BANK OF ENGLAND

Staff Working Paper No. 541

Market beliefs about the UK monetary policy lift-off horizon: a no-arbitrage shadow rate term structure model approach

Martin M Andreasen and Andrew Meldrum

August 2015

BANK OF ENGLAND

Staff Working Paper No. 541

Market beliefs about the UK monetary policy lift-off horizon: a no-arbitrage shadow rate term structure model approach

Martin M Andreasen[1] and Andrew Meldrum[2]

Abstract

We use a no-arbitrage shadow rate term structure model to estimate investors' views about the timing of monetary policy 'lift-off' in the United Kingdom over time. Our estimates show that when the UK policy rate was first cut to 0.5%, in March 2009, investors believed that it would remain at the lower bound only for a short period, with an estimated probability of 70% that the policy rate would rise above 0.75% within twelve months. The estimated median horizon for policy rate lift-off rose sharply in 2012 but fell back to thirteen months by the end of our sample period, in May 2014.

Key words: Shadow rate models, sequential regression estimation, policy lift-off, zero lower bound.

JEL classification: C10, C50, G12.

---

(1) Aarhus University and CREATES. Email: mandreasen@econ.au.dk.
(2) Bank of England. Email: andrew.meldrum@bankofengland.co.uk.

The views expressed in this paper are those of the authors, and not necessarily those of the Bank of England or its committees. The authors are grateful for helpful discussions with Iryna Kaminska and Chris Young and for comments from an anonymous referee for the Bank of England Staff Working Paper series. M Andreasen greatly acknowledges financial support from the Danish Center for Scientific Computing (DCSC). He also appreciates financial support from CREATES — Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation.

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

As in several other advanced economies, UK short-term nominal interest rates have been at historic lows in recent years. In March 2009, the UK's Monetary Policy Committee (MPC) lowered its policy interest rate ('Bank Rate') to 0.5% and has held Bank Rate unchanged since then. Yields on longer maturity government bonds have also fallen to historic lows. For example, the two-year yield fell as low as 0.07% and the 10-year yield to 1.61% in July 2012.

A natural question to ask when the policy rates rate is close to its lower bound is when investors believe that it will start to rise again, commonly referred to as policy 'lift-off'. As has been discussed in a number of recent studies, however, the low level of nominal interest rates significantly complicates the problem of estimating the conditional distributions of future short-term policy rates. Perhaps the most popular class of dynamic term structure models, the multi-factor Gaussian no-arbitrage affine term structure model (ATSM) of Duffie and Kan (1996), does not impose a lower bound on nominal interest rates. These models can therefore imply negative nominal bond yields and a substantial probability of future negative nominal rates when yields are low (e.g. Andreasen and Meldrum (2013); Bauer and Rudebusch (2014)).

There are, however, a number of alternative frameworks that do impose the lower bound within a no-arbitrage dynamic term structure model (DTSM).[1] In recent years, the most widely used framework has been the shadow rate model proposed by Black (1995). Recent examples include Krippner (2012), Priebsch (2013), Christensen and Rudebusch (2013), Andreasen and Meldrum (2014), Bauer and Rudebusch (2014) and Lemke and Vladu (2014). The shadow rate framework is attractive because yields remain approximately affine in the pricing factors when they are far from the lower bound, but are truncated below by a lower bound, at which they can remain for extended periods.

This paper applies the shadow rate framework to UK bond yields and studies how model-implied estimates of policy lift-off in the UK have evolved over time. We use a technique

[1] The performance of some of these models has been compared using Japanese data by Kim and Singleton (2012) and Christensen and D. (2015) and using US data by Christensen and Rudebusch (2013) and Andreasen and Meldrum (2014).

similar to that used by Bauer and Rudebusch (2014) and Lemke and Vladu (2014) to study expectations of policy lift-off in the US and the euro area respectively. We find evidence that in March 2009, when Bank Rate was first lowered to 0.5%, investors did not initially expect it to remain at that level for a long period: the estimated probability that Bank Rate would rise 0.75% within 12 months was around 70% and the median lift-off horizon (the number of months before Bank Rate reaches 0.75%) was 7 months. The median lift-off horizon remained fairly constant for the next three years, before rising markedly in 2012, reaching more than 40 months in the middle of 2012. At this time, there was a significant implied probability of a further *reduction* in the policy rate and the median path for the short rate lay materially below the mean expectation. The median lift-off horizon subsequently fell back to 13 months by the end of our sample period, in May 2014.

Our paper is the first to estimate shadow rate models using UK data.[2] In many important respects, our benchmark four-factor shadow rate model performs very similarly to a standard Gaussian ATSM when measured in terms of in-sample fit and the models' ability to match the standard specification tests proposed by Dai and Singleton (2002). Estimates of the term premia in long-term bond yields from the two models are almost identical, which is consistent with previous findings by Kim and Priebsch (2013) for the US. But while conditional expectations of short-term interest rates are similar from the two models, the ATSM implies a substantial probability of negative nominal interest rates since early 2009 (i.e. when Bank Rate was lowered to 0.5%), making it inappropriate for analysing the conditional distribution of future short-term interest rates and investors' views about the timing of policy lift-off.

The remainder of this paper proceeds as follows. Section 2 outlines the standard multi-factor Gaussian ATSM and the shadow rate extension. Section 3 describes how we apply the SR approach of Andreasen and Christensen (2015) to estimate the models. Section 4 discusses our data set and issues of model specification, including the appropriate number of pricing factors and the level of the lower bound in the shadow rate model. Section 5 compares the results from a four-factor shadow rate model with a benchmark ATSM.

[2] Malik and Meldrum (2014) report estimates of term premia obtained using a shadow rate model but where the coefficient estimates come from an affine model estimated before the recent period of low interest rates.

Section 6 considers the implications of the shadow rate model for the path of policy rates since March 2009, including the estimated lift-off dates. Section 7 concludes.

2 Model

2.1 Gaussian ATSM

We start by setting out the key equations of a standard discrete-time Gaussian ATSM. The first equation specifies the one-period risk-free interest rate r_t to be affine in n_x pricing factors $\mathbf{x}_t$, i.e.

$$r_t = \alpha + \boldsymbol{\beta}' \mathbf{x}_t, \tag{1}$$

where α is a scalar and $\boldsymbol{\beta}$ is an $n_x \times 1$ vector. This specification is typically motivated by referring to a Taylor rule, where the policy rate is determined by a desire to stabilize the inflation and output gap (see Ang and Piazzesi (2003), Hordahl et al. (2008) and Rudebusch and Wu (2008), among others). The second equation describes the dynamics of the pricing factors under the risk-neutral measure $\mathbb{Q}$ as a vector autoregressive (VAR) process, i.e.

$$\mathbf{x}_{t+1} = \boldsymbol{\Phi\mu} + (\mathbf{I} - \boldsymbol{\Phi})\,\mathbf{x}_t + \boldsymbol{\Sigma\varepsilon}^{\mathbb{Q}}_{t+1}, \tag{2}$$

where $\boldsymbol{\varepsilon}^{\mathbb{Q}}_{t+1} \sim \mathcal{NID}(\mathbf{0}, \mathbf{I})$. The mean level of the pricing factors is controlled by $\boldsymbol{\mu}$ of dimension $n_x \times 1$, while the persistence and the conditional volatility of the factors are determined by the $n_x \times n_x$ matrices $\boldsymbol{\Phi}$ and $\boldsymbol{\Sigma}$, respectively. In the absence of arbitrage, the price at time t of an k-period zero-coupon bond is $P_{t,k} = E^{\mathbb{Q}}_t \left[\exp\{-r_t\} P_{t+1,k-1}\right]$. Given the assumptions in (1) and (2), bond prices are exponentially affine in the factors, i.e.

$$P_{t,k} = \exp\left\{A_k + \mathbf{B}'_k \mathbf{x}_t\right\} \tag{3}$$

for $k = 1, 2, ..., K$, where the recursive formulae for A_k and $\mathbf{B}_k$ are easily derived.

The final equation specifies the functional form for the market prices of risk $\mathbf{f}(\mathbf{x}_t)$ with dimension $n_x \times 1$. The relationship between the physical measure $\mathbb{P}$ and the $\mathbb{Q}$ measure is given by $\boldsymbol{\varepsilon}_{t+1}^{\mathbb{Q}} = \boldsymbol{\varepsilon}_{t+1}^{\mathbb{P}} + \mathbf{f}(\mathbf{x}_t)$, and the factor dynamics under $\mathbb{P}$ are therefore

$$\mathbf{x}_{t+1} = \boldsymbol{\Phi}\boldsymbol{\mu} + (\mathbf{I} - \boldsymbol{\Phi})\,\mathbf{x}_t + \boldsymbol{\Sigma}\mathbf{f}(\mathbf{x}_t) + \boldsymbol{\Sigma}\boldsymbol{\varepsilon}_{t+1}^{\mathbb{P}}$$

with $\boldsymbol{\varepsilon}_{t+1}^{\mathbb{P}} \sim \mathcal{NID}(\mathbf{0}, \mathbf{I})$. Following Duffee (2002), to obtain an affine process for the pricing factors under $\mathbb{P}$, we let $\mathbf{f}(\mathbf{x}_t) = \boldsymbol{\Sigma}^{-1}(\mathbf{f}_0 + \mathbf{f}_1\mathbf{x}_t)$, where $\mathbf{f}_0$ has dimension $n_x \times 1$ and $\mathbf{f}_1$ is an $n_x \times n_x$ matrix. This implies the following $\mathbb{P}$ dynamics:

$$\mathbf{x}_{t+1} = \boldsymbol{\Phi}\boldsymbol{\mu} + \mathbf{f}_0 + (\mathbf{I} - \boldsymbol{\Phi} + \mathbf{f}_1)\,\mathbf{x}_t + \boldsymbol{\Sigma}\boldsymbol{\varepsilon}_{t+1}^{\mathbb{P}}. \tag{4}$$

To obtain stationary bond yields with finite first and second unconditional moments, we require the process for $\mathbf{x}_t$ to be stationary, i.e. that all eigenvalues of $\mathbf{I} - \boldsymbol{\Phi} + \mathbf{f}_1$ are inside the unit circle.

The pricing factors are considered to be latent (i.e. unobserved) and a set of normalization restrictions are therefore needed to identify the model. We require i) $\boldsymbol{\beta} = \mathbf{1}$, ii) $\boldsymbol{\mu} = \mathbf{0}$, iii) $\boldsymbol{\Phi}$ to be diagonal, and iv) $\boldsymbol{\Sigma}$ to be triangular.[3] This identification scheme constrains the $\mathbb{Q}$ dynamics for the pricing factors, whereas the $\mathbb{P}$ dynamics are unrestricted. The latter is convenient when the model is estimated by the SR approach, as explained in Section 3.1.2.

