BANK OF ENGLAND

Staff Working Paper No. 575

Long-run priors for term structure models

Andrew Meldrum and Matt Roberts-Sklar

December 2015

BANK OF ENGLAND

Staff Working Paper No. 575

Long-run priors for term structure models

Andrew Meldrum[(1)] and Matt Roberts-Sklar[(2)]

Abstract

Dynamic no-arbitrage term structure models are popular tools for decomposing bond yields into expectations of future short-term interest rates and term premia. But there is insufficient information in the time series of observed yields to estimate the unconditional mean of yields in maximally flexible models. This can result in implausibly low estimates of long-term expected future short-term interest rates, as well as considerable uncertainty around those estimates. This paper proposes a tractable Bayesian approach for incorporating prior information about the unconditional means of yields. We apply it to UK data and find that with reasonable priors it results in more plausible estimates of the long-run average of yields, lower estimates of term premia in long-term bonds and substantially reduced uncertainty around these decompositions in both affine and shadow rate term structure models.

Key words: Affine term structure model, shadow rate term structure model, Gibbs sampler.

JEL classification: C11, E43, G12.

---

(1) Bank of England. Email: andrew.meldrum@bankofengland.co.uk
(2) Bank of England. Email: matt.roberts-sklar@bankofengland.co.uk

The views expressed in this paper are those of the authors and do not necessarily reflect those of the Bank of England or members of its committees. The authors would like to thank Don Kim, Wolfgang Lemke, Peter Spencer, Thomas Werner, participants in seminars at the Bank of England and European Central Bank and an anonymous referee for helpful comments on this paper.

Information on the Bank's working paper series can be found at
www.bankofengland.co.uk/research/Pages/workingpapers/default.aspx

Publications Team, Bank of England, Threadneedle Street, London, EC2R 8AH
Telephone +44 (0)20 7601 4030 Fax +44 (0)20 7601 3298 email publications@bankofengland.co.uk

ISSN 1749-9135 (on-line)

1 Introduction

Dynamic no-arbitrage term structure models are popular tools for analysing the joint dynamics of bond yields of different maturities. Policymakers routinely use these models to estimate expectations of future short-term interest rates and the additional term premia implied by long-term bond yields. For example, as the then Chairman of the Federal Reserve, Ben Bernanke, explained in a speech on long-term interest rates in March 2013: "It is useful to decompose longer-term yields into three components: one reflecting expected inflation over the term of the security; another capturing the expected path of short-term real, or inflation-adjusted, interest rates; and a residual component known as the term premium. Of course, none of these components is observed directly, but there are standard ways of estimating them."

Unfortunately, the uncertainty around the decompositions obtained using these 'standard methods' - dynamic no-arbitrage term structure models - is substantial. This paper provides a simple and tractable method for incorporating prior information about the long-run mean of bond yields, which not only results in more plausible term structure decompositions for the UK but also reduces the estimated uncertainty around those decompositions substantially. This should have obvious appeal to policymakers and others concerned with the long-horizon properties of these models.

In maximally flexible no-arbitrage term structure models, the accuracy of term premium estimates is primarily determined by the accuracy with which we can estimate the dynamics of the pricing factors using the available time series of yields. Term premia are computed as the difference between model-implied yields and the model-implied average expected short-term interest rate over the relevant horizon. In the benchmark model considered in this paper, as is typically the case, the short-term rate is an affine function of a small set of pricing factors, which follow a first-order Gaussian Vector Autoregression (VAR). At very long maturities, the model-implied expectations are largely determined by the estimate of the unconditional mean in the VAR. But the short samples of yields typically available, together with general declines in yields over those samples, means that there is little sample information with which to estimate those unconditional means (a point made previously by Bauer et al. (2012)). Figure 1 plots UK nominal zero-coupon bond yields at selected maturities over the period since October 1992, when the UK first introduced an inflation targeting framework for monetary policy (the majority of studies using US data also tend to use a sample that starts in the 1980s or early 1990s). In common with other advanced economies, UK nominal yields generally fell through this period. This can result in implausibly low estimates of the unconditional mean of yields, which in turn means that long-maturity term premium estimates are likely to be too high.

<Insert Figure 1 here.>

To illustrate why this is the case, the blue line on Figure 2 plots the UK ten-year yield between October 1992 and December 2014; and the solid green line overlays a projection starting in October 1992 from a univariate first-order autoregressive model estimated using the full sample. As pointed out by Sims (2000), OLS estimates of autoregressive models using finite samples have a tendency to exaggerate the component of the sample variation that is deterministic conditional on the initial observation. In this example, the autoregressive model (broadly speaking) interprets the fall in the 10-year yield over the sample as an initial observation a long way above the unconditional mean and a subsequent deterministic reversion, lasting around 20 years, towards that mean. The unconditional mean - a little over 2% (shown by the dashed black line) - is below almost all of the sample data and the initial point is outside the central 95% of the unconditional distribution (shown by the dashed red lines).

<Insert Figure 2 here.>

While there is little information in the *data* with which to estimate the unconditional means of yields, it is nevertheless reasonable to suppose that we do have relevant *prior* information. Ignoring that information, and estimating the model with flat priors implies that we attach a higher prior weight on a steady state value of the short rate that is less than (say) zero than (say) between zero and 10%, which is inconsistent with what we consider to be plausible prior beliefs. If we were working with samples that were highly informative about the mean of yields this would not be too harmful. But when the data are not informative, as is the case in reality, it can result in model estimates that are less plausible *a priori*, as we demonstrate below. The lack of sample information with which to estimate unconditional means also results in the model-implied uncertainty around term premium estimates being extremely high.

Our approach to incorporating prior information about the long run is based on that of Villani (2009), who proposes to specify a prior about the unconditional mean in Bayesian VAR models. To implement this in no-arbitrage term structure models, we rotate the pricing factors into bond yields (which may not necessarily be the same yields used to estimate the model) and specify priors on the unconditional means of those yields. We can then draw the parameters of the time-series dynamics of the factors using the method of Villani (2009) within a Gibbs sampling procedure, that is otherwise very similar to the approach for estimating no-arbitrage affine term structure models proposed by Bauer (2015).

A number of alternative approaches have been proposed previously to address the underlying problem of uninformative samples for estimating the time-series dynamics of the pricing factors in no-arbitrage term structure models. One option is to incorporate additional information in the form of survey expectations of professional economists (proposed by Kim and Orphanides (2012) and applied to UK data by, among others, Joyce et al. (2010) and Guimarães (2014)). In the case of the UK, unfortunately, there are no long-horizon surveys of Bank Rate expectations available; and Malik and Meldrum (2014) show that incorporating short-term surveys can result in markedly inferior performance of affine term structure models against standard specification tests.

A second approach, taken by Cochrane and Piazessi (2008) among others is to impose zero restrictions on the price of risk, in order that estimates of the risk-neutral factor dynamics can inform the time-series dynamics. One approach is choosing zero restrictions for any parameters that are not significantly different from zero.[1] While this may help to identify the time series dynamics, Bauer et al. (2012) show that this does not have an economically meaningful impact on the properties of US term premia.

A third approach, proposed by Bauer et al. (2012), is to use statistical techniques to correct for the small-sample bias of the OLS estimator of the factor dynamics in a classical setting. But that approach is focussed more on the persistence of the factors, rather than

[1] Bauer (2015) recently proposes a Bayesian approach for weighting models with different zero restrictions on the price of risk, in which the prior is specified to shrink towards more parsimonious models.

the intercept in the VAR. While our approach does not address the issue of small-sample bias in a classical setting, our approach for dealing with the intercept in the VAR using a Bayesian setting with informative priors is likely to have an important advantage.[2] Classical bias corrections are typically applied to demeaned data, so the *intercept* is effectively set in order to match the sample mean.[3] This may reduce the problem of underestimating the mean in *some* samples but in general the sample average may also be unlikely *a priori*. Moreover, by calibrating the intercept we are likely to *understate* our true uncertainty about term premium estimates. Estimating the intercept but allowing for prior information to inform that estimate is likely to result in more reasonable estimates of the true uncertainty (conditional on a particular model).

Section 2 of this paper describes our benchmark no-arbitrage affine term structure model. Section 3 describes the techniques we use to estimate it and the choice of priors. Section 4 reports results from the benchmark model. Section 5 shows how we can modify the framework slightly to allow the application to the shadow rate term structure model proposed by Black (1995), which is consistent with a zero lower bound on nominal interest rates; as far as we are aware, ours is the first study to estimate a shadow rate model using Bayesian techniques. Section 6 extends the benchmark model to decompose the term premium on a long-term nominal bond into components compensating investors for real interest rate and inflation risk, using a joint model of nominal and real yields similar to those previously applied to UK data by Joyce et al. (2010), D'Amico et al. (2014) and Guimarães (2014). Section 7 concludes.

2 Model

2.1 Affine term structure model

This section sets out the (entirely standard) benchmark affine term structure model of nominal bond yields. A nominal n-period zero-coupon bond pays £1 at its maturity after

[2] Jarocinski and Marcet (2010) discuss the difference between Bayesian and classical interpretations of bias in OLS estimates of autoregressive models.

