Dynamic term structure models:

The best way to enforce the zero lower bound*

Martin M. Andreasen† Andrew Meldrum‡

Aarhus University and CREATES Bank of England

August 30, 2013

Preliminary version

Abstract

This paper studies whether term structure models for US nominal bond yields should enforce the zero lower bound through a quadratic policy rate or a shadow rate. We address the question by estimating quadratic term structure models (QTSMs) and shadow rate models using the sequential regression (SR) approach. Our findings suggest that QTSMs outperform shadow rate models in terms of in-sample fit with two and three latent factors, but not when a fourth factor is included to properly fit short-term yields. A fourth factor is also required in both models to match the slope coefficient in the Campbell-Shiller regressions for excess returns. Importantly, only shadow rate models reproduce the expectation hypothesis in these regressions when excess returns are risk-adjusted, suggesting that the $\mathbb{Q}$ dynamics are best captured by shadow rate models.

Keywords: Bias correction, Perturbation approximation, Quadratic term structure models, Shadow rate models, SR estimation approach.

JEL: C10, C50, G12.

*M. Andreasen greatly acknowledge financial support from the Danish Center for Scientific Computing (DCSC). He also appreciates financial support to CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation. We also note that the views expressed in the present paper are those of the authors and do not necessarily reflect those of the Bank of England or members of the Monetary Policy Committee or Financial Policy Committee.

†Department of Economics and Business, Aarhus university, Fuglesangs Allé 4, 8210 Aarhus V, Denmark. Email: mandreasen@econ.au.dk. Telephone number: +45 8716 5982.

‡Bank of England, Threadneedle Steet, London, EC2R 8AH, UK. Email: andrew.meldrum@bankofengland.co.uk. Telephone number: +44 (0)20 7601 5607.

1 Introduction

Nominal bond yields have reached historically low levels during the recent financial crisis, with short rates at or close to the zero lower bound (ZLB) in several countries. This development has highlighted a well-known shortcoming of affine term structure models (ATSMs) as they generally are unable to ensure positive bond yields. One way to account for the ZLB is to abandon the affine specification of the policy rate and let this rate be quadratic in the pricing factors with appropriate restrictions. Adopting this extension leads to the well-known class of quadratic term structure models (QTSMs) studied in Ahn, Dittmar & Gallant (2002), Leippold & Wu (2002), Realdon (2006) among others. Another way to enforce the ZLB is to restrict policy rates to be non-negative by the max-function as in the class of shadow rate models suggested by Black (1995). The two ways to account for the ZLB imply different dynamics for bond yields but little is currently known about their relative performance on US bond yields.[1] That is, should dynamic term structure models (DTSMs) for US bond yields enforce the ZLB by an appropriately specified quadratic policy rate or by relying on a shadow rate?

The aim of this paper is to address this question by comparing the in- and out-of-sample performance of QTSMs and shadow rate models with two, three, and four pricing factors on postwar US data. In doing so we face two well-known challenges in the term structure literature. The first relates to multi-factor shadow rate models where no closed-form solution is available for bond yields, and we therefore suggest a fourth-order perturbation approximation where the max-function is replaced by a polynomial approximation. To correct for potential pricing errors induced by the approximated max-function, a novel bias-correction to the perturbation solution is derived from bond yields under perfect foresight. We provide a simple recursive implementation of our perturbation approximation, allowing us to obtain solutions to monthly three- and four-factor models with a 10-year interest rate in about 1.7 and 3.7 seconds, respectively. These approximations are shown to be highly accurate with the root mean squared pricing errors around 5 annualized basis points across all maturities.

The second challenge relates to estimation of non-linear DTSMs with latent pricing factors as implied by QTSMs and shadow rate models. One possibility is to approximate the unknown likelihood function for these models by sequential Monte Carlo methods as in Doucet, de Freitas & Gordon (2001), but this procedure is very time consuming for multi-factor DTSMs and therefore rarely attempted.

[1] Kim & Singleton (2012) explore a similar question on Japanese bond yields in models with two latent pricing factors.