2.2 The shadow rate model

In the shadow rate model suggested by Black (1995), the lower bound is enforced by introducing a shadow interest rate $s(\mathbf{x}_t)$ that is unconstrained by the ZLB and may therefore attain negative values.[4] In the absence of any transaction and storage costs for money, the nominal interest rate cannot be negative because investors can always decide to hold cash. In other words, the nominal interest rate has an option element. This argument motivates the specification

$$r_t = \max(0, s(\mathbf{x}_t)), \tag{5}$$

[3] Other normalization schemes exist, for instance the one recently suggested by Joslin et al. (2011).
[4] The idea of considering a shadow rate is also briefly mentioned in Rogers (1995).

where the policy rate r_t is the non-negative part of the shadow rate. As is standard in recent studies that apply shadow rate models to US data, we let the shadow rate be affine in the pricing factors, i.e.[5]

$$s(\mathbf{x}_t) = \alpha + \boldsymbol{\beta}'\mathbf{x}_t, \tag{6}$$

and continue to assume that the pricing factors follow a Gaussian VAR(1) under both probability measures - i.e. we impose (2) and (4). The identification conditions for the shadow rate model are identical to those for the Gaussian ATSM.

In the shadow rate model, there is no exact solution for long-term bond prices. A number of methods have been suggested to approximate long-term bond prices in these models, including: (i) lattices (Ichiue and Ueno (2007)); (ii) finite-difference methods (Kim and Singleton (2012)); (iii) Monte Carlo integration (Bauer and Rudebusch (2014)); (iv) an option pricing approximation (Krippner (2012) and Christensen and Rudebusch (2013)); and (v) ignoring the Jensen's inequality term to solve a Gaussian model by a truncated normal distribution (Ichiue and Ueno (2013)). In this paper, we use a discrete time version of the method proposed by Priebsch (2013). Note first that k-period bond yields can be written as:

$$y_{t,k} = -\frac{1}{k}\log E_t^{\mathbb{Q}}\left[\prod_{i=0}^{k-1}\exp\left(-r_{t+i}\right)\right]$$

Priebsch (2013) proposes to take a second-order approximation, giving:

$$\begin{aligned} y_{t,k} &\simeq \frac{1}{k}E_t^{\mathbb{Q}}\left[\sum_{i=0}^{k-1} r_{t+i}\right] - \frac{1}{2k}Var_t^{\mathbb{Q}}\left[\sum_{i=0}^{k-1} r_{t+i}\right] \\ &= \frac{1}{k}E_t^{\mathbb{Q}}\left[\sum_{i=0}^{k-1}\max\left\{0, s_{t+i}\right\}\right] - \frac{1}{2k}\left\{\begin{array}{c} E_t^{\mathbb{Q}}\left[\left(\sum_{i=0}^{k-1}\max\left\{0, s_{t+i}\right\}\right)^2\right] \\ -E_t^{\mathbb{Q}}\left[\left(\sum_{i=0}^{k-1}\max\left\{0, s_{t+i}\right\}\right)\right]^2 \end{array}\right\} \end{aligned} \tag{7}$$

Using the results reported by Priebsch (2013) for the truncated Normal distribution, the

[5] Kim and Singleton (2012) and Andreasen and Meldrum (2014) consider models with a quadratic specification for the shadow rate.

expectation of the short rate at period $t+i$ are given by:

$$E_t^{\mathbb{Q}}\left[\max\left\{0, s_{t+i}\right\}\right] = \mu_{t,t+i}\Phi\left(\frac{\mu_{t,t+i}}{\sigma_{t,t+i}}\right) + \sigma_{t,t+i}\phi\left(\frac{\mu_{t,t+i}}{\sigma_{t,t+i}}\right), \tag{8}$$

where $\mu_{t,t+i} = E_t^{\mathbb{Q}}\left[s_{t+i}\right]$ and $\sigma^2_{t,t+i} = Var_t^{\mathbb{Q}}\left[s_{t+i}\right]$, both of which are straightforward to compute given (2) and (6). Here, $\phi\left(.\right)$ is the probability density function of the standard Normal distribution; and $\Phi\left(.\right)$ is the cumulative density function of the standard Normal distribution. The expectation of the squared future short rate is given by:

$$\begin{aligned} E_t\left[\max\left\{0, s_{t+i}\right\}^2\right] \quad = \quad & \left(\mu_{t,t+i}\mu_{t,t+j} + \sigma_{t,t+i,t+j}\right)\Phi_2^d\left(-\zeta_{t,t+i}, -\zeta_{t,t+j}; \chi_{t,t+i,t+j}\right) \\ & +\sigma_{t,t,+j}\mu_{t,t+i}\phi\left(\zeta_{t,t+j}\right)\Phi\left(\frac{\zeta_{t,t+i} - \chi_{t,t+i,t+j}\zeta_{t,t+j}}{\sqrt{1-\chi^2_{t,t+i,t+j}}}\right) \\ & +\sigma_{t,t,+i}\mu_{t,t+j}\phi\left(\zeta_1\right)\Phi\left(\frac{\zeta_{t,t+j} - \chi_{t,t+i,t+j}\zeta_{t,t+i}}{\sqrt{1-\chi^2_{t,t+i,t+j}}}\right) \\ & +\sigma_{t,t,+i}\sigma_{t,t,+j}\sqrt{\frac{1-\chi^2_{t,t+i,t+j}}{2\pi}} \\ & \times\phi\left(\sqrt{\frac{\zeta^2_{t,t+i} - 2\chi_{t,t+i,t+j}\zeta_{t,t+i}\zeta_{t,t+j} + \zeta^2_{t,t+j}}{1-\chi^2_{t,t+i,t+j}}}\right) \end{aligned} \tag{9}$$

where $\zeta_{t,t+i} = \frac{\mu_{t,t+i}}{\sigma_{t,t+i}}$; $\chi_{t,t+i,t+j} = \frac{\sigma_{t,t+i,t+j}}{\sigma_{t,t+i}\sigma_{t,t+j}}$; and $\Phi_2^d\left(z_1, z_2; \chi\right) = 1 - \Phi\left(z_1\right) - \Phi\left(z_2\right) + \Phi_2\left(z_1, z_2; \chi\right)$. Substituting (8) and (9) into (7) therefore provides a second-order approximation to long-term yields.

3 The estimation procedure

One way to estimate non-linear DTSMs with latent pricing factors, as in the shadow rate model, is to approximate the unknown likelihood function by sequential Monte Carlo methods (see Doucet et al. (2001) and De Rossi (2004)). This procedure is very time consuming for multi-factor DTSMs. A computationally more feasible alternative is to use a non-linear extension of the Kalman filter and a quasi-maximum likelihood (QML) approach, but its asymptotic properties are generally unknown. We overcome these difficulties by using the sequential SR approach by Andreasen and Christensen (2015), which has known asymptotic properties and is faster to implement than the QML approach. We also emphasize that the asymptotic properties of the SR approach hold under weaker restrictions than typically considered for likelihood-based inference. In this section we present the SR approach and describe how the latent pricing factors and model parameters are estimated in the models considered.

3.1 The SR approach

The SR approach may be applied to DTSMs where bond yields are potentially non-linear functions of latent pricing factors and measured with errors $v_{t,k}$, i.e.

$$y_{t,k} = g_k\left(\mathbf{x}_t; \boldsymbol{\theta}_1\right) + v_{t,k}, \tag{10}$$

where the subscript k index the maturity of the bond yields. The functional relationship between the pricing factors and bond yields is parameterized by $\boldsymbol{\theta}_1 \equiv \begin{bmatrix} \boldsymbol{\theta}_{11}' & \boldsymbol{\theta}_{12}' \end{bmatrix}'$ containing the risk-neutral paramters. Elements in $\boldsymbol{\theta}_{11}$ may only be determined from the measurement equations in (10), whereas $\boldsymbol{\theta}_{12}$ may be obtained from (10) and the factor dynamics under the $\mathbb{P}$ measure. For the Gaussian ATSM, the g-function is linear in the pricing factors, i.e. $g_k^{ATSM}\left(\mathbf{x}_t; \boldsymbol{\theta}_1^{ATSM}\right) \equiv -\frac{1}{k}\left(A_k + \mathbf{B}_k'\mathbf{x}_t\right)$, and we have $\boldsymbol{\theta}_{11}^{ATSM} \equiv \begin{bmatrix} \alpha & diag(\boldsymbol{\Phi})' \end{bmatrix}'$ with $\boldsymbol{\theta}_{12}^{ATSM} \equiv \begin{bmatrix} vech\left(\boldsymbol{\Sigma}\right)' \end{bmatrix}'$. In the shadow rate model, $g_k^{SH}\left(\mathbf{x}_t; \boldsymbol{\theta}_1^{SH}\right)$ is an unknown non-linear mapping from the pricing factors to bond yields with $\boldsymbol{\theta}_1^{SH} = \boldsymbol{\theta}_1^{ATSM}$. It is important to stress that the SR approach does not impose any distributional assumptions on the measurement errors $v_{t,k}$, which furthermore may display heteroskedasticity and correlation in both the cross-section and the time series dimensions.

The SR approach allows the pricing factors under the $\mathbb{P}$ measure to evolve according to a general Markov process of the form

$$\mathbf{x}_{t+1} = \mathbf{h}\left(\mathbf{x}_t, \boldsymbol{\epsilon}_{t+1}^{\mathbb{P}}; \boldsymbol{\theta}_{11}, \boldsymbol{\theta}_2\right). \tag{11}$$

The $\mathbf{h}$-function may depend on $\boldsymbol{\theta}_{11}$ and $\boldsymbol{\theta}_2 \equiv \begin{bmatrix} \boldsymbol{\theta}_{22}' & \boldsymbol{\theta}_{12}' \end{bmatrix}'$, where $\boldsymbol{\theta}_{22}$ must be determined from the factor dynamics in (11). Both the ATSM and shadow rate model have a linear and unrestricted transition function which we represent by

$$\mathbf{x}_{t+1} = \mathbf{h}_0 + \mathbf{h}_{\mathbf{x}}\mathbf{x}_t + \boldsymbol{\Sigma}\boldsymbol{\varepsilon}^{\mathbb{P}}_{t+1}, \tag{12}$$

where $\mathbf{h}_0 \equiv \boldsymbol{\Phi}\boldsymbol{\mu} + \mathbf{f_0}$, $\mathbf{h}_{\mathbf{x}} \equiv \mathbf{I} - \boldsymbol{\Phi} + \mathbf{f}_1$, and $\boldsymbol{\varepsilon}^{\mathbb{P}}_{t+1} \sim \mathcal{NID}\left(\mathbf{0}, \mathbf{I}\right)$. Hence, given the parametrization of the $\mathbf{h}$-function in (12), we have $\boldsymbol{\theta}_{22} \equiv \begin{bmatrix} \mathbf{h}_0' & vec\left(\mathbf{h}_{\mathbf{x}}\right)' \end{bmatrix}'$ for the models considered.

The subsequent sections describe how the latent pricing factors $\{\mathbf{x}_t\}_{t=1}^{T}$ and the model parameters $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ are estimated in the SR approach using a three-step procedure.