[3] Adrian et al. (2013) do not apply a small-sample bias correction but do nevertheless calibrate the intercept in the VAR so that the unconditional mean of the pricing factors matches the sample average. Malik and Meldrum (2014) apply the same approach to UK data.

n periods. In the absence of arbitrage, the time-t price ($P_t^{(n)}$) is equal to the expected discounted present value of the price at time $t+1$:

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

where i_t is the one-period nominal risk-free rate and expectations are taken with respect to the risk-neutral probability measure, denoted $\mathbb{Q}$. The short-term rate is an affine function of an $K \times 1$ vector of pricing factors $\mathbf{x}_t$:

$$i_t = \delta_0 + \boldsymbol{\delta}_1' \mathbf{x}_t. \tag{2}$$

The factors follow a first-order Gaussian Vector Autoregression (VAR) under $\mathbb{Q}$:

$$\begin{aligned} \mathbf{x}_{t+1} &= \boldsymbol{\mu}^{\mathbb{Q}} + \boldsymbol{\Phi}^{\mathbb{Q}} \mathbf{x}_t + \mathbf{v}_{t+1}^{\mathbb{Q}}, \\ \mathbf{v}_t^{\mathbb{Q}} &\sim i.i.d. \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Sigma}\right). \end{aligned} \tag{3}$$

Given the above assumptions, nominal bond yields are affine functions of the factors:

$$y_t^{(n)} = -\frac{1}{n}\left(a_n + \mathbf{b}_n' \mathbf{x}_t\right), \tag{4}$$

where a_n and $\mathbf{b}_n$ follow the standard recursive equations

$$a_n = a_{n-1} + \mathbf{b}_{n-1}' \boldsymbol{\mu}^{\mathbb{Q}} + \frac{1}{2} \mathbf{b}_{n-1}' \boldsymbol{\Sigma} \mathbf{b}_{n-1} - \delta_0 \tag{5}$$

$$\mathbf{b}_n' = \mathbf{b}_{n-1}' \boldsymbol{\Phi}^{\mathbb{Q}} - \boldsymbol{\delta}_1', \tag{6}$$

with the initial conditions $a_0 = 0$ and $\mathbf{b}_0 = \mathbf{0}$. As has been discussed widely elsewhere (e.g. Dai and Singleton (2000); Hamilton and Wu (2012)) the model is not identified without additional parameter restrictions. We adopt the normalisation $\boldsymbol{\delta}_1 = \mathbf{1}_{(K\times 1)}$, $\boldsymbol{\mu}^{\mathbb{Q}} = \mathbf{0}_{(K\times 1)}$ and $\boldsymbol{\Phi}^{\mathbb{Q}} = diag\left\{[\phi_1, \phi_2, ..., \phi_K]\right\}$, with $1 > \phi_1 > \phi_2 > ... > \phi_K > 0$.

Following Duffee (2002), we assume that the market prices of risk are affine in the pricing factors, which implies that the pricing factors also follow a first-order Gaussian VAR under the real-world probability measure:

$$\begin{aligned} \mathbf{x}_{t+1} &= \boldsymbol{\mu} + \boldsymbol{\Phi}\mathbf{x}_t + \mathbf{v}_{t+1} \quad (7) \\ \mathbf{v}_t &\sim i.i.d.\mathcal{N}\left(\mathbf{0}, \boldsymbol{\Sigma}\right). \end{aligned}$$

As is standard, we define the term premium component of an n-period yield as the difference between the model-implied yield and the average expected short-term rate over the lifetime of the bond:

$$TP_t^{(n)} = y_t^{(n)} - \frac{1}{n}\sum_{i=0}^{n-1} E_t i_{t+i}. \quad (8)$$

3 Estimation

3.1 Data and factor structure

The nominal yields we use to estimate the model have maturities of 1, 12, 24, 36, 48, 60, 84 and 120 months. All except the one-month nominal rate are estimated using the smoothed cubic spline technique of Anderson and Sleath (2001) and are published by the Bank of England.[4] As this dataset does not consistently include nominal maturities shorter than one year, we augment it by using Bank Rate, the United Kingdom monetary policy interest rate, as a proxy for the one-month rate.

As is standard in the dynamic term structure literature, our benchmark model has three pricing factors. We assume that three yields (collected in the vector $\mathbf{y}_{1,t} = \left[y_t^{(12)}, y_t^{(36)}, y_t^{(120)}\right]'$ are observed without error and the remaining 5 yields ($\mathbf{y}_{2,t} = \left[y_t^{(1)}, y_t^{(24)}, y_t^{(48)}, y_t^{(60)}, y_t^{(84)}\right]'$) are observed with errors $\mathbf{w}_t$. This means that the measurement equations of the model can be written as

$$\begin{aligned} \mathbf{y}_{1,t} &= \mathbf{A}_1 + \mathbf{B}_1\mathbf{x}_t \quad (9) \\ \mathbf{y}_{2,t} &= \mathbf{A}_2 + \mathbf{B}_2\mathbf{x}_t + \mathbf{w}_t \quad (10) \\ \mathbf{w}_t &\sim i.i.d.\mathcal{N}\left(0, \mathbf{R}_w\right) \end{aligned}$$

[4] The data are available from: http://www.bankofengland.co.uk/statistics/pages/yieldcurve/default.aspx.

where the definitions of $\mathbf{A}_1$, $\mathbf{B}_1$, $\mathbf{A}_2$ and $\mathbf{B}_2$ follow from (4). Conditional on values of δ_0, $\boldsymbol{\delta}_1$, $\boldsymbol{\mu}^{\mathbb{Q}}$, $\boldsymbol{\Phi}^{\mathbb{Q}}$ and $\boldsymbol{\Sigma}$, we can use the procedure of Chen and Scott (1993) to invert (9) and recover the pricing factors, i.e. $\mathbf{x}_t = \mathbf{B}_1^{-1}(\mathbf{y}_{1,t} - \mathbf{A}_1)$.[5]

3.2 Gibbs sampling procedure

We estimate the model using Bayesian methods, splitting the parameters into six blocks and using a Gibbs sampler to draw from the conditional posteriors of each in turn: (i) the parameters governing the dynamics of the factors under the time series measure ($\boldsymbol{\Phi}$); (ii) the intercepts under the time-series dynamics ($\boldsymbol{\mu}$); (iii) the parameters governing the dynamics of the factors under $\mathbb{Q}$ ($\boldsymbol{\Phi}^{\mathbb{Q}}$); (iv) the intercept in the short rate equation (δ_0); (v) the factor shock covariance matrix $\boldsymbol{\Sigma}$; and (vi) the covariance matrix of measurement errors $\mathbf{R}_w$. The following sub-sections of the paper explain how each parameter block is drawn in turn. To obtain initial values for the chain we first estimate the parameters by maximum likelihood using the Chen and Scott (1993) procedure. We draw 10,000 times, discarding the first 5,000 draws as burn-in.

The approach for blocks (iii)-(vi) is very similar to that proposed by Bauer (2015).[6] The most substantial innovation in this paper is the process for drawing the parameters of the time-series dynamics ((i) and (ii)). Whereas Bauer draws the parameters of the market prices of risk which relate the time-series and risk-neutral factor dynamics, we instead draw the time-series dynamics directly.

3.2.1 Time series dynamics ($\boldsymbol{\mu}$ and $\boldsymbol{\Phi}$)

A typical approach to specifying a prior for a Bayesian VAR would be to assume that (conditional on $\boldsymbol{\Sigma}$) $\boldsymbol{\mu}$ and $\boldsymbol{\Phi}$ are jointly Normally distributed under the prior. In our case, however, it is not obvious how to specify a meaningful prior over $\boldsymbol{\mu}$. But it *is* reasonable to believe that we have prior information about the long-run mean of bond yields. To

[5] Bauer (2015) uses the normalisation of Joslin et al. (2011), which means that he can treat the factors as observed principal components of yields within a Gibbs sampler. It is not obvious, however, that this normalisation can be applied to the shadow rate model we estimate in Section 5. We choose to use the Dai and Singleton (2000) normalisation for consistency across the different models reported in this paper.

[6] Other studies that have estimated dynamic term structure models using Bayesian methods include Chib and Ergashev (2009), Ang et al. (2011) and Andreasen and Meldrum (2013).

implement such a long-run prior in an affine term structure model, we assume there are K independent linear combinations of bond yields about which we have some prior information, which we denote

$$\mathbf{x}_t^* = \mathbf{W}'\mathbf{y}_t \tag{11}$$

where $\mathbf{W}$ is a $K \times N$ matrix of full rank. We can write the reduced-form time-series dynamics of these yields as

$$\begin{aligned} \mathbf{x}_{t+1}^* &= \boldsymbol{\mu}^* + \boldsymbol{\Phi}^*\mathbf{x}_t^* + \mathbf{v}_{t+1}^*, \\ \mathbf{v}_t^* &\sim i.i.d.\mathcal{N}\left(\mathbf{0}, \boldsymbol{\Sigma}^*\right). \end{aligned} \tag{12}$$

Using (7), (11) and (12), we can solve for the structural parameters $\boldsymbol{\mu}$, $\boldsymbol{\Phi}$ and $\boldsymbol{\Sigma}$ in terms of $\boldsymbol{\mu}^*$, $\boldsymbol{\Phi}^*$ and $\boldsymbol{\Sigma}^*$:

$$\boldsymbol{\mu} = \left(\mathbf{W}'\mathbf{B}\right)^{-1}\left(\boldsymbol{\mu}^* - \mathbf{W}'\mathbf{A} + \boldsymbol{\Phi}^*\mathbf{W}'\mathbf{A}\right) \tag{13}$$

$$\boldsymbol{\Phi} = \left(\mathbf{W}'\mathbf{B}\right)^{-1}\boldsymbol{\Phi}^*\mathbf{W}'\mathbf{B} \tag{14}$$

$$\boldsymbol{\Sigma} = \left(\mathbf{W}'\mathbf{B}\right)^{-1}\boldsymbol{\Sigma}^*\left(\mathbf{B}'\mathbf{W}\right)^{-1} \tag{15}$$

We can also re-write (12) in terms of deviations from the unconditional mean of $\mathbf{x}_t^*$, $\boldsymbol{\gamma} = E\left[\mathbf{x}_t^*\right] = \left(\mathbf{I} - \boldsymbol{\Phi}^*\right)^{-1}\boldsymbol{\mu}^*$:

$$\widetilde{\mathbf{x}}_{t+1}^* = \mathbf{x}_{t+1}^* - \boldsymbol{\gamma} = \boldsymbol{\Phi}^*\widetilde{\mathbf{x}}_t^* + \mathbf{v}_{t+1}^*. \tag{16}$$

