[6]Our results are robust to alternative orderings of the variables.

measure lowers the CPI inflation by 0.15% after a year, and increases the unemployment rate by about 0.07% at its maximum. The contractionary shock induces more accommodative monetary policy with the Fed lowering its policy rate by 0.2% in half a year. On the other hand, positive inflation and monetary policy shocks will induce higher uncertainty. An unemployment shock increases uncertainty first and then reduces uncertainty. This might be because higher unemployment follows by lower inflation and a lower policy rate, which in turn decrease uncertainty.

Historical decomposition We next conduct a historical decomposition asking the following question: what would have happened to inflation, unemployment, and the monetary policy rate if there were no uncertainty shocks. In the left panels of Figure 5, the blue lines are the observed data and the red dashed lines are the counterfactual estimates. The red dashed lines measure what the value of that variable would have been if there were no uncertainty shocks. On the right side of Figure 5, we examine the contribution that uncertainty shocks make to each series by plotting the difference between the blue and red lines in the left panel. From top to bottom are inflation, unemployment, and the monetary policy rate.

It is apparent that uncertainty shocks from the term structure of interest rates are an economically important contributor to all the variables we study. In the early 1980s, uncertainty added over 2% to inflation, pushing it to over 14%. Again in the 1990s, uncertainty shocks contributed over 1.5% to inflation around the 1990 recession and immediately afterwards. Recently, uncertainty shocks are negatively associated with inflation. For example, after the Great Recession, uncertainty shocks contributed almost -1% to the already very low level of inflation, bringing "deflation risk" to the table.

For the unemployment rate, the biggest contribution of the uncertainty shock came in the second recession in the 1980s, when it added about 1% to the unemployment rate. Following the Great Recession, uncertainty shocks have helped keep the unemployment rate stubbornly high.

As for the policy rate, the uncertainty shock contributed about 2.5% before the 1980 recession to the 19% federal funds rate then. Recently, during and after the Great Recession, uncertainty shocks added another -1% to the already historically low policy rate, accelerating the speed at which the federal funds rate hit its zero lower bound.