3.1.1 The SR approach: Step 1

The latent pricing factors are estimated by running the cross-section regressions

$$\hat{\mathbf{x}}_t\left(\boldsymbol{\theta}_1\right) = \arg\min_{\mathbf{x}_t \in \mathcal{X}_t} Q_t = \frac{1}{2n_{y,t}} \sum_{j=1}^{n_{y,t}} \left(y_{t,j} - g_j\left(\mathbf{x}_t; \boldsymbol{\theta}_1\right)\right)^2 \tag{13}$$

for $t = 1, 2, ..., T$, where $n_{y,t}$ refers to the number of bond yields in time period t. The estimated factors are denoted $\{\hat{\mathbf{x}}_{2,t}\left(\boldsymbol{\theta}_1\right)\}_{t=1}^{T}$ because they are computed for a given $\boldsymbol{\theta}_1$. These regressions have a closed-form solution for the Gaussian ATSM with g_j^{ATSM} being linear in the pricing factors. For the shadow rate model, the regressions in (13) are non-linear and solved using the Levenberg-Marquardt method with the pricing factors from the previous time period $\hat{\mathbf{x}}_{2,t-1}\left(\boldsymbol{\theta}_1\right)$ serving as ideal starting values for $t = 2, 3, ..., T$.[6]

The model parameters $\boldsymbol{\theta}_1$ are obtained by pooling all squared residuals from (13) and minimizing their sum with respect to $\boldsymbol{\theta}_1$, i.e.

$$\hat{\boldsymbol{\theta}}_1^{step1} = \arg\min_{\boldsymbol{\theta}_1 \in \Theta_1} Q_{1:T}^{step1} = \frac{1}{2N} \sum_{t=1}^{T} \sum_{j=1}^{n_{y,t}} \left(y_{t,j} - g_j\left(\hat{\mathbf{x}}_t\left(\boldsymbol{\theta}_1\right); \boldsymbol{\theta}_1\right)\right)^2, \tag{14}$$

[6] The main input for Levenberg-Marquardt optimizer is the Jacobian $\partial \mathbf{g}\left(\mathbf{x}_t; \boldsymbol{\theta}_1\right) / \partial \mathbf{x}_t'$. For the shadow rate model, the Jacobian is obtained by numerical differentiation using a first-order approximation as in Ichiue and Ueno (2013) but otherwise the second-order approximation by Priebsch (2013) is applied in the optimizer. Using the second-order approximation to also compute the Jacobian in the optimizer gives identical results but is somewhat slower than the adopted procedure.

where $N \equiv \sum_{t=1}^{T} n_{y,t}$. Given standard regularity conditions, Andreasen and Christensen (2015) show consistency and asymptotic normality of $\hat{\boldsymbol{\theta}}_1^{step1}$, i.e.

$$\sqrt{N}\left(\hat{\boldsymbol{\theta}}_1^{step1} - \boldsymbol{\theta}_1^{o}\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \left(\mathbf{A}_o^{\boldsymbol{\theta}_1}\right)^{-1} \mathbf{B}_o^{\boldsymbol{\theta}_1} \left(\mathbf{A}_o^{\boldsymbol{\theta}_1}\right)^{-1}\right), \tag{15}$$

where the superscript "o" denotes the true value. These asymptotic properties are derived by letting the number of bond yields in each time period $n_{y,t}$ tend to infinity, i.e. $N \rightarrow \infty$. The expected value of the average Hessian matrix $\mathbf{A}_o^{\boldsymbol{\theta}_1}$ may be estimated consistently by

$$\hat{\mathbf{A}}^{\boldsymbol{\theta}_1} = \frac{1}{N}\sum_{t=1}^{T}\sum_{j=1}^{n_{y,t}}\left(\hat{\boldsymbol{\Psi}}_{t,j}^{\boldsymbol{\theta}_1}\right)\left(\hat{\boldsymbol{\Psi}}_{t,j}^{\boldsymbol{\theta}_1}\right)', \tag{16}$$

where

$$\boldsymbol{\Psi}_{t,j}^{\boldsymbol{\theta}_1}\left(\boldsymbol{\theta}_1\right) \equiv \frac{\partial \hat{\mathbf{x}}_{2,t}'\left(\boldsymbol{\theta}_1\right)}{\partial \boldsymbol{\theta}_1}\frac{\partial g_j\left(\hat{\mathbf{x}}_{2,t}\left(\boldsymbol{\theta}_1\right);\boldsymbol{\theta}_1\right)}{\partial \mathbf{x}_{2,t}\left(\boldsymbol{\theta}_1\right)} + \frac{\partial g_j\left(\hat{\mathbf{x}}_{2,t}\left(\boldsymbol{\theta}_1\right);\boldsymbol{\theta}_1\right)}{\partial \boldsymbol{\theta}_1} \tag{17}$$

and $\hat{\boldsymbol{\Psi}}_{t,j}^{\boldsymbol{\theta}_1} \equiv \boldsymbol{\Psi}_{t,j}^{\boldsymbol{\theta}_1}\left(\hat{\boldsymbol{\theta}}_1^{step1}\right)$. The average of the score function $\mathbf{B}_o^{\boldsymbol{\theta}_1}$ is estimated using an extension of the Newey-West estimator that is robust to heteroskedasticity in the time dimension, cross-section correlation, and autocorrelation in $v_{t,k}$. That is

$$\hat{\mathbf{B}}^{\boldsymbol{\theta}_1} = \frac{1}{N}\sum_{t=1}^{T}\sum_{j=1}^{n_{y,t}}\hat{\sigma}_t^2\left(\hat{\boldsymbol{\Psi}}_{t,j}^{\boldsymbol{\theta}_1}\right)\left(\hat{\boldsymbol{\Psi}}_{t,j}^{\boldsymbol{\theta}_1}\right)' \tag{18}$$

$$+\sum_{\substack{k_T=-w_T\\k_T\neq 0}}^{w_T}\sum_{\substack{k_D=-w_D\\k_D\neq 0}}^{w_D}\left(1-\frac{|k_T|}{1+w_T}\right)\left(1-\frac{|k_D|}{1+w_D}\right)\left(\hat{\boldsymbol{\Psi}}_{t,j}^{\boldsymbol{\theta}_1}\right)\left(\hat{\boldsymbol{\Psi}}_{t+k_T,j+k_D}^{\boldsymbol{\theta}_1}\right)'\hat{v}_{t,j}\hat{v}_{t+k_T,j+k_D} \tag{19}$$

where

$$\hat{\sigma}_t^2 = \frac{1}{n_{y,t}-n_{x_2}}\sum_{j=1}^{n_{y,t}}\hat{v}_{t,j}^2 \quad \text{for } t = 1, 2, ..., T$$

and $\hat{v}_{t,j} = y_{t,j} - g_j\left(\hat{\mathbf{x}}_t;\hat{\boldsymbol{\theta}}_1^{step1}\right)$. Here, w_D is the bandwidth for bond yields in the cross-section dimension when ordered by duration (i.e. maturity) and w_T is the corresponding bandwidth for the time series dimension. In this paper we set $w_D = 5$ and $w_T = 10$ throughout.

3.1.2 The SR approach: Step 2

We estimate $\boldsymbol{\theta}_2$ in (12) using $\{\hat{\mathbf{x}}_t\}_{t=1}^T$ and moment conditions accounting for the uncertainty $\{\mathbf{u}_t\}_{t=1}^T$ in the estimated pricing factors, i.e. $\hat{\mathbf{x}}_t = \mathbf{x}_t^o + \mathbf{u}_t$, where $\mathbf{x}_t^o$ denotes the true factor value. As in Andreasen and Christensen (2015), we modify the standard moment conditions for VAR models to account for uncertainty in $\{\hat{\mathbf{x}}_t\}_{t=1}^T$ and consider

$$\mathbf{q}_T(\boldsymbol{\theta}_2) \equiv \frac{1}{T-1}\sum_{t=1}^{T-1} \mathbf{q}_t(\boldsymbol{\theta}_2) = \mathbf{0}, \tag{20}$$

where

$$\mathbf{q}_t(\boldsymbol{\theta}_2) \equiv \begin{bmatrix} \hat{\mathbf{w}}_{t+1} \\ vec\left(\hat{\mathbf{w}}_{t+1}\hat{\mathbf{x}}_t' - Cov\left(\mathbf{u}_{t+1}, \mathbf{u}_t\right) + \mathbf{h}_{\mathbf{x}} Var\left(\mathbf{u}_t\right)\right) \\ vech\begin{pmatrix} \hat{\mathbf{w}}_{t+1}\hat{\mathbf{w}}_{t+1}' - Var\left(\hat{\mathbf{w}}_{t+1}\right) - Var\left(\mathbf{u}_t\right) - \mathbf{h}_{\mathbf{x}} Var\left(\mathbf{u}_t\right)\mathbf{h}_{\mathbf{x}}' \\ +Cov\left(\mathbf{u}_{t+1}, \mathbf{u}_t\right)\mathbf{h}_{\mathbf{x}}' + \mathbf{h}_{\mathbf{x}} Cov\left(\mathbf{u}_t, \mathbf{u}_{t+1}\right) \end{pmatrix} \end{bmatrix}$$

and

$$\hat{\mathbf{w}}_{t+1} \equiv \boldsymbol{\Sigma}\hat{\boldsymbol{\varepsilon}}_{t+1}^{\mathbb{P}} \equiv \hat{\mathbf{x}}_{t+1} - \mathbf{h}_0 - \mathbf{h}_{\mathbf{x}}\hat{\mathbf{x}}_t.$$

Note that $\hat{\boldsymbol{\varepsilon}}_{t+1}^{\mathbb{P}}$ refers to the residuals using the true values of $\mathbf{h}_0$ and $\mathbf{h}_{\mathbf{x}}$ but the estimated pricing factors $\hat{\mathbf{x}}_t$. Consistent estimators of $Var(\mathbf{u}_t)$, $Cov(\mathbf{u}_{t+1}, \mathbf{u}_t)$, and $Cov(\mathbf{u}_t, \mathbf{u}_{t+1})$ are provided in Andreasen and Christensen (2015) using output from the first estimation step, and $\boldsymbol{\theta}_2$ can therefore be estimated consistently by generalized methods of moments when the number of time periods T tends to infinity. All models considered in the present paper have unrestricted $\mathbb{P}$ dynamics, and the moment conditions in (20) may then be solved in closed form. The solution is obtained by correcting all second moments for estimation uncertainty in $\{\hat{\mathbf{x}}_t\}_{t=1}^T$ and running the regression[7]

$$\begin{bmatrix} \hat{\mathbf{h}}_{\mathbf{x}}^{step2} & \hat{\mathbf{h}}_0^{step2} \end{bmatrix} = \left(\sum_{t=1}^{T-1} \begin{bmatrix} \hat{\mathbf{x}}_{t+1}\hat{\mathbf{x}}_t' - \widehat{Cov}(\mathbf{u}_{t+1}, \mathbf{u}_t) & \hat{\mathbf{x}}_{t+1} \end{bmatrix}\right) \times \left(\sum_{t=1}^{T-1} \begin{bmatrix} \hat{\mathbf{x}}_t\hat{\mathbf{x}}_t' - \widehat{Var}(\mathbf{u}_t) & \hat{\mathbf{x}}_t \\ \hat{\mathbf{x}}_t' & 1 \end{bmatrix}\right)^{-1} \tag{21}$$

and

$$\begin{aligned} \widehat{Var}(\hat{\mathbf{w}}_{t+1})^{step2} = & \frac{1}{T-1-n_x-1}\sum_{t=1}^{T-1}(\widehat{\mathbf{w}}_{t+1}\left(\widehat{\mathbf{w}}_{t+1}\right)' \\ & -\frac{1}{T-1}\sum_{t=1}^{T-1}\left(\widehat{Var}(\mathbf{u}_t) + \hat{\mathbf{h}}_{\mathbf{x}}\widehat{Var}(\mathbf{u}_t)\hat{\mathbf{h}}_{\mathbf{x}}'\right) \\ & +\frac{1}{T-1}\sum_{t=1}^{T-1}\left(\widehat{Cov}(\mathbf{u}_{t+1}, \mathbf{u}_t)\hat{\mathbf{h}}_{\mathbf{x}}' + \hat{\mathbf{h}}_{\mathbf{x}}\widehat{Cov}(\mathbf{u}_t, \mathbf{u}_{t+1})\right), \end{aligned} \tag{22}$$