Stacking this equation across t gives

$$\widetilde{\mathbf{X}}_+ = \widetilde{\mathbf{X}}_-\boldsymbol{\Phi}^* + \mathbf{V}, \tag{17}$$

where $\widetilde{\mathbf{X}}_+ = \left[\widetilde{\mathbf{x}}_2^*, \widetilde{\mathbf{x}}_3^*, ..., \widetilde{\mathbf{x}}_T^*\right]'$ and $\widetilde{\mathbf{X}}_- = \left[\widetilde{\mathbf{x}}_1^*, \widetilde{\mathbf{x}}_2^*, ..., \widetilde{\mathbf{x}}_{T-1}^*\right]'$. If we assume an independent Normal prior for $\boldsymbol{\phi}^* = vec\left(\boldsymbol{\Phi}^*\right)$**:**

$$\boldsymbol{\phi}^*|\boldsymbol{\Sigma}^* \sim \mathcal{N}\left(\underline{\boldsymbol{\phi}}, \underline{\mathbf{V}}_\phi\right), \tag{18}$$

it is straightforward to generate a draw from the posterior:

$$\boldsymbol{\phi}^{*}|\boldsymbol{\Sigma},\mathbf{X}\sim\mathcal{N}\left(\overline{\boldsymbol{\phi}},\overline{\mathbf{V}}_{\phi}\right), \tag{19}$$

where

$$\begin{aligned}\overline{\mathbf{V}}_{\phi} &= \left(\underline{\mathbf{V}}_{\phi}^{-1}+\boldsymbol{\Sigma}^{*-1}\otimes\widetilde{\mathbf{X}}'_{-}\widetilde{\mathbf{X}}_{-}\right)^{-1}\\ \overline{\boldsymbol{\phi}} &= \overline{\mathbf{V}}_{\phi}\left(\underline{\mathbf{V}}_{\phi}^{-1}\underline{\boldsymbol{\phi}}+\left(\boldsymbol{\Sigma}^{*-1}\otimes\mathbf{I}\right)vec\left(\widetilde{\mathbf{X}}'_{-}\widetilde{\mathbf{X}}_{+}\right)\right).\end{aligned}$$

We set $\underline{\mathbf{V}}_{\phi}^{-1}=\mathbf{0}_{K^2}$ but impose a prior that yields are stationary by rejecting any draws that imply eigenvalues that are outside the unit circle.

Turning to the intercept, we can re-write (12), substituting $\boldsymbol{\gamma}$ for $\boldsymbol{\mu}^{*}$:

$$(\mathbf{I}-\boldsymbol{\Phi}^{*})^{-1}\left(\mathbf{x}^{*}_{t+1}-\boldsymbol{\Phi}^{*}\mathbf{x}^{*}_{t}\right)=\boldsymbol{\gamma}+(\mathbf{I}-\boldsymbol{\Phi}^{*})^{-1}\mathbf{v}^{*}_{t+1}. \tag{20}$$

Stacking this across t, we can re-write it as

$$\boldsymbol{\Xi}=\left(\mathbf{X}_{+}-\mathbf{X}_{-}\boldsymbol{\Phi}^{*}\right)(\mathbf{I}-\boldsymbol{\Phi}^{*})^{-1\prime}=\boldsymbol{\iota}_{T}\boldsymbol{\gamma}'+\mathbf{V}(\mathbf{I}-\boldsymbol{\Phi}^{*})^{-1\prime}, \tag{21}$$

where $\mathbf{X}_{+}=[\mathbf{x}^{*}_{2},\mathbf{x}^{*}_{3},...,\mathbf{x}^{*}_{T}]'$, $\mathbf{X}_{-}=\left[\mathbf{x}^{*}_{1},\mathbf{x}^{*}_{2},...,\mathbf{x}^{*}_{T-1}\right]'$ and $\boldsymbol{\iota}_{T}$ is a $T\times 1$ vector of ones. As proposed by Villani (2009), we assume a Normal prior for $\boldsymbol{\gamma}$:

$$\boldsymbol{\gamma}|\boldsymbol{\Sigma}^{*},\boldsymbol{\Phi}^{*}\sim\mathcal{N}\left(\underline{\boldsymbol{\gamma}},\underline{\mathbf{V}}_{\gamma}\right). \tag{22}$$

As discussed below, the benchmark model estimated in this paper uses three pricing factors. We specify the following mean and variance for the long-run prior over $\mathbf{x}^{*}_{t}=\left[y^{(1)}_{t},y^{(60)}_{t},y^{(120)}_{t}\right]'$:

$$\begin{aligned}\underline{\boldsymbol{\gamma}} &= \begin{bmatrix}\frac{4.5}{1200} & \frac{5}{1200} & \frac{5.5}{1200}\end{bmatrix}',\\ \underline{\mathbf{V}}_{\gamma} &= diag\left\{\begin{bmatrix}\frac{0.25}{1200^2} & \frac{0.5}{1200^2} & \frac{1}{1200^2}\end{bmatrix}\right\}.\end{aligned}$$

The prior mean of the unconditional average of the nominal one-month rate is 4.5% (in annualised percentage points). We can rationalise this as reflecting, for example, a 2% expected inflation rate and an average short-term real interest rate of 2.5%. The prior variance is such that there is a 95% probability that the unconditional mean lies between 3.5% and 5.5%. The prior means for the unconditional averages of longer maturity nominal yields are higher, consistent with a term structure that slopes upwards on average.[7] Reflecting our uncertainty about the average size and sign of term premia, however, the prior variance is also increasing with maturity. For the 10-year nominal yield, it implies that there is a roughly 68% probability that the average ten-year yield will be between 4.5% and 6.5%. Below we also report results from a model with a flat prior over $\boldsymbol{\gamma}$, which is equivalent to setting $\underline{\mathbf{V}}_{\gamma}^{-1} = \mathbf{0}$.

Given this prior, it is straightforward to draw from the posterior, which is given by:

$$\boldsymbol{\phi}|\boldsymbol{\Sigma}^*, \boldsymbol{\Phi}^*, \mathbf{X}_+ \sim \mathcal{N}\left(\overline{\boldsymbol{\gamma}}, \overline{\mathbf{V}}_{\gamma}\right), \tag{23}$$

where

$$\begin{aligned}
\overline{\mathbf{V}}_{\gamma} &= \left(\underline{\mathbf{V}}_{\gamma}^{-1} + T\left(\left(\mathbf{I}-\boldsymbol{\Phi}^*\right)^{-1}\boldsymbol{\Sigma}^*\left(\mathbf{I}-\boldsymbol{\Phi}^*\right)^{-1\prime}\right)^{-1}\right)^{-1} \\
\overline{\boldsymbol{\gamma}} &= \overline{\mathbf{V}}_{\gamma}\left(\underline{\mathbf{V}}_{\gamma}^{-1}\underline{\boldsymbol{\gamma}} + \left(\left(\mathbf{I}-\boldsymbol{\Phi}^*\right)^{-1}\boldsymbol{\Sigma}^*\left(\mathbf{I}-\boldsymbol{\Phi}^*\right)^{-1\prime}\right)^{-1} vec\left(\boldsymbol{\iota}_T'\boldsymbol{\Xi}\right)\right).
\end{aligned}$$

In summary, the algorithm for drawing the values of $\boldsymbol{\mu}$ and $\boldsymbol{\Phi}$ at the i^{th} step in the Gibbs sampler is:

- conditional on the $i-1^{th}$ draw $\left\{\boldsymbol{\Sigma}^{(i-1)}, \delta_0^{(i-1)}, \boldsymbol{\Phi}^{\mathbb{Q}(i-1)}\right\}$, compute the implied value of $\boldsymbol{\Sigma}^{*(i-1)}$ using (15);
- draw $\boldsymbol{\Phi}^{*(i)}$ from the posterior distribution (19);
- draw $\boldsymbol{\gamma}^{(i)}$ from the posterior distribution (23); and
- compute the implied values of $\boldsymbol{\mu}^{(i)}$ and $\boldsymbol{\Phi}^{(i)}$ using (13) and (14).

[7] Chib and Ergashev (2009) also assume a prior that involves an upward sloping term structure on average.

3.3 $\mathbb{Q}$ parameters (δ_0 and $\mathbf{\Phi}^{\mathbb{Q}}$)

We draw the parameters governing the $\mathbb{Q}$ dynamics of the factors ($\mathbf{\Phi}^{\mathbb{Q}}$) and the short-term interest rate (δ_0) using Metropolis-within-Gibbs steps, very similar to those proposed by Bauer (2015). We parameterise $\mathbf{\Phi}^{\mathbb{Q}}$ as $\mathbf{\Phi}^{\mathbb{Q}} = \mathbf{I} + diag\left\{\boldsymbol{\phi}^{\mathbb{Q}}\right\}$, where $\phi_i^{\mathbb{Q}} = \sum_{j=1}^{i} \theta_j$ and restrict $-1 < \theta_j < 0$. We assume an independent prior over θ_j:

$$1 + \theta_j \sim \mathcal{B}\left(a, b\right)$$

where $\mathcal{B}$ denotes the density of a beta distribution and we set $a = 1000$ and $b = 10$. In initial investigations with a flat prior (as used by Bauer (2015)), we found that the posterior distributions for a number of parameters became extremely wide and the Gibbs sampler spent extremely long periods exploring regions with θ_j close to zero, where the likelihood surface becomes extremely flat. Our prior is nevertheless consistent with all factors being highly persistent under $\mathbb{Q}$ (the prior mean of θ_j is approximately -0.01) but relative to a flat prior downweights the possibility that θ_j is greater than about -10^{-5}.