A computational 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.[2] We overcome these difficulties by using the sequential regression (SR) approach by Andreasen & Christensen (2013), where latent pricing factors are obtained by cross-section regressions and model parameters are found by a three-step moment-based estimation procedure. The SR approach gives consistent and asymptotically normal estimates and these properties hold under weaker restrictions than typically imposed in likelihood-based inference for DTSMs. For instance, the SR approach allows measurement errors in bond yields to display heteroskedasticity and correlation in both the cross-section and time series dimension. This estimation procedure is particularly well-suited in our context because the considered QTSMs and shadow rate models only differ in their risk-neutral distributions, which may be estimated independently of their physical distributions in the SR approach. Hence, the ability of these models to match in-sample bond yields reported below hold for *any* considered functional form of the market price of risk. We finally highlight the computational efficiency of the SR approach which allow us to estimate three- and four-factor QTSMs in about 10 to 20 minutes. The estimation time for the shadow rate models based on our perturbation approximation is slightly longer but still feasible. Hence, the framework we suggest for solving and estimating DTSMs enforcing the ZLB remains conveniently tractable.

We highlight the following results from our work using monthly US data from July 1961 to May 2013. First, a quadratic policy rate is the best way to enforce the ZLB in two- and three-factor models when measured by the in-sample fit. On the other hand, in four-factor models, a quadratic policy rate and a shadow rate specification deliver broadly the same fit of bond yields, and we are therefore unable to rank the two mechanisms for enforcing the ZLB in this case. Second, the QTSM requires four pricing factors to match the unconditional first and second moments of bond yields, whereas only three factors are required in the shadow rate model. We also find that both models rely on four pricing factors to match the dynamics of bond yields under the $\mathbb{P}$ measure and hence pass the LPY(i) test of Dai & Singleton (2002). Importantly, only shadow rate models pass the closely related

[2] Recent applications of the procedure in DTSMs enforcing the ZLB may be found in Ichiue & Ueno (2007), Kim & Singleton (2012), Bauer & Rudebusch (2013), Christensen & Rudebusch (2013), and Ichiue & Ueno (2013). For some ATSMs without the ZLB restriction, the findings by Duan & Simonato (1999) and de Jong (2000) suggest that the bias in a QML approach based on the extended Kalman filter may be small. We refer to Andreasen (Forthcoming) for a discussion of the asymptotic properties related to a QML approach when estimating non-linear state space models.

LPY(ii) test, suggesting that the $\mathbb{Q}$ dynamics is best captured by a shadow rate specification. Thrid, the shadow rate model appears to offer the best forecasting performance, where it out-performs a benchmark Gaussian ATSM at short maturity and the QTSM at longer maturities. In conclusion, our work therefore suggests that DTSMs for US bond yields should enforce the ZLB by adopting a shadow rate specification instead of a quadratic policy rate, possibly using four pricing factors.

The rest of the paper is organized as follows. Section 2 presents the considered DTSMs, and Section 3 introduces the perturbation method for computing bond yiels in a general class of DTSMs, including shadow rate models. We describe how the considered DTSMs may be estimated by the SR approach in Section 4. In-sample results are reported in Section 5 and the out-of-sample results are presented in Section 6. Concluding comments are provided in Section 7.

2 Dynamic term structure models

We start by describing the Gaussian ATSM in Section 2.1 which serves as our benchmark. The next two subsections present a QTSM and a shadow rate model, respectively, where we restrict focus to pricing factors with Gaussian distributions under both the risk-neutral and physical measure as in the Gaussian ATSM. That is, we consider an affine specification for the market price of risk. We do not study the multivariate version of the model by Cox, Ingersoll & Ross (1985) with independent pricing factors or its extension with correlated factors as in the $A_m(m)$ model by Dai & Singleton (2000), although such models also account the ZLB. The main reason being that the $A_m(m)$ model is unable to reproduce key properties of term premia in the US, whereas these properties are nicely matched by the Gaussian ATSM as shown by Dai & Singleton (2002). In addition, Kim & Singleton (2012) find that the in-sample fit of the QTSM and the shadow rate model clearly outperforms the $A_m(m)$ model on Japanese bond yields.