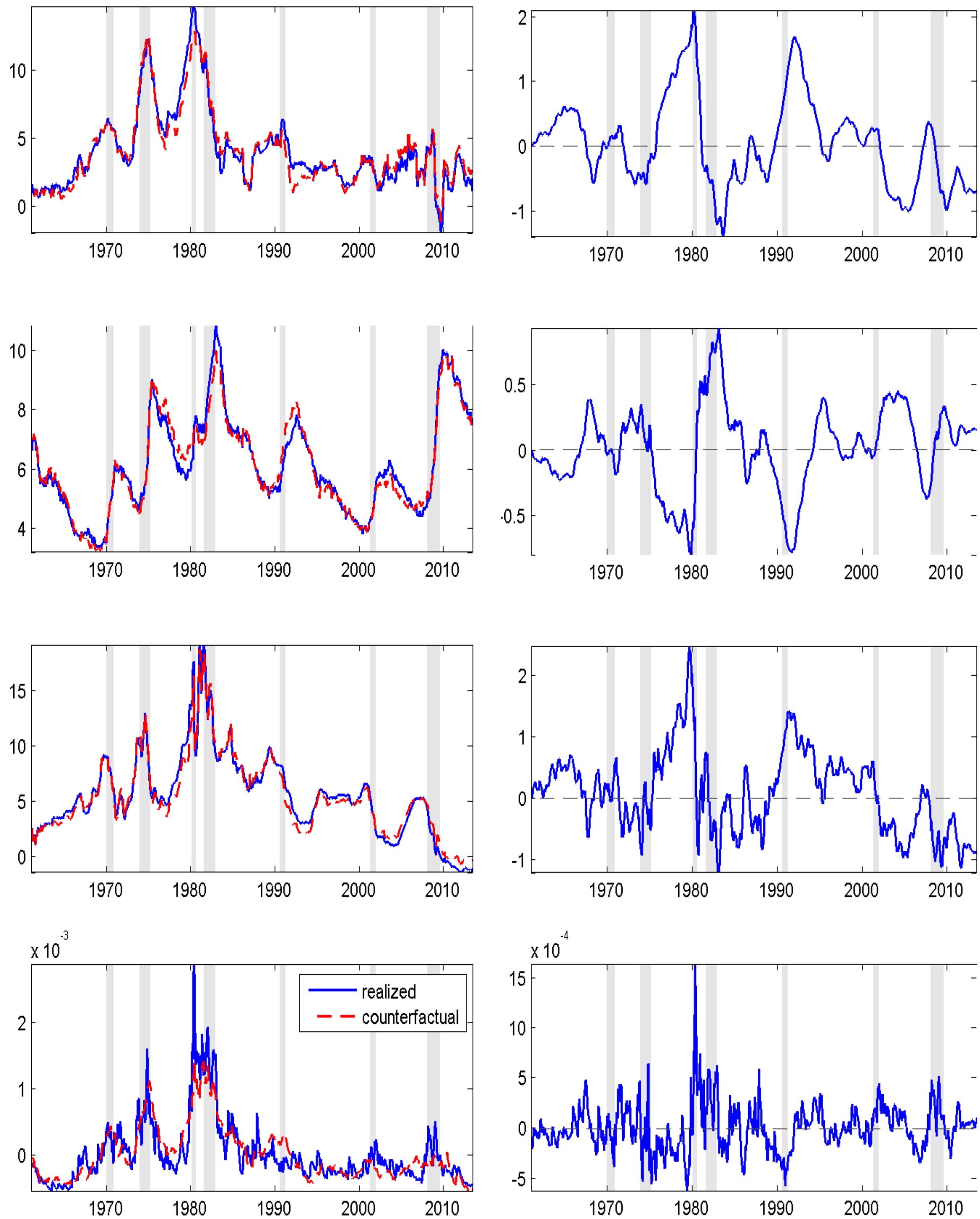
Granger-causality tests Next, we calculate Granger-causality tests, to evaluate if the parameters relating macro-variables and lagged uncertainty are statistically significant. First, we test the coefficient relating inflation to lagged uncertainty, and the p-value for the t-test is 1.1e-7. Uncertainty Granger-causes future inflation. Similarly, the p-value on the coefficients relating unemployment and monetary policy to lagged uncertainty are 1.1e-6 and 5.9e-7 respectively. Again, uncertainty Granger-causes unemployment and monetary policy. Finally, we test the system of three equations jointly by performing a likelihood ratio test with Sims' small-sample correction Hamilton(1994). Consistent with the previous results, we find a p-value of 5.7e-13, again rejecting the null hypothesis and finding uncertainty Granger-causes changes of future macro variables.

Summary Based on our analysis from the VAR, positive uncertainty shocks are contractionary. They increased inflation further when it was historically high in the early 1980s. It accelerated the zero lower bound, and added to concerns over deflation after the Great Recession. It pushed the unemployment rate higher when the economy was already in deep recessions.

5 Conclusion

We developed a new affine term structure model with stochastic volatilities. Our model captured the factor structure in the term structure of interest rate volatilities. At the same time, bond prices are the same as Gaussian ATSMs. We found that we needed three factors to describe the volatilities together with three factors for yields themselves. The common

Figure 5: Historical decomposition



Historical decomposition. Left column: Actual data in blue and counterfactual in red. In descending order, inflation, unemployment rate, monetary policy rate, and uncertainty shocks. Right column: difference between the data and counterfactual estimate.

movement in our volatilities displayed strikingly high correlations with measures for policy uncertainty, monetary policy uncertainty, inflation uncertainty, and GDP uncertainty. Thus it provided a new measure to capture the economy-wide uncertainty. We used this uncertainty measure to assess the economic impact of uncertainty. We found uncertainty contractionary, and that it deepened recessions. For example, towards the end of the Great Recession, uncertainty accelerated the zero lower bound for the short term interest rate, added to concerns over deflation, and contributed to higher unemployment.