At the i^{th} draw in the chain we draw a candidate parameter vector $\boldsymbol{\theta}'$ according to

$$\boldsymbol{\theta}' \sim \mathcal{T}_5\left(\boldsymbol{\theta}^{(i-1)}, \mathbf{\Omega}_\theta\right), \tag{24}$$

where $\mathcal{T}_5$ denotes the density of a multivariate Student's t-distribution with five degrees of freedom; $\boldsymbol{\theta}^{(i-1)}$ is the $i-1^{th}$ draw in the chain; and the proposal covariance $\mathbf{\Omega}_\theta$ is set equal to minus the inverse hessian of the $\mathbb{Q}$-likelihood, with respect to $\boldsymbol{\theta}$, evaluated at the initial values of the chain and tuned to achieve a reasonable Metropolis acceptance rate.[8] The procedure for sampling δ_0 is exactly analogous (except that the prior is flat).

Since we work with the normalisation of Dai and Singleton (2000), rather than that of Joslin et al. (2011), the procedure for drawing the parameters of the cross-section is slightly more complicated compared with Bauer (2015), in that we do not treat the factors as fixed. As described above, we invert them using the Chen and Scott (1993) procedure, so the

[8] Bauer (2015) allows for an adaptive proposal distribution by recomputing the hessian periodically through the chain. We find that this step is not necessary using UK data, so drop it to save on computational time.

values of the factors will depend on the values of $\mathbf{\Phi}^{\mathbb{Q}}$, δ_0 and $\mathbf{\Sigma}$ and so will be different for each draw.

3.4 Factor covariance ($\mathbf{\Sigma}$)

The procedure for drawing $\mathbf{\Sigma}$ is the same as that used by Bauer (2015). We assume a flat prior over the elements of $\mathbf{\Sigma}$ and use another Metropolis-within-Gibbs step to draw from the posterior. At the i^{th} draw in the chain, we draw a proposal $\mathbf{\Sigma}'$ according to

$$\mathbf{\Sigma}' \sim \mathcal{IW}\left(\nu, \mathbf{\Psi}_{\Sigma}^{(i)}\right), \tag{25}$$

where $\mathcal{IW}$ denotes the density of an inverse Wishart distribution; the shape parameter ν is tuned to achieve a reasonable acceptance rate; and the scale parameters $\mathbf{\Psi}_{\Sigma}^{(i)}$ are set such that the mean of the proposal distribution is equal to $\mathbf{\Sigma}^{(i-1)}$.

3.5 Measurement error covariance ($\mathbf{R}_w$)

We assume an independent inverse Wishart prior for $\mathbf{R}_w$:

$$\mathbf{R}_w \sim \mathcal{IW}\left(\underline{\nu}_w, \underline{\mathbf{\Psi}}_w\right) \tag{26}$$

with $\nu_w = N + 2$ and $\mathbf{\Psi}_w = 0.05 \times \mathbf{I}_N$ (i.e. a mean variance for each yield of five basis points). The posterior is given by

$$\mathbf{R}_w|\mathbf{Y}, \mathbf{X}, \delta_0, \boldsymbol{\theta}, \mathbf{\Sigma} \sim \mathcal{IW}\left(\overline{\nu}_w, \overline{\mathbf{\Psi}}_w\right), \tag{27}$$

where

$$\begin{aligned} \overline{\nu}_w &= \underline{\nu}_w + T \\ \overline{\mathbf{\Psi}}_w &= \underline{\mathbf{\Psi}}_w + \sum_{t=1}^{T} \mathbf{w}_t \mathbf{w}_t'. \end{aligned}$$

This differs very slightly from Bauer (2015), who assumes that the measurement error is independent across yields and has the same variance for all maturities (i.e. $\mathbf{R}_w = \sigma^2 \mathbf{I}_N$). We prefer to relax this assumption, partly to allow for a different measurement error variance for the proxy for the one-month yield and partly to allow for different measurement error variances across the nominal and real curves in the joint model reported in Section 6.

4 Results

4.1 Parameter estimates

Table 1 reports parameter estimates for the benchmark model with the long-run prior.[9] As is standard, the factors are highly persistent under the risk-neutral dynamics (the largest eigenvalue of $\mathbf{\Phi}^{\mathbb{Q}}$ has a posterior mean of 0.998). The factors are also persistent under the time-series measure, but the posterior distributions for the parameters Φ are much wider than those of the risk-neutral equivalent.

The posterior mean of the long-run mean of the short-term interest rate in the model with the long-run prior is 4.3%; and the posterior mean unconditional yield curve is upward sloping, with an average 10-year yield of 5.2%. Table 2 illustrates the impact of the long-run prior. With a flat prior over $\boldsymbol{\gamma}$, the unconditional mean of yields is implausibly low: at the posterior mean, the average short-tem rate is -4.2%, rising to only 0.3% for the 10-year yield. The posterior distributions are also much wider with a flat prior - for example, the central 90% of the posterior distribution covers the region between -14.3% and 5.7%.

<Insert Table 1 here.>

<Insert Table 2 here.>

4.2 Term premium estimates

In the model with flat priors over the long-run mean $\boldsymbol{\gamma}$, the fact that yields revert to implausibly low long-run means is likely to lead the models to underestimate the component of yields that reflects expected future policy rates. Between October 1992 and December 2014, the model-implied average expected short-term interest rates over ten-year horizons (shown in Figure 3) were on average around 2.8%. In late 2011, it was approximately 0%,

[9] In the interests of space, the table omits the parameters of the measurement error covariance matrix $\mathbf{R}_w$.

implying that the 10-year yield of around 3% was entirely made up of a term premium (Figure 4). And the uncetainty around these point estimates is wide. For example, the average width of the 80% posterior probability interval for the 10-year term premium is 2.7 percentage points (Figure 5).

<Insert Figure 3 here.>

<Insert Figure 4 here.>

<Insert Figure 5 here.>

The broad *changes* in the posterior mean term premium are similar in the model with the long-run prior (and are very similar to those reported previously by Malik and Meldrum (2014)). But the average expected short rate over a 10-year horizon between October 1992 and December 2014 is around one percentage point higher compared with the model with the flat prior, at 3.7% (Figure 6) and the term premium is correspondingly lower (Figure 7). The 80% posterior probability interval is also considerably narrower. For example, for the term premium component of yields it has a average width of around 1.9 percentage points (Figure 5).

<Insert Figure 6 here.>

<Insert Figure 7 here.>

5 Long-run priors in a shadow rate term structure model

One potential drawback of a Gaussian affine term structure model over our sample is that the model is not consistent with a lower bound on nominal interest rates. When interest rates are close to zero, as has been the case towards the end of our sample, this means that the model can imply a significant probability of negative nominal interest rates (a point made previously by a number of studies, including Andreasen and Meldrum (2013) and Bauer and Rudebusch (2014)). A potential concern could therefore be that the results in the previous section were driven by the fact that we were estimating an affine model over a period that ended with very low short-term interest rates. To demonstrate that this is not likely to be the case, this section shows that we can apply a similar long-run prior in a model that does impose the zero bound on nominal interest rates, with only minimal changes to the specification, and that term premium estimates from such a model are actually even lower than in the benchmark model.

In the shadow rate model, as proposed by Black (1995), the short-term interest rate is the maximum of zero and a 'shadow rate' of interest (s_t):

$$i_t = \max\{0, s_t\},$$

which is affine in the pricing factors

$$s_t = \delta_0 + \boldsymbol{\delta}_1' \mathbf{x}_t.$$

The risk-neutral (3) and time-series (7) dynamics of the pricing factors are the same as in the affine model. While the shadow rate specification ensures that bond yields are non-negative, unfortunately there are no closed-form expressions for yields as functions of the pricing factors and structural parameters of the model. We therefore use the second-order approximation to yields proposed by Priebsch (2013), applied previously in a discrete-time setting by Andreasen and Meldrum (2014) (for the US) and (for the UK) by Malik and Meldrum (2014) and Andreasen and Meldrum (2015).

Since the mapping between yields and factors is non-linear in the shadow rate model is non-linear, we cannot simply specify priors about the long-run values of bond yields by inverting the pricing factors. We can, however, specify priors on the 'shadow term structure', which is defined as

$$s_t^{(n)} = -\frac{1}{n}\left(a_n + \mathbf{b}_n' \mathbf{x}_t\right),$$

where a_n and $\mathbf{b}_n'$ follow the same recursive equations as in the affine model, i.e. (5) and (6) above. We can think of the shadow term structure as the bond yields that would apply if there were no lower bound on nominal interest rates. This is convenient, since it means that we can specify a long-run prior about the *shadow* term structure in exactly the same way as before.

A related complication when working with the shadow rate model (given the non-linear relationship between yields and factors) is that we can no longer extract the factors using the Chen and Scott (1993) inversion.[10] We instead assume that all N yields ($\mathbf{y}_t$) are observed with additive measurement error, i.e.

$$\begin{aligned} \mathbf{y}_t &= g\left(\mathbf{x}_t; \delta_0, \boldsymbol{\delta}_1, \boldsymbol{\mu}^{\mathbb{Q}}, \boldsymbol{\Phi}^{\mathbb{Q}}, \boldsymbol{\Sigma}\right) + \mathbf{w}_t \\ \mathbf{w}_t &\sim i.i.d.\mathcal{N}\left(0, \mathbf{R}_w\right) \end{aligned} \tag{28}$$

where $g\left(\mathbf{x}_t; \delta_0, \boldsymbol{\delta}_1, \boldsymbol{\mu}^{\mathbb{Q}}, \boldsymbol{\Phi}^{\mathbb{Q}}, \boldsymbol{\Sigma}\right)$ is the non-linear function given by the Priebsch (2013) approximation, and estimate the factors using an adaptation of the single-move procedure proposed by Jacquier et al. (1994). At the i^{th} step in the Gibbs sampler, for each time period in turn we construct a proposal $\mathbf{x}'_t$ according to

$$\mathbf{x}'_t \sim \mathcal{N}\left(\mathbf{x}_t^{(i-1)}, \mathbf{R}_{\mathbf{x}_t}^{CDKF}\right),$$

where $\mathbf{x}_t^{(i-1)}$ is the $i-1^{th}$ draw of the factors at time t and $\mathbf{R}_{\mathbf{x}_t}^{CDKF}$ is the filtered covariance matrix for $\mathbf{x}_t$ obtained using the Central Difference Kalman Filter of Norgaard et al. (2000) evaluated at the initial parameter values. We assume a flat prior over $\mathbf{x}_t$ and initialise the chain at the filtered values obtained by running a single pass of the Central Difference Kalman Filter, again at the initial parameter values.