2.1 The benchmark ATSM

The discrete-time Gaussian ATSM is characterized by three equations. The first 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 some type of Taylor rule as in Clarida, Gali & Gertler (2000). The second equation describes the dynamics of the pricing factors under the risk-neutral measure $\mathbb{Q}$ as a vector autoregressive (VAR) process

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

where $\boldsymbol{\varepsilon}_{t+1}^{\mathbb{Q}} \sim \mathcal{NID}(\mathbf{0}, \mathbf{I})$. The mean level of the pricing factor is controlled by $\boldsymbol{\mu}$ of dimension $n_x \times 1$, while the persistence and 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 given by $P_{t,k} = E_t^{\mathbb{Q}}\left[\exp\{-r_t\}\, P_{t+1,k-1}\right]$. 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 price 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 given by

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

where $\boldsymbol{\varepsilon}_{t+1}^{\mathbb{P}} \sim \mathcal{NID}(\mathbf{0}, \mathbf{I})$. 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. The $\mathbb{P}$ dynamics are then given by

$$\mathbf{x}_{t+1} = \boldsymbol{\Phi}\boldsymbol{\mu} + \mathbf{f_0} + (\mathbf{I} - \boldsymbol{\Phi} + \mathbf{f}_1)\,\mathbf{x}_t + \boldsymbol{\Sigma}\boldsymbol{\varepsilon}_{t+1}^{\mathbb{P}}. \tag{5}$$

The pricing factors are considered to be latent (i.e. unobserved) and a set of normalization restrictions are needed to identify the model. We therefore 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

[3] There exist other normalization schemes, for instance the one recently suggested by Joslin, Singleton & Zhu (2011). We prefer the considered normalization scheme because it is closely related to the one adopted for QTSMs.

estimated by the SR approach, as we explain in Section 4.

2.2 The QTSM

The discrete-time QTSM differs from the Gaussian ATSM by letting the policy rate be quadratic in the pricing factors, i.e.

$$r_t = \alpha + \boldsymbol{\beta}'\mathbf{x}_t + \mathbf{x}_t'\boldsymbol{\Psi}\mathbf{x}_t, \tag{6}$$

where $\boldsymbol{\Psi}$ is a symmetric $n_x \times n_x$ matrix. This specification may also be motivated from a Taylor rule if it displays time-varying parameters as considered in Ang, Boivin, Dong & Loo-Kung (2011). Introducing quadratic terms in the policy rate is useful because they allow the model to enforce the ZLB. The non-negativity conditions for bond yields are i) $\alpha \geq \frac{1}{4}\boldsymbol{\beta}'\boldsymbol{\Psi}^{-1}\boldsymbol{\beta}$ and ii) $\boldsymbol{\Psi}$ to be positive semi-definite (Realdon (2006)). It is worth noting that this way of imposing the ZLB may be applied independently of the chosen dynamics for the pricing factors, and a quadratic policy rule therefore serves as a mechanism to enforce the ZLB.

Given the policy rate in (6), it is convenient to adopt the same specification for the pricing factors as in (2), because it gives the closed-form solution for zero-coupon bonds

$$P_{t,k} = \exp\left\{\tilde{A}_k + \tilde{\mathbf{B}}_k'\mathbf{x}_t + \mathbf{x}_t'\tilde{\mathbf{C}}_k\mathbf{x}_t\right\} \tag{7}$$

for $k = 1, 2, ..., K$, with the recursive formulae for $\tilde{A}_k$, $\tilde{\mathbf{B}}_k$, and $\tilde{\mathbf{C}}_k$ derived in Realdon (2006). Hence, the quadratic terms in (6) imply that all bond yields $y_{t,k} \equiv -\frac{1}{k}\log P_{t,k}$ are quadratic in the pricing factors and bond yields therefore display heteroskedasticity.

For comparability with the benchmark ATSM, we maintain the affine specification for the market price of risk, meaning that the $\mathbb{P}$ dynamics for the pricing factors in the QTSM are given by (5). As in the benchmark ATSM, not all parameters are identified in the QTSM with latent factors. We therefore follow Ahn et al. (2002) and impose the restrictions: i) $\boldsymbol{\Psi}$ is symmetric with diagonal elements equal to one, ii) $\boldsymbol{\mu} \geq \mathbf{0}$, iii) $\boldsymbol{\beta} = \mathbf{0}$, iv) $\boldsymbol{\Phi}$ is diagonal, and v) $\boldsymbol{\Sigma}$ is triangular. This normalization scheme implies an unrestricted $\mathbb{P}$ dynamics for the pricing factors and that the ZLB may be enforced by letting $\alpha = 0$.