with $\hat{\mathbf{\Sigma}}^{step2}$ obtained from a Cholesky decomposition of $\widehat{Var}(\hat{\mathbf{w}}_{t+1})^{step2}$. When T tends to infinity, Andreasen and Christensen (2015) show that the asymptotic distribution of $\boldsymbol{\theta}_2$ is

$$\sqrt{T}\left(\boldsymbol{\theta}_2^{step2} - \boldsymbol{\theta}_2^o\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \left(\mathbf{R}_o^{\boldsymbol{\theta}_2}\mathbf{S}_o^{-1}\left(\mathbf{R}_o^{\boldsymbol{\theta}_2}\right)'\right)^{-1}\right) \tag{23}$$

when using the optimal weighting matrix. Here, $\mathbf{R}_o^{\boldsymbol{\theta}_2} \equiv \frac{\partial \mathbf{q}_T(\boldsymbol{\theta}_2^o)'}{\partial \boldsymbol{\theta}_2}$ and $\mathbf{S}_o \equiv \sum_{\nu=-\infty}^{\infty} E\left[\mathbf{q}_t(\boldsymbol{\theta}_2^o)\,\mathbf{q}_{t-\nu}(\boldsymbol{\theta}_2^o)'\right]$. We estimate $\mathbf{R}_o^{\boldsymbol{\theta}_2}$ using numerical differentiation and $\mathbf{S}_o$ by the Newey-West estimator.

3.1.3 The SR approach: Step 3

The elements in $\mathbf{\Sigma}$ appear in $\boldsymbol{\theta}_{12}$ which are estimated in both the first and second estimation step. Andreasen and Christensen (2015) suggest considering a linear combination of these estimators, i.e.

$$\hat{\boldsymbol{\theta}}_{12}^{step3} = \mathbf{\Lambda}\hat{\boldsymbol{\theta}}_{12}^{step1} + (\mathbf{I} - \mathbf{\Lambda})\,\hat{\boldsymbol{\theta}}_{12}^{step2},$$

[7] Preliminary investigations using our data set showed that the intercept parameters $\mathbf{h}_0$ were weakly identified. It is unsurprising that it is not straightforward to estimate the mean of the pricing factors given the general decline in yields through our sample period. We therefore calibrate the intercept to match the sample mean of the estimated factors. A similar approach is used for an affine term structure model by Adrian et al. (2013).

and determine $\mathbf{\Lambda}$ to minimize the variance $\hat{\boldsymbol{\theta}}_{12}^{step3}$ and hence reduce the efficiency loss from sequential identification. We generally find that $\hat{\mathbf{\Sigma}}^{step1}$ is estimated very inaccurately compared to $\hat{\mathbf{\Sigma}}^{step2}$, meaning that the time series estimate $\hat{\mathbf{\Sigma}}^{step2}$ cannot be improved by adding cross-section information from $\hat{\mathbf{\Sigma}}^{step1}$, i.e. $\mathbf{\Lambda} \approx \mathbf{0}$. Hence, the adopted estimate of $\mathbf{\Sigma}$ after the first two steps is simply given by $\hat{\mathbf{\Sigma}}^{step2}$.

Based on the more accurate estimate of $\mathbf{\Sigma}$ from the second step, it is natural to re-estimate $\boldsymbol{\theta}_{11}$ when conditioned on $\hat{\mathbf{\Sigma}}^{step2}$. That is

$$\hat{\boldsymbol{\theta}}_{11}^{step3} = \arg \min_{\boldsymbol{\theta}_{11} \in \Theta_{11}} Q_{1:T}^{step3} = \frac{1}{2N} \sum_{t=1}^{T} \sum_{j=1}^{n_{y,t}} \left(y_{t,j} - g_j \left(\hat{\mathbf{x}}_t \left(\boldsymbol{\theta}_{11}, \hat{\mathbf{\Sigma}}^{step2} \right); \boldsymbol{\theta}_{11}, \hat{\mathbf{\Sigma}}^{step2} \right) \right)^2. \quad (24)$$

Andreasen and Christensen (2015) show consistency and asymptotic normality of $\hat{\boldsymbol{\theta}}_{11}^{step3}$ with

$$\widehat{Var}\left(\hat{\boldsymbol{\theta}}_{11}^{step3} \right) = \frac{\hat{\mathbf{V}}_{\boldsymbol{\theta}_{11}}^{step3}\left(\hat{\mathbf{\Sigma}}^{step2} \right)}{N} + \hat{\mathbf{K}} \widehat{Var}\left(\hat{\mathbf{\Sigma}}^{step2} \right) \hat{\mathbf{K}}'. \quad (25)$$

The first term $\hat{\mathbf{V}}_{\boldsymbol{\theta}_{11}}^{step3}\left(\hat{\mathbf{\Sigma}}^{step2} \right)/N$ is given by (15) when used on the subset of $\boldsymbol{\theta}_1$ corresponding to $\boldsymbol{\theta}_{11}$. The second term in (25) corrects for estimation uncertainty in $\hat{\mathbf{\Sigma}}^{step2}$ with $\mathbf{K} \equiv \partial \hat{\boldsymbol{\theta}}_{11}^{step3}(\mathbf{\Sigma}) / \partial vech(\mathbf{\Sigma})'$. We estimate $\mathbf{K}$ as suggested in Andreasen and Christensen (2015) and refer to their paper for further details.

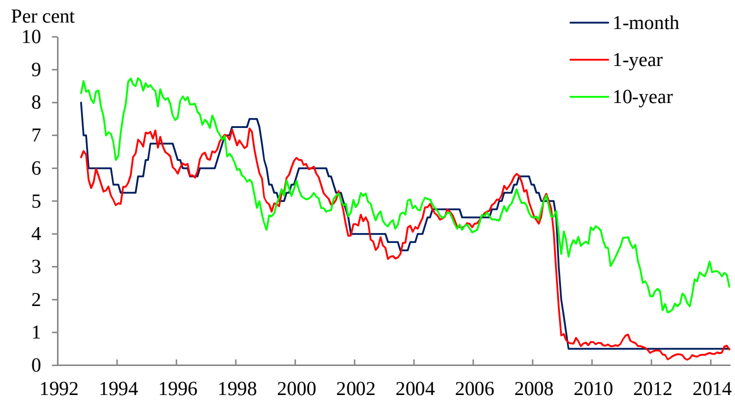
Given the estimated pricing factors $\left\{ \hat{\mathbf{x}}_t \left(\boldsymbol{\theta}_{11}^{step3}, \hat{\mathbf{\Sigma}}^{step2} \right) \right\}_{t=1}^{T}$ from (24), we finally update our estimates of $\boldsymbol{\theta}_2$ using (21) and (22).

4 Data and model specification

4.1 Data

We estimate all the models reported below on UK end-of-month zero-coupon yields with maturities of 12, 18, 24, ..., 120 months for the period October 1992-May 2014. These bond yields are constructed by the Bank of England using the smoothed cubic spline technique of Anderson and Sleath (2001). Data are not consistently available for maturities shorter than 12 months during this period. Unfortunately, there are no ideal alternatives, in part because the UK Treasury bill secondary market is extremely illiquid. We therefore also

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



include the UK policy interest rate (i.e. Bank Rate) as a proxy for the one-period risk-free rate.[8] Selected maturities from our data set are illustrated in Figure 1.

4.2 Number of factors

Most previous studies that estimate shadow rate models have assumed either two or three pricing factors.[9] There is, however, increasing evidence that three principal components of bond yields are insufficient when modelling nominal yields within the Gaussian affine class of models (see e.g. Duffee (2011) and Adrian et al. (2013) for the US and Malik and Meldrum (2014) for the UK). Support for including a fourth pricing factor in a UK shadow rate model can be found by considering the in-sample fit of the model to bond yields. Figure 2 shows the root mean squared fitting error at different maturities from shadow rate models with three and four pricing factors. Average fitting errors from the three-factor model are below five basis points at most maturities, but not at short maturities, where they reach more than 14 basis points for the 12-month yield. Adding a fourth pricing factor allows the model to achieve a superior fit at the short end of the yield curve, with root mean squared errors below three basis points at all maturities.

Figure 3 shows the fitted one-month interest rates from the two models, alongside Bank Rate (our one-month rate proxy). In both models, the one-month interest rates remain

[8] Bank Rate is the interest rate at which the Bank of England remunerates reserves held by commercial banks in accounts at the Bank of England.

[9] One exception is Andreasen and Meldrum (2014), who also estimate models with four factors using US data.

Figure 2: Root mean squared fitting errors by maturity from shadow rate models with three and four pricing factors

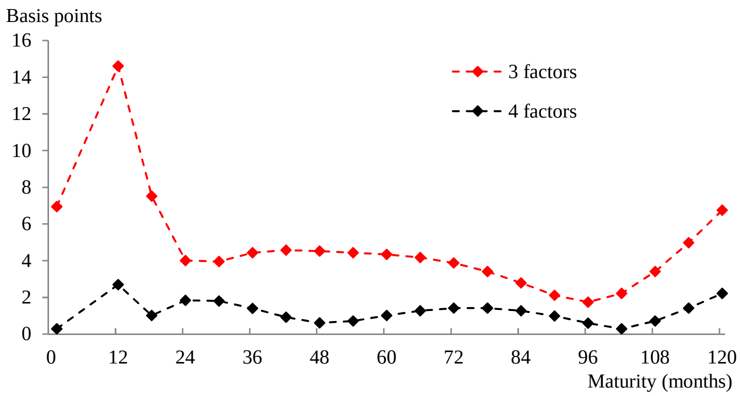


Figure 3: Bank Rate and fitted one-month rates from shadow rate models with three and four pricing factors.