Figure 8 shows estimates of the 10-year term premium from the shadow rate model with the long-run prior. Until the period of near-zero short-term interest rates towards the end of the sample, the posterior mean term premium estimates from the model are very similar to those from the affine model (7). More recently, however, the estimated term premium from the shadow rate model has been lower than that from the affine model. This contrasts slightly with previous findings by Kim and Priebsch (2013) (for the US) and Malik and Meldrum (2014) (for the UK) that long-maturity term premia from shadow rate models are similar to those from affine models. It is also striking that the model-implied uncertainty around the term premium estimates is much narrower than in the affine model during the recent period of very low nominal interest rates.

[10] In the classical literature on shadow rate models, the factors are typically estimated using a non-linear extension of the Kalman filter (e.g. Christensen and Rudebusch (2013), Kim and Priebsch (2013) and Bauer and Rudebusch (2014)) or using non-linear regression (e.g. Andreasen and Meldrum (2014))).

<Insert Figure 8 here.>

6 Decomposition into real and inflation term premia

6.1 Joint model of nominal and real bonds

Finally, in this section, we decompose the term premium on a 10-year nominal bond further, into a real term premium and inflation risk premium. The joint model of nominal and real yields that we use for these purposes is similar in structure to those previously applied to UK data by Joyce et al. (2010), D'Amico et al. (2014) and Guimarães (2014). The pricing of nominal bonds is the same as in the benchmark affine model described above. A real n-period zero-coupon bond pays one unit of a composite consumption good at its maturity after n periods. In the absence of arbitrage, the time-t price ($P_{t,R}^{(n)}$) is equal to the expected discounted present value of the price of an $n-1$-period bond at time $t+1$:

$$P_{t,R}^{(n)} = E_t^{\mathbb{Q}} \left[\exp\left(-r_t\right) P_{R,t+1}^{(n-1)} \right], \tag{29}$$

where $r_t = i_t - E_t \pi_{t+1}$ is the one-period real risk-free rate and π_{t+1} is the rate of inflation between t and $t+1$. The short-term real rate is also an affine function of an $K \times 1$ vector of pricing factors $\mathbf{x}_t$:

$$r_t = \delta_{0,R} + \boldsymbol{\delta}_{1,R}' \mathbf{x}_t. \tag{30}$$

Given the above assumptions, real bond yields are given by:

$$y_{t,R}^{(n)} = -\frac{1}{n} \left(a_{n,R} + \mathbf{b}_{n,R}' \mathbf{x}_t \right), \tag{31}$$

where a_n and $\mathbf{b}_n$ follow the recursive equations

$$a_{n,R} = a_{n-1,R} + \mathbf{b}_{n-1,R}' \boldsymbol{\mu}^{\mathbb{Q}} + \frac{1}{2} \mathbf{b}_{n-1,R}' \boldsymbol{\Sigma} \mathbf{b}_{n-1,R} - \delta_{0,R} \tag{32}$$

$$\mathbf{b}_{n,R}' = \mathbf{b}_{n-1,R}' \boldsymbol{\Phi}^{\mathbb{Q}} - \boldsymbol{\delta}_{1,R}', \tag{33}$$

with the initial conditions $a_{0,R} = 0$ and $\mathbf{b}_{0,R} = \mathbf{0}$. Real term premia are defined as

$$TP_{t,R}^{(n)} = y_{t,R}^{(n)} - \sum_{i=0}^{n-1} E_t r_{t+i}.$$

Inflation breakevens are defined as the difference between the nominal and real yields of the same maturities:

$$\pi_t^{(n)} = y_t^{(n)} - y_{t,R}^{(n)}, \tag{34}$$

and the inflation risk premium as the difference between the nominal and real term premia of the same maturity:

$$IRP_t^{(n)} = TP_t^{(n)} - TP_{t,R}^{(n)}. \tag{35}$$

To estimate the model, in addition to the nominal yields reported above, we also use zero-coupon real yields with maturities of 48, 60, 84 and 120 months, which are estimated by the Bank of England using UK government bonds indexed to the UK Retail Prices Index (RPI).[11] Our benchmark model has five pricing factors: three extracted from the nominal yield curve and two from the real yield curve. Table 3 reports the results of a preliminary principal components analysis, which provides support for this specification. Just two principal components are required to account for 99.9% of the variation in real yields (compared with three for nominal yields).[12]

<Insert Table 3 here.>

In addition to the same three nominal yields assumed to be measured without error in our benchmark nominal model, we also assume that the 48- and 120-month real yields are measured without error and adopt the normalisation $\boldsymbol{\delta}_1 = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \end{bmatrix}'$, $\boldsymbol{\delta}_{1,R} = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 \end{bmatrix}'$, $\boldsymbol{\mu}^{\mathbb{Q}} = \mathbf{0}_{(K\times 1)}$ and $\boldsymbol{\Phi}^{\mathbb{Q}} = diag\left\{\begin{bmatrix} \phi_{N1} & \phi_{N2} & \phi_{N3} & \phi_{R1} & \phi_{R2} \end{bmatrix}\right\}$, with $1 > \phi_{N1} > \phi_{N2} > \phi_{N3} > 0$ and $1 > \phi_{R1} > \phi_{R2} > 0$. This allows us to invert the factors using the Chen and Scott (1993) method.

We specify a prior for the unconditional mean of the vector $\mathbf{x}_t^* = \left[y_t^{(1)}, y_t^{(60)}, y_t^{(120)}, y_{R,t}^{(1)}, y_{R,t}^{(120)}\right]'$.

[11] Real yields are also estimated using the smoothed cubic spline method of Anderson and Sleath (2001) (and the method of Evans (1998) to address the indexation lag that applies to index-linked bonds). Real yields with maturities of less than 48 months are not available consistently through our sample.

[12] Joyce et al. (2012) also adopt a two-factor specification for an affine term structure model of the UK real curve.

The prior for the three nominal rates is the same as in the benchmark model. The prior mean of the unconditional average of the one-month *real* rate (recall that an advantage of our proposed method is that the prior can be formed over yields that are not observed in our sample) is 2%. This implies a prior mean of the one-month inflation rate of 2.5% annualised.[13] We assume the same prior mean for the 10-year real yield (which implies that the positive slope of the nominal term structure under the prior is due to an upward-sloping term structure of inflation breakevens). The prior variances are the same as for the nominal yields of the same maturity.

6.2 Real and inflation premia

Figures 9 and 10 show estimates of real term premia in models with flat priors over γ and our long-run prior respectively. The impact of the long-run prior is similar to that which we observed above for the nominal term premium in our benchmark model: while the broad pattern of movements in the real term premium is the same for the different priors, in the model with the long-run prior the level is lower and the 80% probability interval around the estimates is dramatically smaller for most of the sample.

The impact on the inflation risk premium, on the other hand, is much less pronounced (Figures 11 and 12). In both models, the inflation risk premium starts the sample period at around 2.5% and falls to around zero by the end of the 1990s, with the sharpest falls coming over the period after which the Bank of England was granted operational independence for monetary policy in May 1997.[14] The inflation risk premium in both models drifted up during the mid 2000s before falling sharply during the financial crisis of late 2008. Since then, it has generally been slightly positive. The 80% posterior probability interval around the estimates is broadly similar in both models.

To provide some intuition for this result, Figure 13 plots the 10-year inflation breakeven over our sample. Unlike the nominal and real yields, it appears to be much more obviously

[13] Recall that UK real government bonds are indexed to the RPI, whereas the UK Monetary Policy Committee's inflation target of 2% refers to the Consumer Prices Index (CPI). RPI inflation is on average higher than CPI inflation, partly reflecting a different composition of the basket of goods used to compute the index and partly due to differences in calculation.

[14] This fall in inflation risk premia in the 1990s is a common result in studies using UK data (e.g. Joyce et al. (2010), D'Amico et al. (2014), Guimarães (2014) and Abrahams et al. (2015)). Both Guimarães (2014) and Abrahams et al. (2015) find a lower average level of the inflation risk premium than we do.

stationary, so the sample information for estimating the long-run mean of inflation is likely to be much more informative. Table 4 reports estimates of the posterior distribution of the long-run mean parameters, which confirm this. While the posterior distributions for the long-run means of the nominal and real yields are extremely wide in the model with the flat prior over $\boldsymbol{\gamma}$, those for the long-run mean of the 1-month and 10-year inflation breakevens ($\gamma_1 - \gamma_4$ and $\gamma_3 - \gamma_5$ respectively) are much narrower. These are narrower still in the model with the long-run prior but the difference is not as dramatic as for the nominal and real yields.

<Insert Figure 13 here.>

<Insert Table 4 here.>

7 Conclusions

This paper uses Bayesian techniques to develop a tractable approach for incorporating prior information about the unconditional mean of yields in dynamic no-arbitrage term structure models. We build on the work of Villani (2009) who proposes a way to specify a prior about the unconditional mean in Bayesian VAR models, and Bauer (2015) who uses a similar Bayesian method for estimating affine term structure models. We rotate the term structure model pricing factors into bond yields and specify priors on the unconditional means of those yields. Parameters of the time-series dynamics are then drawn within a Gibbs sampling procedure.