2.3 The shadow rate model

The ZLB may also be enforced in DTSMs by introducing a shadow interest rate $s(\mathbf{x}_t)$ as suggested by Black (1995).[4] This shadow rate is unconstrained by the ZLB and may therefore attain negative values. Absent any transaction and storing costs for money, Black (1995) observes that the nominal interest rate cannot be negative because investors may always decide to hold cash. In other words, the nominal interest rate has an option element. This argument leads to the following specification

$$r_t = \max(0, s(\mathbf{x}_t)), \tag{8}$$

where the actual policy rate r_t is the non-negative part of the shadow rate. As with the quadratic policy rule, the concept of a shadow rate serves as a mechanism to enforce the ZLB and may be applied independently of the functional form for $s(\mathbf{x}_t)$ and the considered factor dynamics.

For comparability with the benchmark ATSM, we let the shadow rate be affine in the pricing factors, i.e.

$$s(\mathbf{x}_t) = \alpha + \boldsymbol{\beta}'\mathbf{x}_t, \tag{9}$$

but other specifications may also be considered (see for instance Kim & Singleton (2012)). For the same reason, we also restrict focus to an affine process for the pricing factors under the risk-neutral and physical measure as in the benchmark ATSM, but other specifications could be considered. That is, we impose (4) and (5) in our shadow rate model. Finally, the identification conditions for the shadow rate model are identical to those for the benchmark ATSM in Section 2.1.

3 A perturbation approximation to shadow rate models

Shadow rate models do not attain closed-form expressions for bond prices, except for one-factor models with a Gaussian or square-root process driving the shadow rate (Gorovoi & Linetsky (2004)). Given that one-factor models typically are considered too stylized, numerical approximations are therefore needed when studying multi-factor shadow rate models. The methods used in the literature include i) lattices (Ichiue & Ueno (2007)), ii) finite-difference methods (Kim & Singleton (2012)), iii) Monte Carlo integration (Bauer & Rudebusch (2013)), iv) an option pricing approximation (Krippner (2012),

[4] The idea of considering a shadow rate is also briefly mentioned in Rogers (1995).

Christensen & Rudebusch (2013)), and v) ignoring Jensen's inequality term to solve a Gaussian model by a truncated normal distribution (Ichiue & Ueno (2013)). Each approximation method has its pros and cons and no consensus has so far emerged on the preferred method.

The present paper introduces an entirely different approximation procedure as we suggest a perturbation method to compute bond prices by Taylor series expansions around the deterministic steady state - often equivalent to the unconditional mean.[5] The perturbation method is attractive because it delivers high accuracy and remains computationally tractable even with three and four pricing factors.

We proceed as follows. Section 3.1 introduces the considered class of DTSMs, and Section 3.2 derives the fourth-order Taylor series approximation to bond prices. Shadow rate models are not differentiable everywhere as required to apply the perturbation method, and we therefore suggest replacing the max-function by a fully differentiable function in Section 3.3. To correct for potential pricing errors induced by the approximated max-function, Section 3.3 also derives a bias correction to the perturbation solution by considering bond yields under perfect foresight. Section 3.4 finally describes how to compute conditional expectations of bond yields and hence term premia by the perturbation method.

3.1 A class of DTSMs

The evolution of the pricing factors in the considered class of DTSMs is given by

$$\mathbf{x}_{t+1} = \mathbf{h}\left(\mathbf{x}_t\right) + \sigma \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+1}, \tag{10}$$

where $\boldsymbol{\epsilon}_{t+1}$ has dimension $n_\epsilon \times 1$. These innovations are assumed to be independent and identically distributed with mean zero and covariance matrix $\mathbf{I}$, $\boldsymbol{\epsilon}_{t+1} \sim \mathcal{IID}\left(\mathbf{0}, \mathbf{I}\right)$. It is also assumed that each element of $\boldsymbol{\epsilon}_{t+1}$ has a symmetric probability distribution with finite fourth moment.[6] We impose no further restrictions on the innovations, meaning that $\boldsymbol{\epsilon}_{t+1}$ may be non-Gaussian. The statistical properties of $\boldsymbol{\epsilon}_{t+1}$ may be specified under any probability measure. In most cases, however, the $\mathbb{Q}$ measure is preferred to the $\mathbb{P}$ measure as it implies higher accuracy because the perturbation