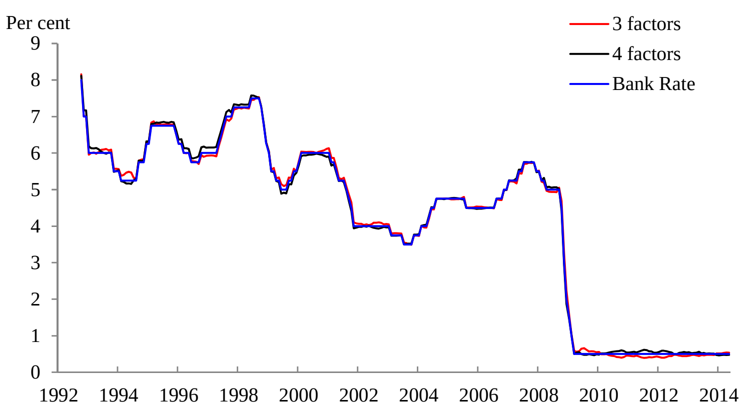


Figure 4: Fitting errors for the 1-month yield from shadow rate models with three and four pricing factors

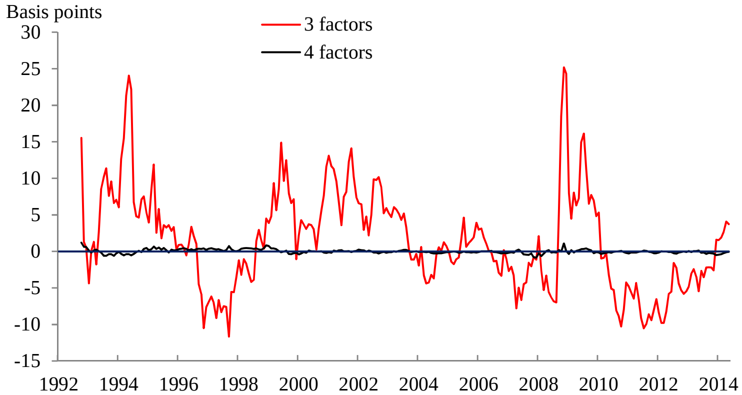
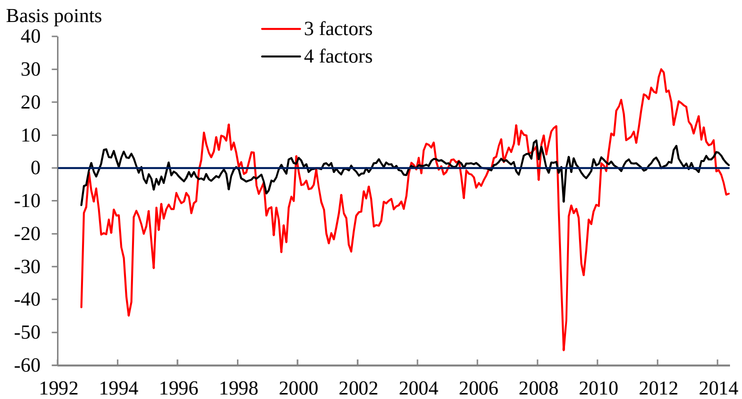


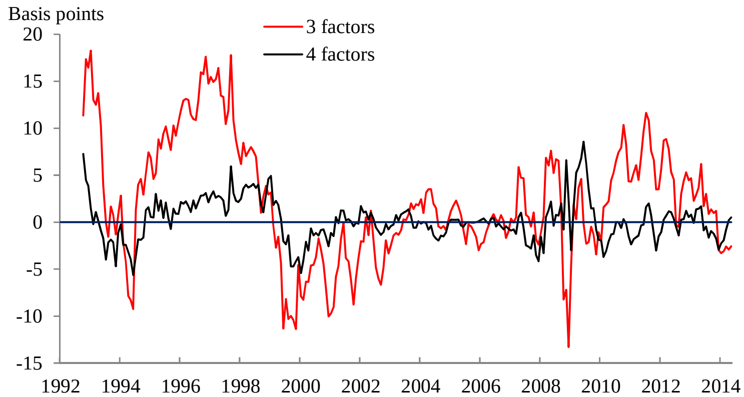
Figure 5: Fitting errors for the 12-month yield from shadow rate models with three and four pricing factors



positive, as required to achieve a close fit to the short end of the yield curve. Figures 4-6 show time series of the model residuals at 1-, 12- and 120-month maturities, respectively. The fitted one-month rates (i.e. the shadow rates) remain positive throughout the sample. The residuals from the three-factor model are typically much larger, particularly at short maturities, and reach a peak of more than 50 basis points in magnitude for the one-year yield.

It is well-known that a good in-sample fit does not necessarily indicate that the model is well-specified. We therefore test the ability of three- and four-factor models to match the specification tests proposed by Dai and Singleton (2002). Following Campbell and Shiller

Figure 6: Fitting errors for the 120-month yield from shadow rate models with three and four pricing factors



(1991), we run the regressions

$$y_{t+1,k-1} - y_{t,n} = \delta_k + \frac{\phi_k}{k-1}(y_{t,k} - r_t) + u_{t,k} \tag{26}$$

for $k = 12, 13, ..., 120$, where $u_{t,k} \sim i.i.d.\,(0, Var\,(u_{t,k}))$. We then explore if the models with different numbers of factors can reproduce the pattern in $\{\phi_k\}_{k=12}^{120}$ and hence capture key moments of the real-world dynamics of bond yields, also known as the LPY(i) test. Following Dai and Singleton (2002), a risk-adjusted version of the Campbell-Shiller regressions in (26) is given by

$$y_{t+1,k-1} - y_{t,k} - (TP_{t+1,k-1} - TP_{t,k-1}) + \frac{1}{k-1}\theta_{t,k-1} = \delta_k^{\mathbb{Q}} + \frac{\phi_k^{\mathbb{Q}}}{k-1}(y_{t,k} - r_t) + v_{t,k} \tag{27}$$

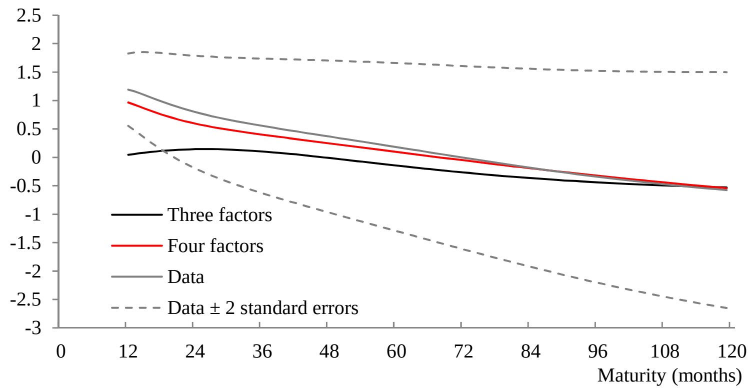
where $v_{t,k} \sim i.i.d.\,(0, Var\,(v_{t,k}))$; $TP_t^{(k)}$ is the term premium, defined as

$$TP_{t,k} = y_{t,k} - \frac{1}{k}\sum_{i=0}^{k-1} E_t\,[r_{t+i}] \tag{28}$$

and $\theta_{t,k-1} = f_{t,k} - E_t\,[r_{t+k}]$ is the forward term premium with $f_{t,k} = -\log\,(P_{t,n+1}/P_{t,n})$. If term premia are correctly specified, then $\phi_k^{\mathbb{Q}} = 1$ for all k. The ability of the models to match these moments is known as the LPY(ii) test and studies whether the models can capture key moments of the $\mathbb{Q}$ dynamics of bond yields.

Figure 7 illustrates the performance of three- and four-factor models against the LPY(i)

Figure 7: LPY(i) test results for shadow rate models with three and four pricing factors. The chart shows estimated slope coefficients $(\widehat{\phi}_k)$ from the regression $y_{t+1,n-1} - y_{t,k} = \delta_k + \phi_k(y_{t,k} - r_t)/(k-1) + u_{t,k}$. Model-implied slope coefficients are estimated using a data set with 100,000 periods simulated from the model, conditional on the point estimates of the parameters.



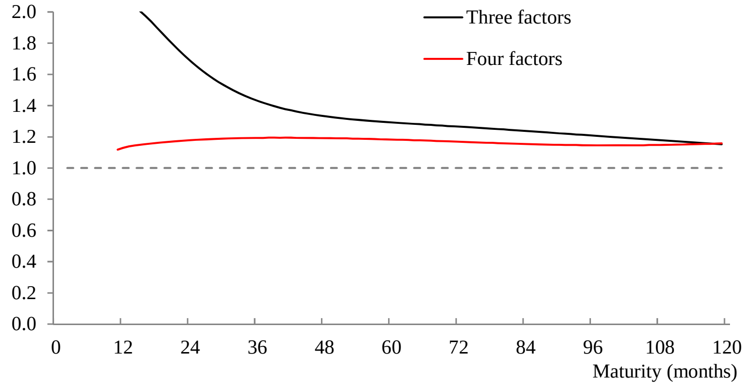
test. Estimates of ϕ_k from a four-factor model are well within an interval of ± 2 estimated standard errors from the estimates obtained on the raw data, whereas those from a three-factor model fall outside this interval at short maturities. This finding is consistent with Malik and Meldrum (2014), who find that a fourth factor helps affine models match the LPY(i) test for the UK and suggests that a fourth factor is required for well-specified time-series dynamics of bond yields.

Turning to the LPY(ii) test, results from which are reported in Figure 8, the estimates of $\phi_k^{\mathbb{Q}}$ from a four-factor model are close to one for all maturities. In a three-factor model, the estimates of $\phi_n^{\mathbb{Q}}$ are higher, particularly at short maturities, which lends further support to the idea that a three-factor model cannot capture the cross-section of bond yields well - particularly the short end of the UK term structure. Overall, we find strong evidence supporting the inclusion of a fourth pricing factor when modelling the short end of the UK term structure.

4.3 The level of the lower bound

Some recent studies using US data have questioned whether it is appropriate to impose that the lower bound for nominal bond yields is exactly zero in shadow rate models since, in practice, US market interest rates have remained at least a few basis point above zero.

Figure 8: LPY(ii) test results for shadow rate models with three and four pricing factors. The chart shows estimated slope coefficients ($\widehat{\phi_k^{\mathbb{Q}}}$) from the regression $y_{t+1,k-1} - y_{t,k} - (TP_{t+1,k-1} - TP_{t,k-1}) + \frac{1}{k-1}\theta_{t,k-1} = \delta_k^{\mathbb{Q}} + \frac{\phi_k^{\mathbb{Q}}}{k-1}(y_{t,k} - r_t) + v_{t,k}$.



Moreover, an increasing number of countries have experienced *negative* nominal interest rates in recent years, including Denmark, the euro area and Switzerland. An alternative specification proposed by Kim and Priebsch (2013) for the US and Lemke and Vladu (2014) for the euro area is to modify (5) to

$$r_t = \max\left(\underline{r}, s\left(\mathbf{x}_t\right)\right)$$

where the lower bound $\underline{r}$ is a free parameter to be freely estimated. Such an extension would have obvious appeal in the UK case, since the Monetary Policy Committee has left Bank Rate unchanged at 0.5% since March 2009. On the other hand, the fact that short-term market rates have been *below* 0.5% during much of the period since March 2009 is inconsistent with the lower bound being so high, with the minimum level of any yield in our sample being 0.07%. Moreover, independent evidence from surveys of professional economists suggests that the perceived probability of future reductions in the policy rate was substantial at times. For example, in a monthly survey of professional economists conducted by Reuters, the maximum probability attributed to a further lowering of the policy rate by any respondent during the period since March 2009 was 40%. Nevertheless, as a robustness check, we have estimated a version of the four-factor shadow rate model in which the lower bound is freely estimated but constrained to be above zero.[10] The estimated lower bound in this model turns out to be below 1 basis point. In our view, this justifies retaining the standard version of the model, with a lower bound equal to zero.