We apply this technique to UK data in a benchmark affine term structure model of nominal bond yields, a shadow rate term structure model, and a jointly estimated affine term structure model of real and nominal bond yields. We find that with reasonable priors we obtain more plausible estimates of the long-run average of yields, lower estimates of term premia in long-term bonds and substantially reduced uncertainty around these decompositions.

References

Abrahams, M., T. Adrian, R. K. Crump, and E. Moench (2015). Decomposing real and nominal yield curves. *Federal Reserve Bank of New York Staff Reports*.

Adrian, T., R. K. Crump, and E. Moench (2013). Pricing the term structure with linear regressions. *Journal of Financial Economics 110*, 110–138.

Anderson, N. and J. Sleath (2001). New estimates of the UK real and nominal yield curves. *Bank of England Working Paper 126*.

Andreasen, M. M. and A. C. Meldrum (2013). Likelihood inference in non-linear term structure models: the importance of the zero lower bound. *Bank of England Working Paper 481*.

Andreasen, M. M. and A. C. Meldrum (2014). Dynamic term structure models: the best way to enforce the zero lower bound. *CREATES Research Paper 2014-47*.

Andreasen, M. M. and A. C. Meldrum (2015). Market beliefs about the UK monetary policy lift-off horizon: a no-arbitrage shadow rate term structure model approach. *Bank of England Staff Working Paper 541*.

Ang, A., J. Boivin, S. Dong, and R. Loo-Kung (2011). Monetary policy shifts and the term structure. *Review of Economic Studies 78*, 429–457.

Bauer, M. D. (2015). Restrictions on risk prices in dynamic term structure models. *Unpublished working paper*.

Bauer, M. D. and G. D. Rudebusch (2014). Monetary policy expectations at the zero lower bound. *Federal Reserve Bank of San Francisco Working Paper 2013-18*.

Bauer, M. D., G. D. Rudebusch, and J. C. Wu (2012). Correcting estimation bias in dynamic term structure models. *Journal of Business & Economic Statistics 30*, 454–467.

Black, F. (1995). Interest rates as options. *Journal of Finance 50*, 1371–1376.

Chen, R.-R. and L. Scott (1993). Maximum likelihood estimation for a multifactor equilibrium model of the term structure of interest rates. *Journal of Fixed Income 3*, 14–31.

Chib, S. and B. Ergashev (2009). Analysis of multi-factor affine yield curve models. *Journal of the American Statistical Association 104*, 1324–1337.

Christensen, J. and G. D. Rudebusch (2013). Modelling yields at the zero lower bound: Are shadow rates the solution? *Federal Reserve Bank of San Francisco Working Paper 2013-07*.

Cochrane, J. H. and M. Piazessi (2008). Decomposing the yield curve. *Unpublished working paper*.

Dai, Q. and K. J. Singleton (2000). Specification analysis of affine term structure models. *Journal of Finance 55*, 1943–1978.

D'Amico, S., D. H. Kim, and M. Wei (2014). Tips from TIPS: the informational content of Treasury Inflation-Protected Security prices. *Federal Reserve Board Finance and Economics Discussion Series 2014-24*.

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

Evans, M. D. D. (1998). Real rates, expected inflation and inflation risk premia. *Journal of Finance 53*, 187–218.

Guimarães, R. (2014). Expectations, risk premia and information spanning in dynamic term structure model estimation. *Bank of England Working Paper 489*.

Hamilton, J. D. and J. C. Wu (2012). Identification and estimation of affine term structure models. *Journal of Econometrics 168*, 315–331.

Jacquier, E., N. G. Polson, and P. E. Rossi (1994). Bayesian analysis of stochastic volatility models. *Journal of Business and Economic Statistics 12*, 413–417.

Jarocinski, M. and A. Marcet (2010). Autoregressions in small samples, priors about observables and initial conditions. *ECB Working Paper 1263*.

Joslin, S., K. J. Singleton, and H. Zhu (2011). A new perspective on Gaussian dynamic term structure models. *Review of Financial Studies 24*, 926–970.

Joyce, M. A. S., I. Kaminska, and P. Lildholdt (2012). Understanding the real rate conundrum: an application of no-arbitrage finance models to the UK real yield curve. *Review of Finance 16*, 837–866.

Joyce, M. A. S., P. Lildholdt, and S. Sorensen (2010). Extracting inflation expectations and inflation risk premia from the term structure: a joint model of the UK nominal and real yield curves. *Journal of Banking and Finance 34*, 281–294.

Kim, D. H. and A. Orphanides (2012). Term structure estimation with survey data on interest rate forecasts. *Journal of Financial and Quantitative Analysis 47*, 241–272.

Kim, D. H. and M. A. Priebsch (2013). Estimation of multi-factor shadow-rate term structure models. *Unpublished working paper*.

Malik, S. and A. C. Meldrum (2014). Evaluating the robustness of UK term structure decompositions using linear regression methods. *Bank of England Working Paper 518.*

Norgaard, M., K. Poulsen, and O. Ravn (2000). Advances in derivative-free state estimation for non-linear systems. *Automatica 36*, 1627–1638.

Priebsch, M. A. (2013). Computing arbitrage-free yields in multi-factor Gaussian shadow-rate term structure models. *Federal Reserve Board Finance and Economics Discussion Series 2013-63.*

Sims, C. (2000). Using a likelihood perspective to sharpen economic discourse: Three examples. *Journal of Econometrics 95*, 443–462.

Villani, M. (2009). Steady-state priors for vector autoregressions. *Journal of Applied Econometrics 24*, 630–650.

Appendix A: Tables and charts

Table 1: Posterior parameter estimates for benchmark affine model of nominal yields with long-run prior

Parameter	5^{th} percentile	Mean	95^{th} percentile
δ_0	0.004	0.005	0.006
θ_1	-0.002	-0.002	-0.002
θ_2	-0.022	-0.020	-0.018
θ_3	-0.045	-0.039	-0.033
$\sigma_{11} \times 10^3$	0.320	0.365	0.416
$\sigma_{21} \times 10^3$	-0.823	-0.655	-0.509
$\sigma_{22} \times 10^3$	0.530	0.635	0.764
$\sigma_{31} \times 10^3$	0.161	0.296	0.444
$\sigma_{32} \times 10^3$	-0.776	-0.639	-0.518
$\sigma_{33} \times 10^3$	0.189	0.216	0.245
$\gamma_1 \times 1200$	3.608	4.340	5.118
$\gamma_2 \times 1200$	4.414	5.128	5.866
$\gamma_3 \times 1200$	4.431	5.231	6.053
ϕ_{11}	0.938	0.966	0.993
ϕ_{12}	-0.021	-0.001	0.018
ϕ_{13}	-0.050	-0.021	0.007
ϕ_{21}	-0.028	0.036	0.101
ϕ_{22}	0.941	0.986	1.027
ϕ_{23}	-0.042	0.028	0.099
ϕ_{31}	-0.067	-0.015	0.036
ϕ_{32}	-0.032	0.003	0.039
ϕ_{33}	0.898	0.956	1.013

Table 2: Prior and posterior estimates of long-run mean parameters in benchmark affine model of nominal yields

Parameter	5^{th} percentile	Mean	95^{th} percentile
(a) Long-run prior			
$y_t^{(1)}$	3.67	4.50	5.32
$y_t^{(60)}$	3.84	5.00	6.16
$y_t^{(120)}$	3.86	5.50	7.14
(b) Model with long-run prior			
$y_t^{(1)}$	3.61	4.34	5.12
$y_t^{(60)}$	4.41	5.12	5.87
$y_t^{(120)}$	4.43	5.23	6.05
(c) Model with flat prior over γ			
$y_t^{(1)}$	-14.26	-4.17	5.65
$y_t^{(60)}$	-10.29	-1.63	6.39
$y_t^{(120)}$	-6.17	0.25	6.47

Estimates of the long-run means of yields under the long-run prior (panel (a)), in the model with a long-run prior (panel (b)) and in the model with a flat prior over γ. All numbers are annualised percentage points.

Table 3: Principal components analysis of nominal and real yields

Principal component	Nominal		Real	
	Proportion	Cumulative	Proportion	Cumulative
1	89.06%	89.06%	99.03%	99.03%
2	9.46%	98.52%	0.96%	99.99%
3	1.42%	99.94%	0.01%	100.00%
4	0.05%	99.99%	0.00%	100.00%
5	0.01%	100.00%	–	–

Table 4: Prior and posterior estimates of long-run mean parameters in joint affine model of real and nomimal yields

Parameter	5^{th} percentile	Mean	95^{th} percentile
(a) Long-run prior			
$y_t^{(1)}$	3.67	4.50	5.32
$y_t^{(60)}$	3.84	5.00	6.16
$y_t^{(120)}$	3.86	5.50	7.14
$y_{t,R}^{(1)}$	1.78	2.00	2.82
$y_{t,R}^{(120)}$	0.36	2.00	3.64
$\pi_t^{(1)}$	1.34	2.50	3.66
$\pi_t^{(120)}$	1.17	3.50	5.83
(b) Model with long-run prior			
$y_t^{(1)}$	3.89	4.48	5.07
$y_t^{(60)}$	4.59	5.24	5.92
$y_t^{(120)}$	4.55	5.31	6.08
$y_{t,R}^{(1)}$	1.69	2.32	2.99
$y_{t,R}^{(120)}$	1.65	2.13	2.61
$\pi_t^{(1)}$	1.61	2.16	2.67
$\pi_t^{(120)}$	2.72	3.18	3.61
(c) Model with flat prior over $\boldsymbol{\mu}$			
$y_t^{(1)}$	-9.21	-1.92	6.10
$y_t^{(60)}$	-7.92	-1.95	6.76
$y_t^{(120)}$	-4.88	-0.78	6.69
$y_{t,R}^{(1)}$	-13.60	-4.95	4.07
$y_{t,R}^{(120)}$	-6.61	-2.73	3.11
$\pi_t^{(1)}$	1.42	3.03	4.35
$\pi_t^{(120)}$	1.20	1.95	3.93

Estimates of the long-run means of yields under the long-run prior (panel (a)), in the model with a long-run prior (panel (b)) and in the model with a flat prior over γ. All numbers are annualised percentage points.