[5] This approximation method is widely used in stochastic consumption-based equilibrium models (see Judd & Guu (1997), Schmitt-Grohé & Uribe (2004), Aruoba, Fernandez-Villaverde & Rubio-Ramirez (2006), among others).

[6] Symmetric probability distributions imply that derivatives of bond prices taken k times with respect to $\mathbf{x}$ and n times with respect to the perturbation parameter σ are zero when n is an uneven integer. It is, however, straightforward to allow for non-symmetric probability distributions in the perturbation method and account for skewness or rare disasters (see for instance Andreasen (2012)).

approximation is computed at the deterministic steady state. For instance, simple algebra shows that a second-order approximation under the $\mathbb{Q}$ measure reproduces the exact solution for bond prices in the benchmark ATSM, whereas a fourth-order expansion is needed under the $\mathbb{P}$ measure.

The matrix $\boldsymbol{\eta}$ has dimension $n_x \times n_\epsilon$ and denotes the square root of the covariance matrix for the innovations. An auxiliary parameter $\sigma \geq 0$ scales this matrix and allow us to switch between the stochastic model ($\sigma = 1$) and the deterministic model ($\sigma = 0$). A defining feature of the perturbation method is to compute the Taylor series expansions in $\mathbf{x}_t$ *and* σ. Although the approximation is carried out at the deterministic steady state where $\sigma = 0$ and $\mathbf{x}_{t+1} = \mathbf{x}_t = \mathbf{x}_{ss}$, the derived Taylor series expansions are still able to capture effects of uncertainty by letting $\sigma = 1$.[7]

The function $\mathbf{h}(\mathbf{x}_t)$ in (10) may be non-linear and is required to be four times differentiable for all $\mathbf{x}_t \in \mathbb{R}^{n_x}$ in order to derive the fourth-order Taylor series expansion for bond prices. No further restrictions are imposed on the $\mathbf{h}$-function, meaning that $\mathbf{x}_t$ may be non-stationary. Additional lags of the pricing factors may be included in (10) by an appropriate extension of $\mathbf{x}_t$. Similarly, the assumption that innovations only enter linearly in (10) is also without loss of generality, because a system with non-linearities between pricing factors and innovations may be rewritten into an extended system with only linear innovations.[8] Hence, (10) also accommodates factor dynamics with time-varying volatility, for instance when specified by stochastic volatility or GARCH.

The fundamental asset pricing equation gives the well-known recursive expression for zero-coupon bond prices at time t with k periods to maturity

$$P_{t,k} = \mathbb{E}_t \left[M\left(\mathbf{x}_t, \mathbf{x}_{t+1}\right) P_{t+1,k-1} \right]. \tag{11}$$

Here, $\mathbb{E}_t$ is the conditional expectation given information available at time t and under the same probability measure as the innovations in (10). The stochastic discount factor is denoted by $M(\mathbf{x}_t, \mathbf{x}_{t+1})$ and required to be four times differentiable for all $(\mathbf{x}_t, \mathbf{x}_{t+1}) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_x}$ to derive the fourth-order Taylor series expansion. As an illustration, consider the stochastic discount factor for the benchmark

[7] Note that the steady state is a fixed-point in (10) and that this point corresponds to the unconditional mean in linear systems.

[8] Two examples are provided in Andreasen, Fernandez-Villaverde & Rubio-Ramirez (2013)

ATSM, which is $\exp\{-r_t\}$ under the $\mathbb{Q}$ measure and

$$M(\mathbf{x}_t, \mathbf{x}_{t+1}) = \exp\left\{-r_t - \frac{1}{2}(\mathbf{f}_0 + \mathbf{f}_1\mathbf{x}_t)'(\mathbf{f}_0 + \mathbf{f}_1\mathbf{x}_t) - (\mathbf{f}_0 + \mathbf{f}_1\mathbf{x}_t)'\sigma\boldsymbol{\epsilon}_{t+1}\right\} \tag{12}$$

when the pricing is done under the $\mathbb{P}$ measure.