5 Comparison of the ATSM and shadow rate models

5.1 Parameter estimates and model fit

Since our paper is the first to apply a shadow rate model to UK data, in this section we compare its performance with a benchmark Gaussian ATSM. Parameter estimates for the shadow rate model are reported in Table 1 and those for the ATSM in Table 2. The estimated parameters from the two models are very similar, which is not surprising given that the models have similar structures when yields are far away from the lower bound, which they have been for the majority of our sample. The eigenvalues of $\mathbf{h}_x$ in the shadow rate model are 0.9925, 0.9482, 0.9013 and 0.9013, whereas they are 0.9921, 0.9488, 0.9015 and 0.9015 in the affine model. The eigenvalues of $\mathbf{I} - \mathbf{\Phi}$, i.e. the parameters determining the factor persistence under the risk-neutral measure, are also similar across the models. There is, however, a difference between the estimates of α from the two models. This implies that the unconditional mean of the short-term interest rate is a little lower in the shadow rate model, at 3.73%, than in the ATSM, at 4.59%. But the estimated standard errors for α are large and this difference is not statistically significant.

Such small differences in the estimates of α and $\mathbf{\Phi}$ are unsurprisingly not associated with substantial differences in the in-sample fit of the models to bond yields. Figure 9 shows that the root mean squared pricing error by maturity for the two models are almost identical. Fitting errors for 1-, 12- and 120-month yields (Figures 10-12) are very close throughout the sample. The largest differences are observed during the second half of 2012 and early 2013, which - as we will discuss below - corresponds to the period when bond yields were at their lowest level during our sample. Table 3 illustrates that while the shadow rate

[10] The cases of negative nominal interest rates mentioned above suggest that the true lower bound may be below zero. We do not consider this possibility because it seems unlikely that it would be possible to estimate a negative lower bound with reasonable precision from the time series of yields, given that they have remained positive throughout our sample.

Table 1: Parameter estimates for four-factor shadow rate model

Parameter	Estimate	Standard error	Parameter	Estimate	Standard error
α	0.0034	0.0035	$h_{x,34}$	-0.1806^{**}	0.0200
ϕ_{11}	3.776×10^{-8}	–	$h_{x,41}$	0.0175^{*}	0.0068
ϕ_{22}	0.0200^{**}	0.0007	$h_{x,42}$	0.0212^{**}	0.0040
ϕ_{33}	0.0940^{**}	0.0031	$h_{x,43}$	0.0236^{**}	0.0061
ϕ_{44}	0.2488^{**}	0.0038	$h_{x,44}$	0.9321^{**}	0.0069
$h_{x,11}$	0.9504^{**}	0.0141	σ_{11}	$3.095^{**} \times 10^{-4}$	1.637×10^{-5}
$h_{x,12}$	-0.0028	0.0084	σ_{21}	$-4.013^{**} \times 10^{-4}$	2.564×10^{-5}
$h_{x,13}$	-0.0119	0.0104	σ_{22}	$4.774^{**} \times 10^{-4}$	1.878×10^{-5}
$h_{x,14}$	0.0189	0.0154	σ_{31}	$8.215^{**} \times 10^{-5}$	2.228×10^{-5}
$h_{x,21}$	0.0538^{*}	0.0212	σ_{32}	$-5.679^{**} \times 10^{-4}$	2.586×10^{-5}
$h_{x,22}$	1.0004^{**}	0.0174	σ_{33}	$4.699^{**} \times 10^{-4}$	2.105×10^{-5}
$h_{x,23}$	0.0374^{*}	0.0163	σ_{41}	$1.724^{*} \times 10^{-5}$	8.133×10^{-6}
$h_{x,24}$	0.0213	0.0345	σ_{42}	$1.033^{*} \times 10^{-4}$	4.577×10^{-6}
$h_{x,31}$	-0.0182	0.0177	σ_{43}	$-4.361^{**} \times 10^{-4}$	4.330×10^{-6}
$h_{x,32}$	-0.0406^{**}	0.0124	σ_{44}	$1.224^{**} \times 10^{-4}$	5.090×10^{-6}
$h_{x,33}$	0.8576^{**}	0.0149			

Preliminary investigations showed that allowing for non-stationary dynamics under $\mathbb{Q}$ resulted in substantially inferior performance against the LPY(i) test, so we restrict the $\mathbb{Q}$ dynamics to the stationary region. The estimate of ϕ_{11} is essentially at the boundary once we have imposed this constraint, meaning that estimated asymptotic standard errors would be invalid, so we calibrate the parameter to the value shown and re-estimate all other parameters.

model achieves a slightly better fit to short- and long-term yields over the full sample, the difference is larger during the period since March 2009, when Bank Rate was 0.5%. That said, these differences are not large, as both models achieve a close fit to bond yields over the full sample period.

Turning to performance against the LPY tests, the shadow rate model and ATSM perform very similarly against the LPY(i) test (Figure 13). The ATSM does slightly better against the LPY(ii) tests, with estimates of $\phi_k^{\mathbb{Q}}$ close to one at all maturities (Figure 14). But these differences in performance against the LPY tests are not associated with substantial differences in estimates of term premia from the two models. Figure 15 shows estimates of the term premium in 10-year bond yields (as defined in (28)). If the only reason for estimating term structure models is to obtain estimates of long-maturity term premia, it is not obvious that we need to account for the zero lower bound in the case of the UK. This result is similar to findings by Kim and Priebsch (2013) for long-term yields in the US.

The fact that the Gaussian ATSM and the shadow rate model deliver very similar in-

Figure 9: Root mean squared fitting errors by maturity from a four-factor ATSM and shadow rate model

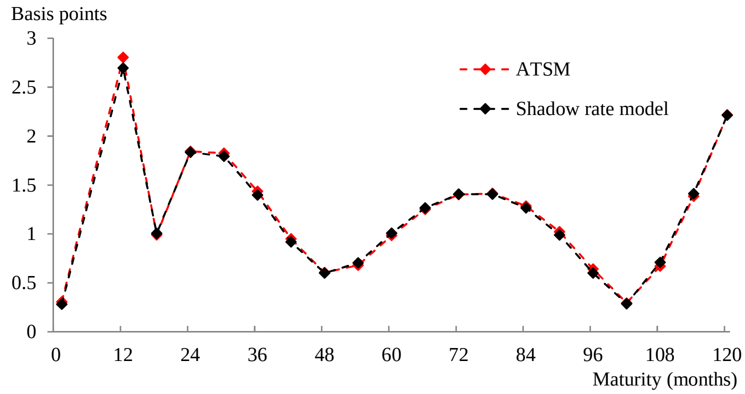


Figure 10: Fitting errors for the 1-month yield from a four-factor ATSM and shadow rate model

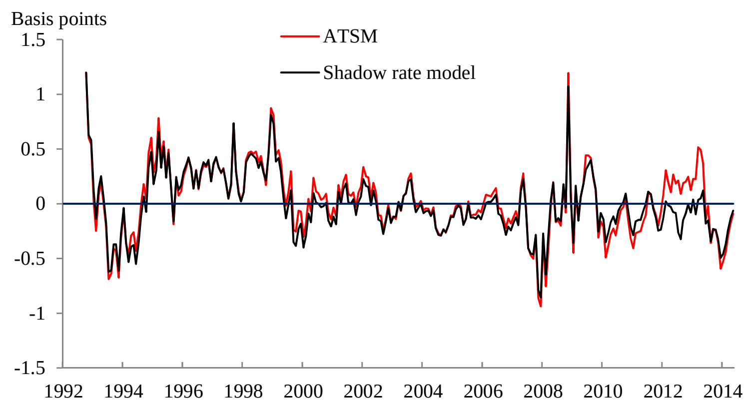


Figure 11: Fitting errors for the 12-month yield from a four-factor ATSM and shadow rate model

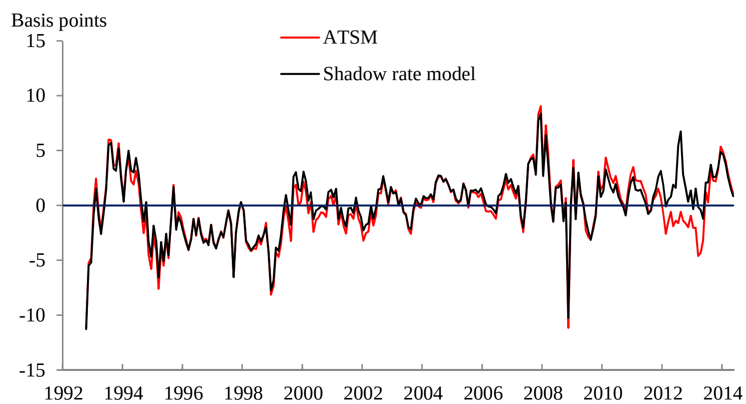


Figure 12: Fitting errors for the 120-month yield from a four-factor ATSM and shadow rate model

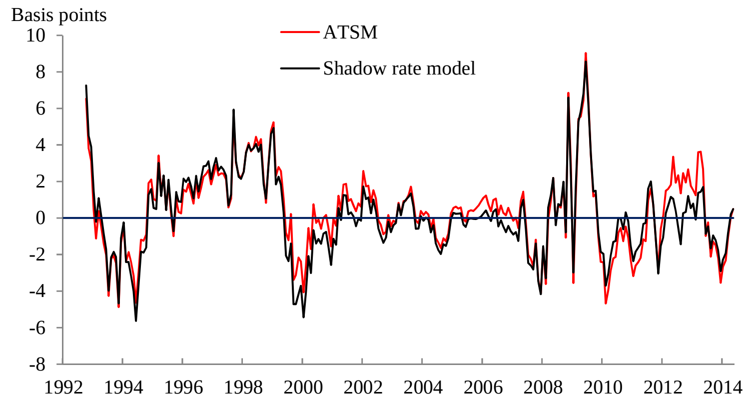


Figure 13: LPY(i) test results for a four-factor ATSM and shadow rate model. The chart shows estimated slope coefficients ($\widehat{\phi}_k$) from the regression $y_{t+1,n-1} - y_{t,k} = \delta_k + \phi_k(y_{t,k} - r_t)/(k-1) + u_{t,k}$. Model-implied slope coefficients are estimated using a data set with 100,000 periods simulated from the model, conditional on the point estimates of the parameters.

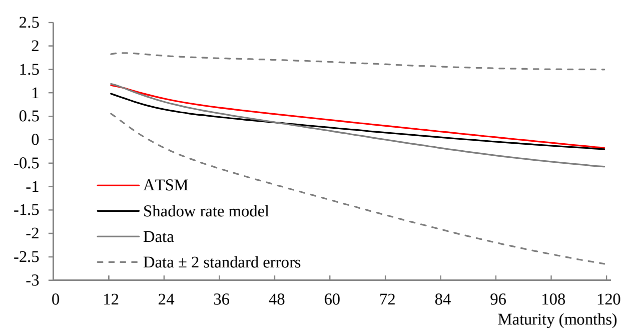


Figure 14: LPY(ii) test results for a four-factor ATSM and shadow rate model. The chart shows estimated slope coefficients $(\widehat{\phi_k^{\mathbb{Q}}})$ from the regression $y_{t+1,k-1} - y_{t,k} - (TP_{t+1,k-1} - TP_{t,k-1}) + \frac{1}{k-1}\theta_{t,k-1} = \delta_k^{\mathbb{Q}} + \frac{\phi_k^{\mathbb{Q}}}{k-1}(y_{t,k} - r_t) + v_{t,k}$.