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.

Ang, Andrew, and Monika Piazzesi (2003) “A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables” *Journal of Monetary Economics* 50, 745–787.

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

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

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

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

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

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

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

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

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.

Cox, John C., Jonathan E. Ingersoll, and Stephen A. Ross (1985) “A theory of the term structure of interest rates” *Econometrica* 53, 385–407.

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 (2013) “Estimation of non-Gaussian affine term structure models” Working paper, University of Chicago, Booth School of Business.

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

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

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

Diebold, Francis X., and Glenn D. Rudebusch (2013) *Yield Curve Modeling and Forecasting.* Princeton University Press, Princeton, NJ.

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

Duffee, Gregory R. (2011) "Information in (and not in) the term structure" *The Review of Financial Studies* 24, 2895–2934.

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

Durbin, James, and Siem Jan Koopman (2012) *Time Series Analysis by State Space Methods* Oxford University Press, Oxford, UK 2 edition.

Gürkaynak, Refet S., and Jonathan H. Wright (2012) "Macroeconomics and the term structure" *Journal of Economic Literature* 50, 331–367.

Hamilton, James D (1994) *Time Series Analysis* Princeton University Press, Princeton, NJ.

Hamilton, James D., and Jing Cynthia Wu (2012a) "Identification and estimation of Gaussian affine term structure models." *Journal of Econometrics* 168, 315–331.

Hamilton, James D., and Jing Cynthia Wu (2012b) "Testable implications of affine term structure models." *Journal of Econometrics* forthcoming.

Heidari, Massoud, and Liuren Wu (2003) "Are interest rate derivatives spanned by the term structure of interest rates?" *The Journal of Fixed Income* 13, 75–86.

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

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

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

Li, Haitao, and Feng Zhao (2006) "Unspanned stochastic volatility: evidence from hedging interest rate derivatives." *The Journal of Finance* 61, 341–378.

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

Liu, Jun S., Wing Hung Wong, and Augustine Kong (1994) "Covariance structure of the Gibbs sampler with applications to the comparisons of estimators and augmentation schemes." *Biometrika* 81, 27–40.

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

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

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

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

Piazzesi, Monika (2010) "Affine term structure models" in *Handbook of Financial Econometrics*, edited by Y. Ait-Sahalia and L. P. Hansen Elsevier, New York pages 691–766.

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

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

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

Trolle, Anders B., and Eduardo S. Schwartz (2009) "A general stochastic volatility model for the pricing of interest rate derivatives." *The Review of Financial Studies* 22, 2007–2057.

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

Wu, Jing Cynthia, and Dora Fan Xia (2013) "Measuring the macroeconomic impact of monetary policy at the zero lower bound." Working paper, University of Chicago, Booth School of Business.

Appendix A MCMC and particle filtering algorithms

Appendix A.1 MCMC algorithm

Appendix A.1.1 Conditionally Gaussian state space form

We write the model as a conditionally linear, Gaussian state space model, using the notation from Durbin and Koopman(2012). The parameter δ_0 enters the bond loadings A linearly and we can decompose this vector as $A = \tilde{A} + \iota\delta_0$, where ι is a vector of ones. Also, our MCMC algorithm works with the unconditional mean $\tilde{\mu}_g = (I_G - \Phi_g)^{-1}\mu_g$ of the factors. Define the state vector as $\alpha_t = \left(g_t', \delta_0, \tilde{\mu}_g'\right)'$ and write the model as

$$
\begin{aligned}
Y_t &= Z\alpha_t + d + \eta_t & \eta_t &\sim \mathrm{N}(0, H), & \text{(A.1)}\\
\alpha_{t+1} &= T\alpha_t + c + R\varepsilon_{g,t+1} & \varepsilon_{g,t+1} &\sim \mathrm{N}(0, Q_t), & \text{(A.2)}
\end{aligned}
$$