Figure 1: UK end-month zero-coupon bond yields, October 1992-December 2014

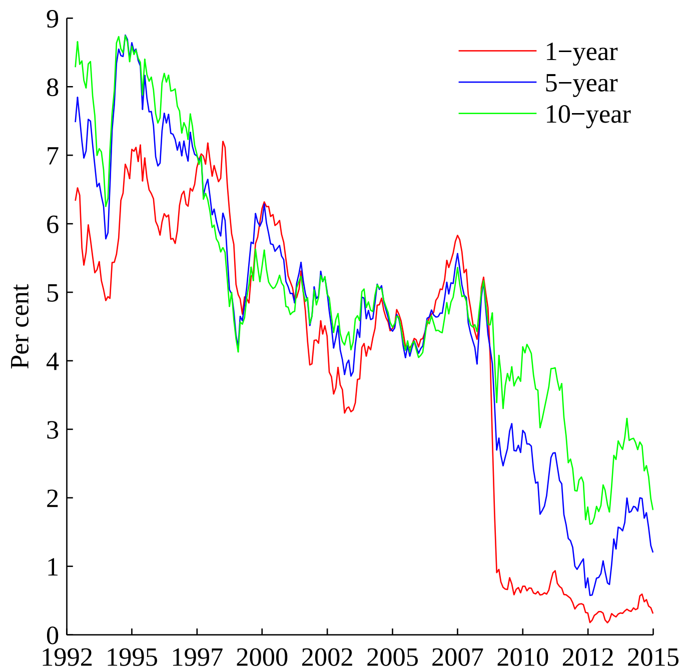

Figure 2: UK 10-year zero-coupon bond yield with AR(1) model projection from October 1992

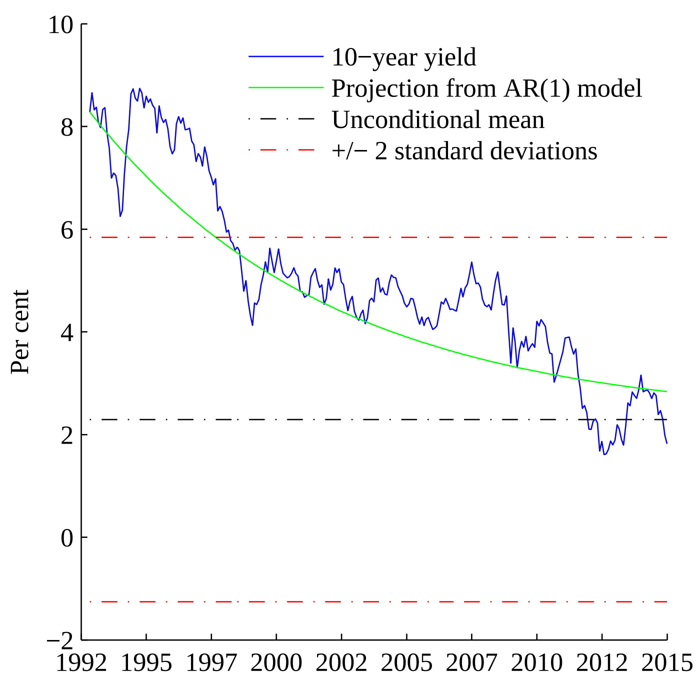

Figure 3: Average expected short-term interest rates over a 10-year horizon from the affine model with flat priors over the time-series dynamics

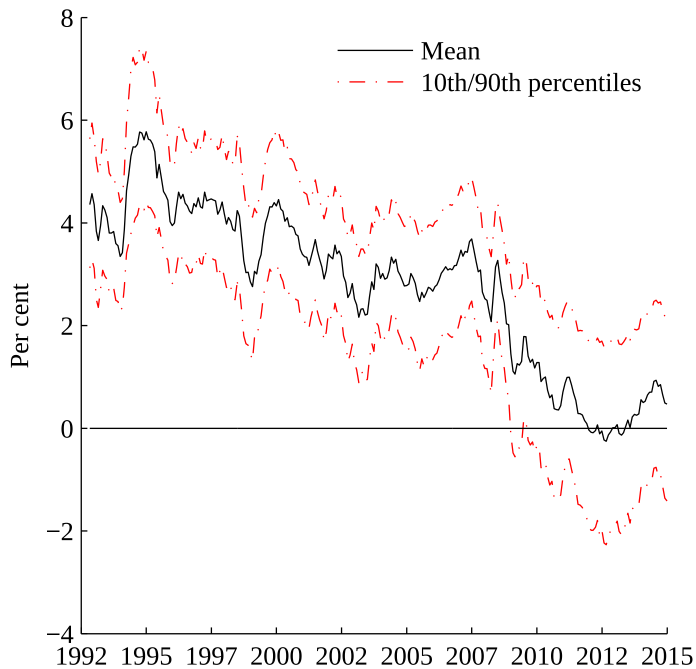

Figure 4: Term premium component of the 10-year yield implied by the affine model with flat priors over the time-series dynamics

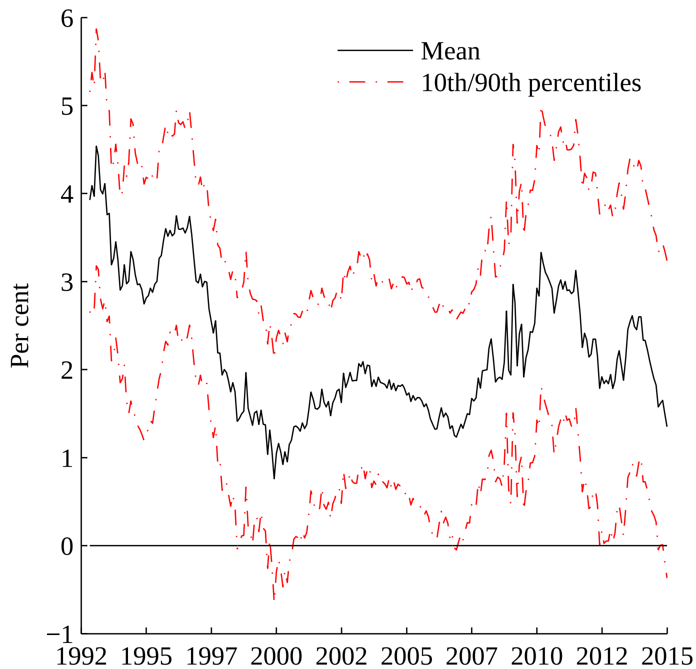

Figure 5: Width of the 80% probability interval for the 10-year term premium in the model with the long-run prior and a flat prior over γ

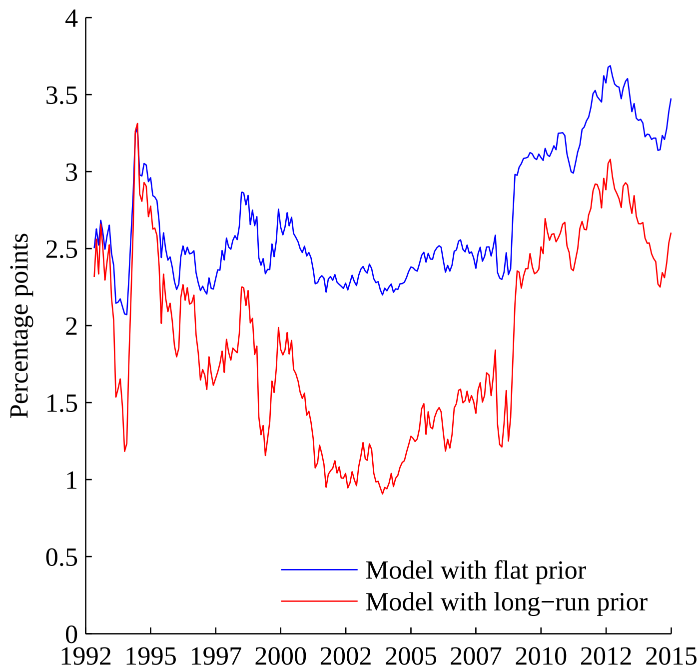

Figure 6: Average expected short-term interest rates over a 10-year horizon from the affine model with long-run prior

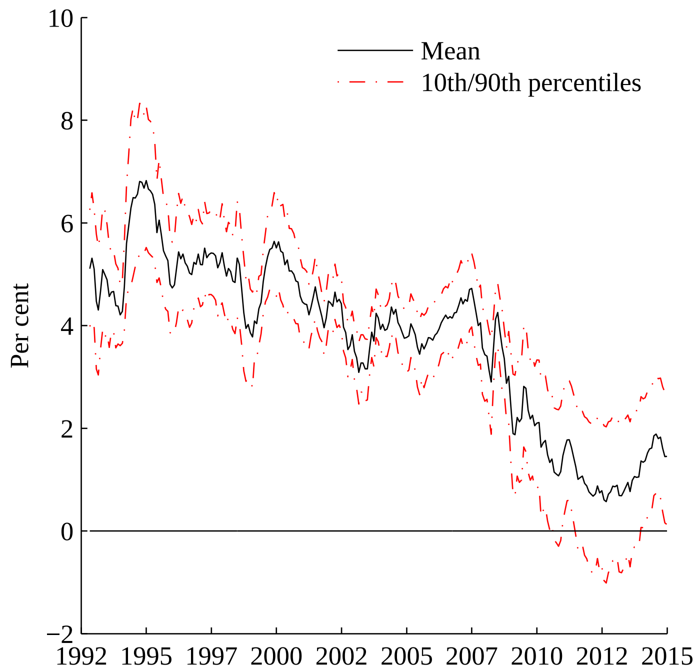

Figure 7: Term premium component of the 10-year yield implied by the affine model with long-run prior