The true solution for bond prices in the class of DTSMs considered is $P_{t,k} = P_k(\mathbf{x}_t, \sigma)$. The perturbation method approximates these unknown functions by Taylor series expansions around the deterministic steady state. When applying these approximations, it is often useful to adopt a log-transformation of bond prices as continuously compounded interest rates are linear functions of $\log P_{t,k}$. We therefore re-write (11) as

$$\exp\left(p^{t,k}\right) = \mathbb{E}_t\left[M(\mathbf{x}_t, \mathbf{x}_{t+1})\exp\left(p^{t+1,k-1}\right)\right] \tag{13}$$

with $p^{t,k} \equiv \log P_{t,k}$ and apply (13) for the approximation. A fourth-order Taylor series expansion of $p^{t,k}$ in $\mathbf{x}_t$ and σ is given by

$$\begin{aligned} p^k(\mathbf{x}_t, \sigma) \approx\ & p^k(\mathbf{x}_{ss}, 0) + \textstyle\sum_{\alpha_1=1}^{n_x}\left(p^k_{\mathbf{x}}(\alpha_1) + \frac{3}{6}p^k_{\sigma\sigma\mathbf{x}}(\alpha_1)\sigma^2\right)\bar{\mathbf{x}}_t(\alpha_1) \\ & + \frac{1}{2}\textstyle\sum_{\alpha_1=1}^{n_x}\sum_{\alpha_2=1}^{n_x}\left(p^k_{\mathbf{xx}}(\alpha_1,\alpha_2) + \frac{6}{24}p^k_{\sigma\sigma\mathbf{xx}}(\alpha_1,\alpha_2)\sigma^2\right)\bar{\mathbf{x}}_t(\alpha_1)\bar{\mathbf{x}}_t(\alpha_2) \\ & + \frac{1}{6}\textstyle\sum_{\alpha_1=1}^{n_x}\sum_{\alpha_2=1}^{n_x}\sum_{\alpha_3=1}^{n_x} p^k_{\mathbf{xxx}}(\alpha_1,\alpha_2,\alpha_3)\bar{\mathbf{x}}_t(\alpha_1)\bar{\mathbf{x}}_t(\alpha_2)\bar{\mathbf{x}}_t(\alpha_3) \\ & + \frac{1}{6}\textstyle\sum_{\alpha_1=1}^{n_x}\sum_{\alpha_2=1}^{n_x}\sum_{\alpha_3=1}^{n_x}\sum_{\alpha_4=1}^{n_x} p^k_{\mathbf{xxxx}}(\alpha_1,\alpha_2,\alpha_3,\alpha_4)\bar{\mathbf{x}}_t(\alpha_1)\bar{\mathbf{x}}_t(\alpha_2)\bar{\mathbf{x}}_t(\alpha_3)\bar{\mathbf{x}}_t(\alpha_4) \\ & + \frac{1}{2}p^k_{\sigma\sigma}\sigma^2 + \frac{1}{24}p^k_{\sigma\sigma\sigma\sigma}\sigma^4 \end{aligned} \tag{14}$$

where $\bar{\mathbf{x}}_t \equiv \mathbf{x}_t - \mathbf{x}_{ss}$ for $k = 1, 2, ..., K$.[9] Here, $p^k_{\sigma^j\mathbf{x}^n}$ refers to the derivative of $p^{t,k}$ with respect to σ taken j times and with respect to $\mathbf{x}_t$ taken n times. These derivatives are evaluated at the deterministic steady state. Due to the log-transformation of bond prices, the perturbation approximation to bond yields is simply given by

$$y_k^{app}(\mathbf{x}_t) = -\frac{1}{k}p^k(\mathbf{x}_t, \sigma). \tag{15}$$

[9] Only non-zero terms are included in (14). See Schmitt-Grohé & Uribe (2004) and Andreasen (2012) for further details.