The state space matrices are defined as: $H = \Omega$, $d = \tilde{A}$, $c = 0$, $Q_t = L_g \mathrm{diag}\left[\exp\left(\Gamma h_t\right)\right] L_g'$ and

$$
R = \begin{pmatrix} I_G \\ 0_{(G+1)\times G} \end{pmatrix} \qquad Z = \begin{pmatrix} B & \iota & 0_{N\times G} \end{pmatrix} \qquad T = \begin{pmatrix} \Phi_g & 0_{G\times 1} & I - \Phi_g \\ 0_{(G+1)\times G} & & I_{G+1} \end{pmatrix}
$$

with the initial condition $\alpha_1 \sim \mathrm{N}(a_1, P_1)$. The mean and covariance matrix for the initial condition are

$$
a_1 = \left(\mu_{\tilde{\mu}_g}', \mu_{\delta_0}', \mu_{\tilde{\mu}_g}'\right)' \qquad P_1 = \begin{pmatrix} L_g \mathrm{diag}\left[\exp\left(\Gamma h_0\right)\right] L_g' + V_{\tilde{\mu}_g} & 0 & V_{\tilde{\mu}_g} \\ 0 & V_{\delta_0} & 0 \\ V_{\tilde{\mu}_g} & 0 & V_{\tilde{\mu}_g} \end{pmatrix}
$$

A discussion of the priors for the parameters of the model are in Appendix B.

Many steps of the MCMC algorithm use the conditionally, Gaussian state space form of (A.1) and (A.2). In particular, the draws for the parameters $\Phi_g^{\mathbb{Q}}, L_g, \Omega$ and Γ will be conducted by independence Metropolis Hastings (IMH). Importantly, we draw these parameters and use the Kalman filter to marginalize out the parameters $(\tilde{\mu}_g, \delta_0)$ and the Gaussian state variables $g_{1:T}$ from the MH acceptance ratio. We will repeatedly apply the IMH algorithm along the lines of Chib and Greenberg(1994). Suppose we separate the parameter vector $\theta = (\psi, \psi^-)$ and we want to draw a subset of the parameters ψ conditional on the remaining parameters ψ^-. For a generic parameter ψ, we do the following

- Use the Kalman filter to maximize the log-posterior $p\left(\psi | y_{1:T}, h_{0:T-1}, \psi^-\right) \propto p\left(y_{1:T} | h_{0:T-1}, \psi, \psi^-\right) p(\psi)$.

- Let $\hat{\psi} = \text{vech}(\psi)$ be the posterior mode and H_{ψ}^{-1} be the inverse Hessian at the mode.
- Draw a proposal $\psi^* \sim q(\psi)$ from a Student's t distribution with mean $\hat{\psi}$, scale matrix H_{ψ}^{-1}, and 5 degrees of freedom.
- The proposal ψ^* is accepted with probability $\alpha = \frac{p\left(y_{1:T}|h_{0:T-1},\psi^*,\psi^-\right)p(\psi^*)q\left(\psi^{(j-1)}\right)}{p\left(y_{1:T}|h_{0:T-1},\psi^{(j-1)},\psi^-\right)p\left(\psi^{(j-1)}\right)q(\psi^*)}$ where the likelihood $p(y_{1:T}|h_{0:T-1},\psi,\psi^-)$ is calculated by the Kalman filter.

A similar algorithm has been used by Chib and Ergashev(2009) for Gaussian models with macroeconomic factors but no stochastic volatility. For Gaussian ATSMs, our algorithm has two differences from theirs. First, we marginalize out (δ_0, μ_g) when drawing other parameters of the model. Secondly, they parameterize the model in terms of market price of risk parameters and use a different rotation of the state vector. Consequently, more parameters enter the bond loadings, which causes the log-likelihood surface to be more complicated.