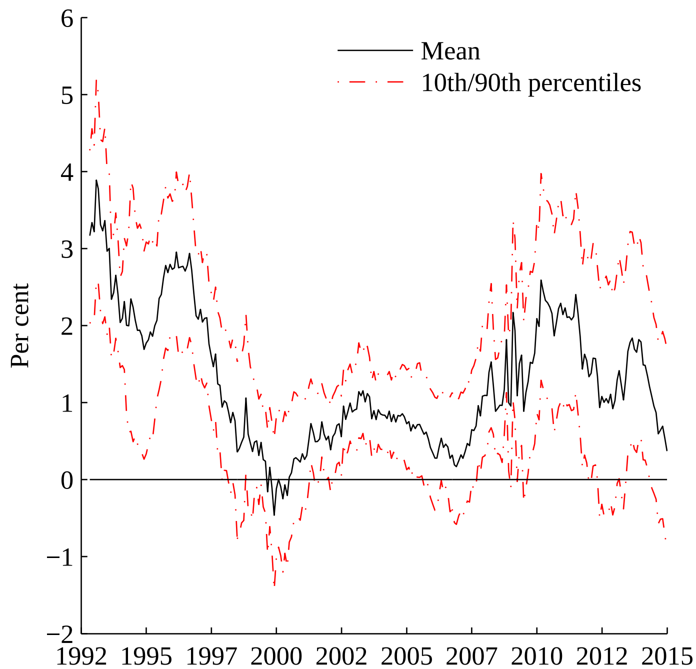

Figure 8: Term premium component of the 10-year yield implied by the shadow rate model with long-run prior

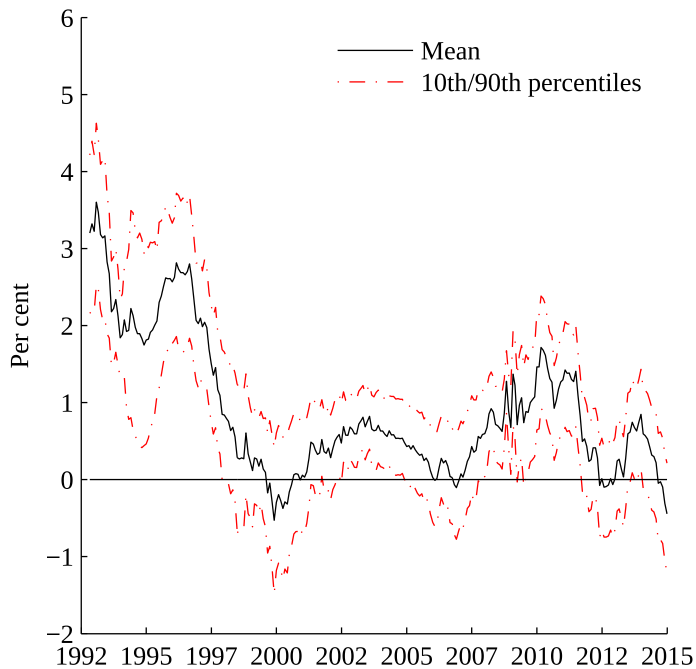

Figure 9: Real term premium component of the 10-year yield implied by the affine model with flat priors over the time-series dynamics

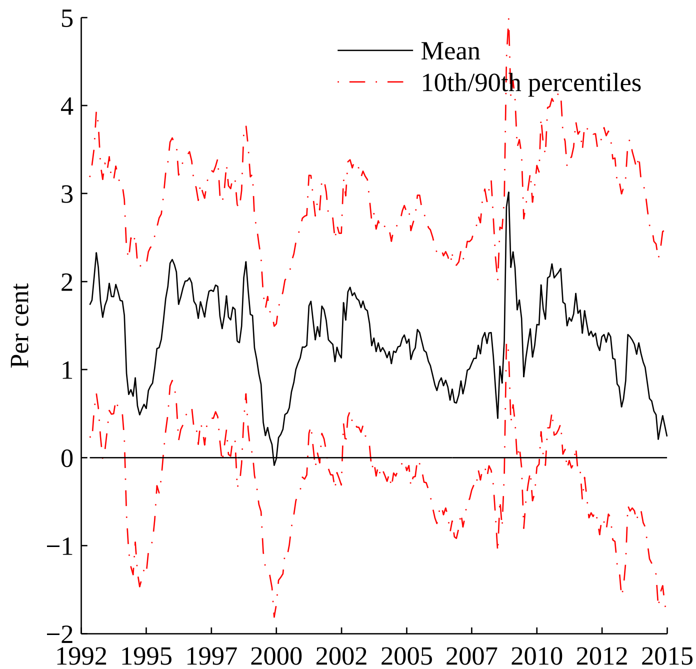

Figure 10: Real term premium component of the 10-year yield implied by the affine model with long-run prior

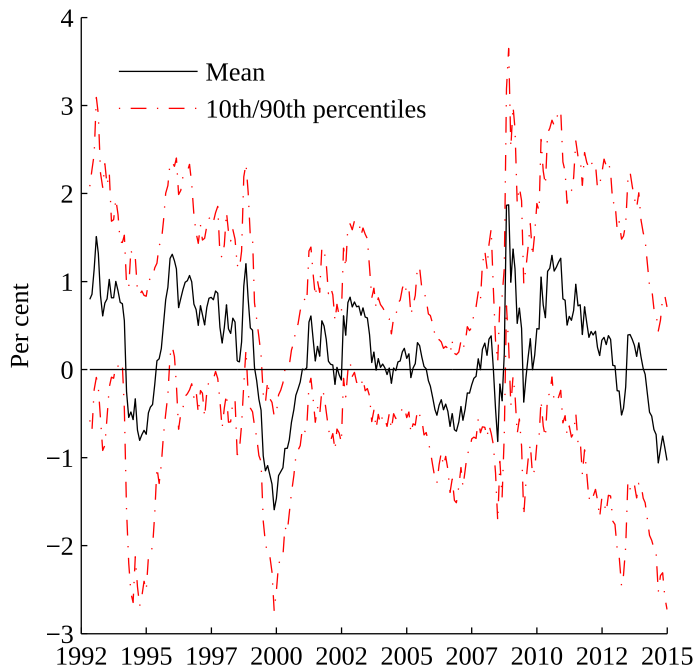

Figure 11: Inflation risk premium component of the 10-year yield implied by the affine model with flat priors over the time-series dynamics

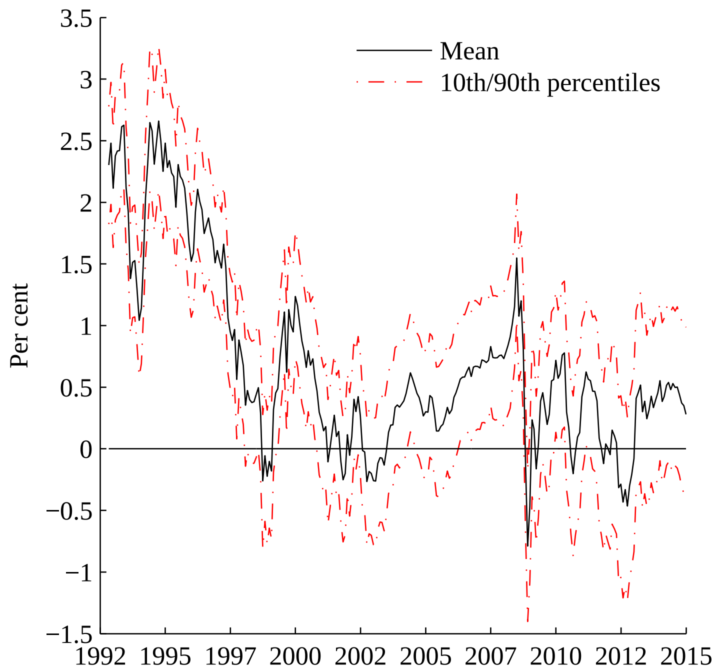

Figure 12: Inflation risk premium component of the 10-year yield implied by the affine model with long-run prior

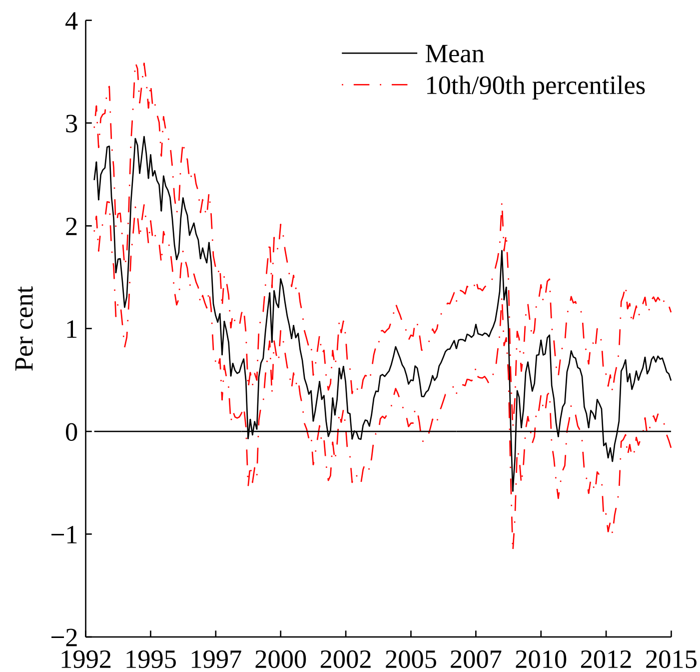

Figure 13: 10-year inflation breakeven