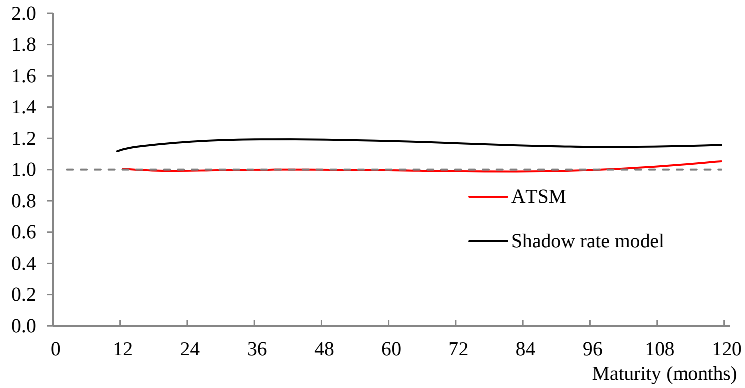


Figure 15: Estimates of ten-year term premia from a four-factor ATSM and shadow rate model

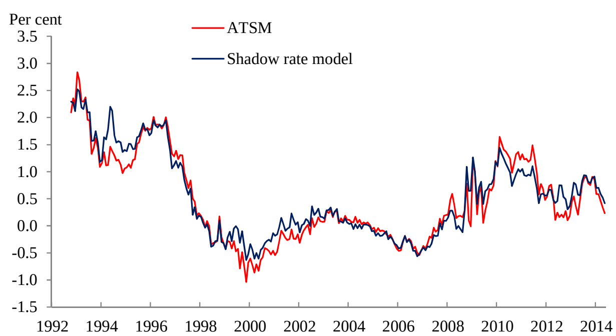


Table 2: Parameter estimates for four-factor ATSM

Parameter	Estimate	Standard error	Parameter	Estimate	Standard error
α	0.0041	0.0105	$h_{x,34}$	-0.1666^{**}	0.0198
ϕ_{11}	3.156×10^{-9}	–	$h_{x,41}$	0.0187^{**}	0.0078
ϕ_{22}	0.0175^{**}	0.0012	$h_{x,42}$	0.0171^{**}	0.0033
ϕ_{33}	0.0856^{**}	0.0055	$h_{x,43}$	0.0224^{**}	0.0056
ϕ_{44}	0.2507^{**}	0.0189	$h_{x,44}$	0.9210^{**}	0.0084
$h_{x,11}$	0.9473^{**}	0.0200	σ_{11}	$3.183^{**} \times 10^{-4}$	1.674×10^{-5}
$h_{x,12}$	-0.0032	0.0080	σ_{21}	$-4.265^{**} \times 10^{-4}$	2.250×10^{-5}
$h_{x,13}$	-0.0178	0.0112	σ_{22}	$4.305^{**} \times 10^{-4}$	1.749×10^{-5}
$h_{x,14}$	0.0155	0.0177	σ_{31}	$8.671^{**} \times 10^{-5}$	2.004×10^{-5}
$h_{x,21}$	0.0522^{*}	0.0240	σ_{32}	$-4.780^{**} \times 10^{-4}$	2.301×10^{-5}
$h_{x,22}$	0.9974^{**}	0.0133	σ_{33}	$4.278^{**} \times 10^{-4}$	1.986×10^{-5}
$h_{x,23}$	0.0360^{*}	0.0158	σ_{41}	$2.784^{**} \times 10^{-5}$	8.643×10^{-6}
$h_{x,24}$	0.0184	0.0380	σ_{42}	$5.965^{**} \times 10^{-5}$	4.665×10^{-6}
$h_{x,31}$	-0.0097	0.0208	σ_{43}	$-3.950^{**} \times 10^{-4}$	4.332×10^{-6}
$h_{x,32}$	-0.0271^{**}	0.0093	σ_{44}	$1.208^{**} \times 10^{-4}$	5.017×10^{-6}
$h_{x,33}$	0.8748^{**}	0.0160			

Preliminary investigations showed that allowing for non-stationary dynamics under $\mathbb{Q}$ resulted in substantially inferior performance against the LPY(i) test, so we restrict the $\mathbb{Q}$ dynamics to the stationary region. The estimate of ϕ_{11} is essentially at the boundary once we have imposed this constraint, meaning that estimated asymptotic standard errors would be invalid, so we calibrate the parameter to the value shown and re-estimate all other parameters.

Table 3: Root mean squared fitting errors from four-factor ATSM and shadow rate model.

	October 1992-May 2014		March 2009-May 2014	
Maturity (months)	ATSM	Shadow rate	ATSM	Shadow rate
1	0.31	0.28	0.27	0.20
12	2.86	2.71	2.37	2.24
120	2.25	2.24	2.79	2.38

All figures report root mean squared fitting errors in basis points.

sample fit, display nearly equal performance against out-of-sample specification tests, and have similar mean projections for future short-term rates (and hence similar term premium estimates) does not imply that the additional complexity of estimating shadow rate models is not justified. The analysis in the following section is concerned with model-implied conditional distributions of short-term interest rates around the mean level at times when bond yields are close to zero. These conditional distributions for short rates can be very different between the models, even at times when yields are not close to the lower bound, which means that an ATSM is likely to be unsuitable for these purposes. To illustrate this point, Figure 16 shows the model-implied conditional probabilities of negative short rates 12 and 120 months ahead from the ATSM (conditional on the parameter and factor estimates obtained at Step 3 of the SR method). During most of the sample, the probability that the short rate would be negative in 12 months' time is neglible. Since March 2009, however, this probability has risen to more than 40% in February 2012, when the 1-year bond yield was 0.46%. Similar findings have been reported previously for the US by Andreasen and Meldrum (2013) and Bauer and Rudebusch (2014). Moreoever, the fact that the probability of negative short rates at *short* forecast horizons is essentially zero when yields are away from the lower bound does not imply that we can necessarily ignore the lower bound at these points. Even in the mid-1990s, when the 10-year yield was typically above 5%, the model-implied probability of negative short rates 10 years ahead was around 15%. In more recent years this probability has risen, reaching 30% in February 2012. In the shadow rate model, of course, these probabilities are zero by construction.

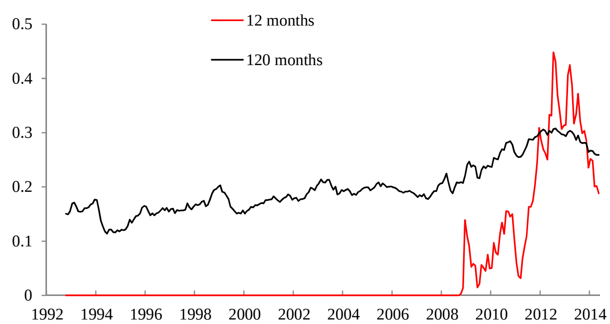
6 Implications of the model for the path of future policy rates

6.1 Policy lift-off dates

Even if the level of the lower bound implied by the data is below 0.5%, with Bank Rate having been held constant at 0.5% since March 2009, a natural question is how long it is before policy rates are expected first to rise *above* 0.5%.[11] One common metric has been to report the maturity at which *forward* rates first increase by 25 basis points above the current policy rate, which appeals to the historical tendency for policy rates to move in multiples of 25 basis points. There are, however, at least two potential drawbacks to this approach. First, it measures the time at which the expected policy rate reaches a particular level, i.e. $\min\left\{h|E_t^{\mathbb{Q}}\left[r_{t+h}\right] \geq r_t + 0.25\%\right\}$, which is not the same as the time at which the policy rate is expected to reach a certain level. Second, forward rates are risk-neutral expectations of

[11] For example, in a speech by Bank of England Deputy Governor Charlie Bean, he said that asset purchases "should be seen as just an emergency weapon for use when policy rates reach the effective lower bound", available at: http://www.bankofengland.co.uk/publications/Pages/news/2014/081.aspx.

Figure 16: Estimates of the conditional probabilities of negative one-month rates at different forecast horizons from a four-factor ATSM



future policy rates and are therefore affected by risk premia, whereas we prefer estimates of the real-world expectation of the lift-off date.

To compute expected lift-off dates for the UK, we follow an approach that is similar to the method proposed by Bauer and Rudebusch (2014). At each point in the sample since March 2009, we simulate the model 100,000 times, conditional on the parameter values and the factor estimates obtained at Step 3 of the SR procedure. More precisely, let $k = 1$ and $j = 1$ and consider the following steps for each time period t.

1. Draw factor disturbances $\boldsymbol{\varepsilon}_{t+k}^{(j)}$ and compute $\mathbf{x}_{t+k}^{(j)} = \mathbf{h}_0 + \mathbf{h}_x \mathbf{x}_{t+k-1}^{(j)} + \boldsymbol{\Sigma} \boldsymbol{\varepsilon}_{t+k}^{(j)}$.

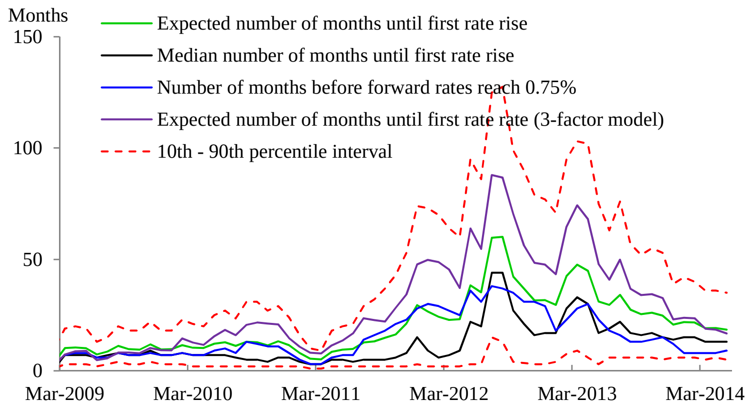
2. Compute the model-implied fitted short rate $r_{t+k}^{(j)} = \max \left\{ 0, \alpha + \boldsymbol{\beta}' \mathbf{x}_{t+k}^{(j)} \right\}$. If $r_{t+k}^{(j)} \geq 0.75\%$, go to step 3; otherwise let $k = k + 1$ and return to step 1.

3. Save the number of periods before the first rate rise, $k^{(j)}$, then let $k = 1$ and $j = j + 1$, and if $j < 100,000$ return to step 1.

Following Bauer and Rudebusch (2014), we use the median of the draws $\left\{ k^{(j)} \right\}_{j=1}^{J}$ as our estimate of the lift-off date (but we also report the mean of the draws, i.e. the expected lift-off horizon). The most substantial difference compared with their method is that they use the risk-neutral measure, whereas we take advantage of the fact that we can compute real-world probabilities using dynamic no-arbitrage term structure models and simulate lift-off dates under the real-world probability measure.[12] Our estimates of the lift-off dates are shown in Figure 17, alongside the 10^{th} and 90^{th} percentiles of the draws. The figure also reports the simpler metric of the number of months before the forward rate reaches 0.75%. Two main findings are apparent. First, while the model-implied lift-off date has been fairly close to the simpler metric for most of the period during which Bank Rate has been at 0.5%, the differences have been substantial at times. The most notable period of difference was a period during late 2011 and early 2012, when the model predicted lift-off much sooner than the simpler metric. For example, in February 2012, the model-implied median lift-off date

[12] Our estimated risk-neutral dynamics of the pricing factors are extremely persistent, with a largest eigenvalue that is very close to one (Table 1). This means that lift-off horizons simulated from the risk-neutral distribution would be extremely long.