Appendix A.1.2 MCMC algorithm

Our MCMC algorithm proceeds as follows:

STEP 1: Draw from the joint distribution $p(g_{1:T}, \tilde{\mu}_g, \delta_0, \Phi_g^{\mathbb{Q}}|y_{1:T}, h_{0:T-1}, L_g, \Omega, \Phi_g)$. First, draw $\Phi_g^{\mathbb{Q}}$, while marginalizing out $(g_{1:T}, \tilde{\mu}_g, \delta_0)$. Conditional on the draw of $\Phi_g^{\mathbb{Q}}$, draw the state variables and parameters $(g_{1:T}, \tilde{\mu}_g, \delta_0)$ in one block.

STEP 1a: Conditional on $h_{0:T-1}$ and the remaining parameters of the model, form the state space model (A.1)-(A.2). Draw $\Phi_g^{\mathbb{Q}}$ using an independence Metropolis-Hastings step as explained above.

STEP 1b: Conditional on $\Phi_g^{\mathbb{Q}}$, draw $(g_{1:T}, \tilde{\mu}_g, \delta_0)$ using the simulation smoother for model (A.1)-(A.2).

STEP 2: Conditional on $g_{1:T}, h_{0:T-1}, L_g$, and $\tilde{\mu}_g$, the model is a vector autoregression. Under a multivariate normal prior for Φ_g, the full conditional posterior for Φ_g is multivariate normal. The draw is standard for a Bayesian VAR; see, e.g. Del Negro and Schorfheide(2011).

STEP 3: Conditional on the current values of $(g_{1:T}, \mu_g, \Phi_g, L_g)$, calculate the observations $\tilde{g}_t = L_g^{-1}(g_t - \mu_g - \Phi_g g_{t-1})$. Using $\tilde{g}_{1:T}$, form a linear state space model.

$$\hat{g}_t = \Gamma h_t + \varepsilon_t^* \qquad \varepsilon_t^* \sim p(\varepsilon_t^*) \tag{A.3}$$

$$h_{t+1} = \tilde{\mu}_h + \Phi_h(h_t - \tilde{\mu}_h) + \varepsilon_{h,t+1} \qquad \varepsilon_{h,t+1} \sim \text{N}(0, \Sigma_h) \tag{A.4}$$

where the error $\varepsilon_{it}^* = \log\left(\varepsilon_{it}^2\right)$ and $\hat{g}_{it} = \log\left(\tilde{g}_{it}^2\right)$ for $i = 1, \ldots, G$. The distribution of the non-Gaussian error $p(\varepsilon_t^*)$ can be accurately approximated by a mixture of normals $p(\varepsilon_t^*) \approx \sum_{j=1}^{10} \pi_j \text{N}(\mu_j, \sigma_j^2)$ with known component weights π_j, means μ_j, and variances σ_j^2. In practice, we use the components of the mixture from Omori, Chib, Shephard, and Nakajima(2007). Conditional on a vector of indicator variables, the model is a conditionally linear, Gaussian state space model. In practice, we draw the unconditional mean $\tilde{\mu}_h$ instead of the intercept μ_h.

STEP 3a: Draw mixture indicators for the error ε_t^* as in Omori et al. (2007).

STEP 3b: The parameter $\tilde{\mu}_h$ enters the bond loadings A through our assumption that the covariance matrix under $\mathbb{Q}$ is equal to the covariance matrix under $\mathbb{P}$. Draw $\tilde{\mu}_h$ by IMH using the state space model (A.3) and (A.4).

STEP 3c: Draw Φ_h by IHM using the state space model (A.3) and (A.4).

STEP 3d: Draw Σ_h by IHM using the state space model (A.3) and (A.4).

STEP 3e: Draw the latent variables $h_{0:T-1}$ using the simulation smoother for the model (A.3) and (A.4). This last step is a draw from the joint distribution of Σ_h and $h_{0:T-1}$.

STEP 4: Conditional on $h_{0:T-1}$ and the remaining parameters of the model, form the state space model (A.1)-(A.2). Draw the diagonal elements of Ω using the IMH algorithm as explained above. In this step, our proposal distribution is for the log-variances instead of the variances. This requires adding a Jacobian term to the Metropolis-Hastings acceptance ratio.

STEP 5: Conditional on $h_{0:T-1}$ and the remaining parameters of the model, form the state space model (A.1)-(A.2). Draw L_g using the IMH algorithm as explained above.

STEP 6: Conditional on $h_{0:T-1}$ and the remaining parameters of the model, form the state space model (A.1)-(A.2). Draw Γ using the IMH algorithm as explained above.

Appendix A.2 Particle filter

The particle filter we implement is the mixture Kalman filter of Chen and Liu(2000). For a survey of particle filtering see Creal(2012). Let $g_{t|t-1}$ denote the conditional mean and $P_{t|t-1}$ the conditional covariance matrix of the predictive distribution $p(g_t|Y_{1:t-1}, h_{0:t-1}; \theta)$ of a conditionally linear, Gaussian state space model. Similarly, let $g_{t|t}$ denote the conditional mean and $P_{t|t}$ the conditional covariance matrix of the filtering distribution $p(g_t|Y_{1:t}, h_{0:t}; \theta)$. Conditional on knowing the volatilities $h_{0:T}$, these quantities can be calculated by the Kalman filter. Using the notation from Durbin and Koopman(2012), the linear, Gaussian state space model is

$$
\begin{aligned}
Y_t &= Zg_t + d + \eta_t & \eta_t \sim \mathrm{N}(0, H) \\
g_{t+1} &= Tg_t + c + R\varepsilon_{g,t+1} & \varepsilon_{g,t+1} \sim \mathrm{N}(0, Q_t)
\end{aligned}
$$

where $H = \Omega$, $Z = B$, $d = A$, $c = (I_G - \Phi_g)\,\tilde{\mu}_g$, $T = \Phi_g$, $R = I$ and with $Q_t = L_g \mathrm{diag}\left[\exp\left(\Gamma h_t\right)\right] L_g'$. 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 $i = 1, \ldots, N$, set $w_0^{(i)} = \frac{1}{N}$ and

- Draw $h_0^{(i)} \sim p\left(h_0; \theta\right)$ and calculate $Q_1^{(i)} = L_g \mathrm{diag}\left[\exp\left(\Gamma h_0^{(1)}\right)\right] L_g'$.
- Set $P_{1|0}^{(i)} = Q_1^{(i)}$, $g_{1|0}^{(i)} = \tilde{\mu}_g$ and $\ell_0 = 0$.

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

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

- Calculate $Q_t^{(i)} = L_g \mathrm{diag}\left[\exp\left(\Gamma h_t^{(i)}\right)\right] L_g'$.
- Run the Kalman filter:

$$
\begin{aligned}
v_t^{(i)} &= Y_t - Zg_{t|t-1}^{(i)} - d \\
F_t^{(i)} &= ZP_{t|t-1}^{(i)} Z' + H \\
K_t^{(i)} &= P_{t|t-1}^{(i)} Z' \left(F_t^{(i)}\right)^{-1} \\
g_{t|t}^{(i)} &= g_{t|t-1}^{(i)} + K_t^{(i)} v_t^{(i)} \\
P_{t|t}^{(i)} &= P_{t|t-1}^{(i)} - K_t^{(i)} Z P_{t|t-1}^{(i)} \\
g_{t+1|t}^{(i)} &= Tg_{t|t}^{(i)} + c \\
P_{t+1|t}^{(i)} &= TP_{t|t}^{(i)} T' + Q_t^{(i)}
\end{aligned}
$$

- Draw from the transition density: $h_{t+1}^{(i)} \sim p(h_{t+1}|h_t^{(i)}; \theta)$.
- Calculate the weight: $\log\left(w_t^{(i)}\right) = \log\left(\hat{w}_{t-1}^{(i)}\right) - 0.5N\log\left(2\pi\right) - 0.5\log|F_t^{(i)}| - \frac{1}{2} v_t^{(i)\prime}\left(F_t^{(i)}\right)^{-1} v_t^{(i)}$.