Figure 17: Estimates of the lift-off dates from a four-factor shadow rate model



was August 2012 (i.e. six months ahead), while the forward curve did not reach 0.75% until July 2014, a difference of 23 months.[13]

We can gain further insight into this result by considering the time series of forward rates at different maturities, shown in Figure 18. Starting in early 2011, forward rates at medium maturites (the chart shows a 2-year rate) began to fall quite substantially. During this period, the number of months before forward rates reached 0.75% rose substantially (Figure 17). But falls in shorter-term forward rates were much smaller (Figure 18 shows a 6-month rate) and remained above 0.5% until December 2011. It was not until these shorter-term forward rates fell further that the estimated median lift-off horizon started to rise substantially. Even though the horizon at which the forward curve reached 0.75% was becoming shorter during 2011, the fact that short maturity forwards were not falling substantially meant that the probability associated with short-term rate rises did not rise substantially.

Second, the degree of model-implied uncertainty around the lift-off date has been extremely large at times, as illustrated by the dashed lines in Figure 17 corresponding to the 10^{th} and 90^{th} percentiles of the simulated lift-off distribution. For example, continuing to focus on February 2012, the model implied a 10% probability that lift-off would not occur before December 2017 (i.e. a horizon of 70 months). This suggests that we need to be cautious about the precision attached to *any* estimate of the lift-off date, particularly when

[13] Figure 17 also reports estimates of the expected lift-off date from a three-factor model. Our preferred model is the four-factor model, for reasons discussed above. But the estimates from the three-factor model are within the 80% confidence interval from the four-factor model.

Figure 18: Estimates of the lift-off dates from a four-factor shadow rate model alongside model-implied fitted forward rates

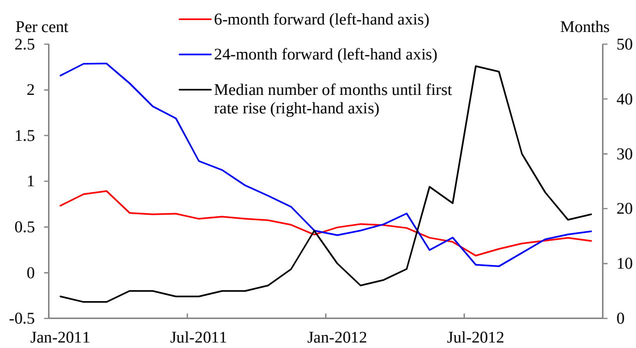
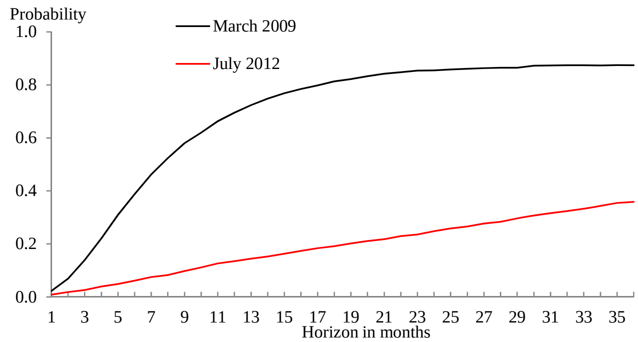


Figure 19: Estimated probabilities of lift-off at different horizons in March 2009 and July 2012



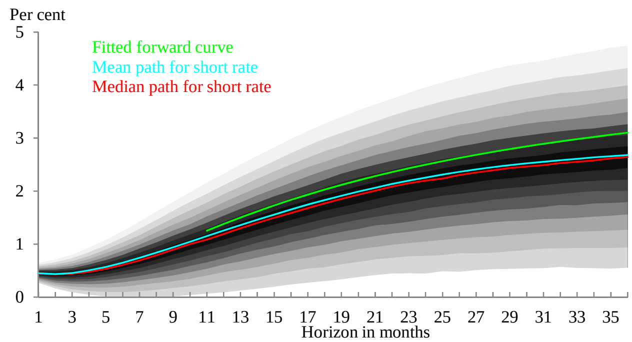
the estimated lift-off date is further into the future. The area covered by the dashed lines in Figure 17 became larger as medium-term forward rates fell in 2011 and 2012.

Rather than focussing on a median (or expected) date for lift-off, which may be well into the future, an alternative metric is to consider the model-implied probabilities of lift-off within a certain fixed horizon. Figure 19 shows these probabilities on two different dates: March 2009, when the MPC first reduced Bank Rate to 0.5%; and July 2012, which is when short-term bond yields reached their lowest point in our sample. In March 2009, investors believed that short rates would remain at the lower bound for a relatively short period, with a probability around 70% that the short rate would be above 0.75% within 12 months. By July 2012, these probabilities had fallen substantially and the probability of lift-off within 12 months was around 15%.

6.2 Policy rate paths

When bond yields are very low, the proximity of the lower bound means that the conditional distribution of short rates is likely to be positively skewed. In a Gaussian ATSM these distributions are always symmetric, so the mean and median are identical. This is not the case in the shadow rate model, which is better suited to capture the conditional skew

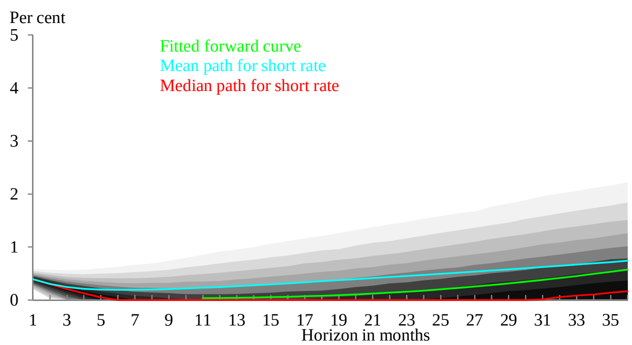
Figure 20: Fan-chart showing conditional probability distributions for short-term interest rates at different horizons in March 2009. The blue line shows the expected path of the short rate under the real-world probability measure (the gap between the green and blue lines is therefore the forward term premium). The red line shows an estimate of the median path for the short rate under the real-world probability measure. The grey shading denotes intervals of 5 centiles around the median, in total covering the region between the 10^{th} and 90^{th} percentiles.



caused by the lower bound. To illustrate the impact of this asymmetry on interest rate expectations, Figures 20 and 21 report the conditional distributions at different maturities in March 2009 and July 2012 respectively. In each chart, the green line shows the fitted forward rate at each maturity - i.e. the expected path of the short rate under $\mathbb{Q}$. The blue line shows the expected path of the short rate under the real-world probability measure (the gap between the green and blue lines is therefore the forward term premium). The red line shows an estimate of the median path for the short rate under the real-world probability measure. The grey shading denotes intervals of 5 centiles around the median, in total covering the region between the 10^{th} and 90^{th} percentiles.

In March 2009, the conditional distribution was close to symmetric despite the recent lowering of Bank Rate to 0.5%, with only small differences between the mean and the mode. The risk-free rate was expected to rise back above 2% within the next two years and the probability that interest rates would fall to zero was very low. In July 2012, however, when forward rates at longer maturities had fallen further, the conditional distribution was much

Figure 21: Fan-chart showing conditional probability distributions for short-term interest rates at different horizons in July 2012. The blue line shows the expected path of the short rate under the real-world probability measure (the gap between the green and blue lines is therefore the forward term premium). The red line shows an estimate of the median path for the short rate under the real-world probability measure. The grey shading denotes intervals of 5 centiles around the median, in total covering the region between the 10^{th} and 90^{th} percentiles.



more positively skewed, with a probability of more than 50% that short rates would be zero in a year's time and an expected path for the short rate that was substantially above the median.

7 Conclusions

This paper estimates a dynamic no-arbitrage shadow rate term structure model for the UK using the sequential regression approach proposed by Andreasen and Christensen (2015). In many important respects, the estimated four-factor shadow rate model has very similar implications compared to a four-factor ATSM. The in-sample fit of the two models is almost identical and they also have similar performance against the standard specification tests proposed by Dai and Singleton (2002). Term premium estimates from the models are similar, which is consistent with previous findings by Kim and Priebsch (2013) for the US and Malik and Meldrum (2014) for the UK. But the probability of negative UK short-term interest rates implied by such an ATSM can be material (which is consistent with previous findings by Andreasen and Meldrum (2013) and Bauer and Rudebusch (2014) for the US). This means such a model is likely to be inappropriate for analysing questions that relate to the conditional distribution of future short rates when yields are low, such as the date when policy rates are expected to lift off from the lower bound.

We estimate lift-off dates using the shadow rate model using a similar technique to that proposed by Bauer and Rudebusch (2014). We show that initially after the UK's Monetary Policy Committee cut its policy rate to 0.5% in March 2009, investors did not expect a long stay at the lower bound. The lift-off horizon rose sharply in mid-2012, when short-term forward rates fell substantially below 0.5% and the conditional distribution of short-term rates was substantially skewed. But the level of uncertainty around the lift-off horizon has been wide at times, particularly when the lift-off date is likely to be well into the future.

References

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 B. J. Christensen (2015). The SR approach: A new estimation procedure for non-linear and non-Gaussian dynamic term structure models. *Journal of Econometrics 184*, 420–451.

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.*

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

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*.

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

Campbell, J. Y. and R. J. Shiller (1991). Yield spreads and interest rate movements: a bird's eye view. *NBER Working Paper 3153*.

Christensen, J. and R. G. D. (2015). Estimating shadow-rate term structure models with near-zero yields. *Journal of Financial Econometrics 13*, 226–259.

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*.

Dai, Q. and K. J. Singleton (2002). Expectation puzzles, time-varying risk premia, and affine models of the term structure. *Journal of Financial Economics 63*, 415–441.

De Rossi, G. (2004). Maximum likelihood estimation of the Cox-Ingersoll-Ross model using particle filters. *Working Paper*.

Doucet, A., N. de Freitas, and N. Gordon (2001). Sequential monte carlo methods in practice. *Springer*.

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

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

Duffie, D. and R. Kan (1996). A yield-factor model of interest rates. *Mathematical Finance 6*, 379–406.

Hordahl, P., O. Tristani, and D. Vestin (2008). The yield curve and macroeconomic dynamics. *The Economic Journal 118*, 1937–1970.

Ichiue, H. and Y. Ueno (2007). Equilibrium interest rate and the yield curve in a low interest rate enviroment. *Bank of Japan Working Paper 07-E-18*.

Ichiue, H. and Y. Ueno (2013). Estimating term premia at the zero bound: An analysis of Japanese, US, and UK yields. *Bank of Japan Working Paper 13-E-8*.

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.

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

Kim, D. H. and K. J. Singleton (2012). Term structure models and the zero bound: an empirical investigation of japanese yields. *Journal of Econometrics 170*, 32–49.

Krippner, L. (2012). Modifying Gaussian term structure models when interest rates are near the zero lower bound. *Reserve Bank of New Zealand, Discussion Paper Series*.

Lemke, W. and A. L. Vladu (2014). A shadow-rate term structure model for the euro area. *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*.

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*.

Rogers, L. C. G. (1995). Which model for term-structure of interest rates should one use? *Mathematical Finance, IMA 65*, 93–116.

Rudebusch, G. D. and T. Wu (2008). A macro-finance model of the term structure, monetary policy, and the economy. *The Economic Journal 118*, 906–926.