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

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

STEP 4: Calculate the effective sample size $E_t = \frac{1}{\sum_{i=1}^N \left(w_t^{(i)}\right)^2}$.

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

STEP 6: Increment time and return to STEP 1.

Appendix B Prior distributions

Throughout this discussion, a normal distribution is defined as $x \sim \mathrm{N}(\mu_x, V_x)$, where μ_x is the mean and V_x is the covariance matrix. The prior distributions on the parameters of the model are as follows.

- The parameter δ_0 has a normal prior $\delta_0 \sim \mathrm{N}(0.008, 0.1)$.
- The diagonal elements of $\Phi_g^{\mathbb{Q}}$ each have a beta distribution $\phi_{g,ii}^{\mathbb{Q}} \sim \mathrm{beta}(a_i, b_i)$ for $i = 1, \dots G$. For the three factor model, we set $a_1 = 100, b_1 = 1, a_2 = 200, b_2 = 10, a_3 = 100, b_3 = 60$.
- The parameter $\tilde{\mu}_g$ is multivariate normal $\tilde{\mu}_g \sim \mathrm{N}(0_{G\times 1}, \mathrm{I}_G * 0.05)$.
- The parameter Φ_g has a truncated multivariate normal distribution $\mathrm{vec}(\Phi_g) \sim \mathrm{N}\left(\mu_{\Phi_g}, \mathrm{I}_{G^2} * 0.05\right)$. We set $\mu_{\Phi_g} = (0.98, 0, 0, 0, 0.95, 0, 0, 0, 0.85)'$. The distribution has been truncated to the stationary region.
- The parameter $\tilde{\mu}_h$ is multivariate normal $\tilde{\mu}_h \sim \mathrm{N}(0_{H\times 1}, I_H)$.
- The parameter Φ_h has a truncated multivariate normal distribution $\mathrm{vec}(\Phi_h) \sim \mathrm{N}(\mu_{\Phi_h}, \mathrm{I}_{H^2} * 0.025)$. We set $\mu_{\Phi_h} = (0.98, 0, 0, 0, 0.98, 0, 0, 0, 0.98)'$. The distribution has been truncated to the stationary region.
- The parameter Σ_h has an inverse Wishart distribution $\Sigma_h \sim \mathrm{InvWishart}(\nu, \Omega)$. We set $\nu = H + 1$ and $\Omega = (\nu - H - 1) * \mathrm{I}_H * 0.2$. The distribution is defined such that $\mathbb{E}[\Sigma_h] = \frac{\Omega}{\nu - H - 1}$.
- For the models with $H = 1, 2$ factors, we estimate parameters in the matrix Γ. For $H = 1$, we have $\Gamma' = \begin{pmatrix} \gamma_{1,1} & 1 & \gamma_{3,1} \end{pmatrix}$. These parameters have normal distributions $\gamma_{i,1} \sim \mathrm{N}(1, 0.5)$. For $H = 2$, we have $\Gamma' = \begin{pmatrix} 1 & 0 & \gamma_{3,1} \\ 0 & 1 & \gamma_{3,2} \end{pmatrix}$. These parameters have normal distributions $\gamma_{3,1} \sim \mathrm{N}(0.5, 0.5)$ and $\gamma_{3,2} \sim \mathrm{N}(0.5, 0.5)$.
- The off-diagonal, lower triangular elements of L_g have a normal distribution $L_{g,ij} \sim \mathrm{N}(0, 1)$.