3.2 Computing the perturbation approximation to DTSMs

One possibility for computing the unknown terms in (14) uses (10) and (13) for $k = 1, 2, ..., K$ to jointly find all bond price derivatives by standard solution algorithms for the perturbation method.[10] In the context of DTSMs, this procedure is not numerically efficient because the $\mathbf{h}$-function is taken to be unknown and the recursive relationship between bond prices is ignored. We therefore present a fully efficient two-step solution algorithm tailored to the considered class of DTSMs.

Step 1: The first step computes all derivatives for bond prices with one period to maturity. Letting $k = 1$ in (13), we get

$$p^{t,1} = \log\left(\mathbb{E}_t\left[M\left(\mathbf{x}_t, \mathbf{x}_{t+1}\right) \times 1\right]\right), \tag{16}$$

as zero-coupon bonds pay one unit of currency at maturity. When the pricing is done under $\mathbb{Q}$, the stochastic discount factor does not depend on $\mathbf{x}_{t+1}$ and (16) simplifies to $p^{t,1} = \log\left(M\left(\mathbf{x}_t\right)\right)$. The terms $p^1_{\mathbf{x}}$, $p^1_{\mathbf{xx}}$, $p^1_{\mathbf{xxx}}$, and $p^1_{\mathbf{xxxx}}$ therefore follow by simple differentiation, and all derivatives with respect to the perturbation parameter σ are zero, i.e. $p^1_{\sigma\sigma} = 0$, $p^1_{\sigma\sigma\sigma\sigma} = 0$, $p^1_{\sigma\sigma\mathbf{x}} = \mathbf{0}$, and $p^1_{\sigma\sigma\mathbf{xx}} = \mathbf{0}$.[11] If the pricing is done under under $\mathbb{P}$, we suggest to specify the probability distribution of $\boldsymbol{\epsilon}_{t+1}$ and manually compute $\mathbb{E}_t\left[M\left(\mathbf{x}_t, \mathbf{x}_{t+1}\right)\right]$. Based on this expression, all required bond price derivatives for p^1 then follow by simple differentiation.

Step 2: The second step uses the perturbation method to recursively compute the remaining bond price derivatives. Based on (10) and (13) for a *given* k, we define the function

$$\begin{aligned} F^k\left(\mathbf{x}_t, \sigma\right) \quad \equiv \quad & E_t[\exp\left(p^{t,k}\left(\mathbf{x}_t, \sigma\right)\right) - M\left(\mathbf{h}\left(\mathbf{x}_t, \sigma\right) + \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}, \mathbf{x}_t\right) \\ & \times \exp\left(p^{t+1,k-1}\left(\mathbf{h}\left(\mathbf{x}_t, \sigma\right) + \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}, \sigma\right)\right)]. \end{aligned} \tag{17}$$

The relationship for bond prices in (13) always holds, meaning that $F^k\left(\mathbf{x}_t, \sigma\right) = 0$ and $F^k_{\mathbf{x}^i\sigma^j}\left(\mathbf{x}, \sigma\right) = 0$ for all values of $\mathbf{x}$ and σ. For consumption-based equilibrium models, Andreasen & Zabczyk (2010) show that these conditions determine the remaining bond price derivatives up to third order. The same property clearly holds for DTSMs, and we provide the expressions for bond price derivatives up

[10] They include *Dynare* and *Dynare++* by Kamenik (2005), *Perturbation AIM* by Swanson, Anderson & Levin (2005), and codes accompanying Schmitt-Grohé & Uribe (2004).

[11] When computing $p^1_{\mathbf{x}}$, $p^1_{\mathbf{xx}}$, $p^1_{\mathbf{xxx}}$, and $p^1_{\mathbf{xxxx}}$ from $p^{t,1} = \log\left(M\left(\mathbf{x}_t\right)\right)$ we recommend using software for symbolic differentiation. In MATLAB, this is carried out by using the "diff" or "jacobian" function.

to fourth order in Appendix A. These Ricatti equations have a simple recursive structure and the remaining bond price derivatives are therefore easily obtained by simple summations.[12]

3.3 The perturbation approximation and shadow rate models

Additional considerations are needed when applying the perturbation method to shadow rate models because the max-function in the policy rate and hence $M(\mathbf{x}_t, \mathbf{x}_{t+1})$ are not differentiable everywhere. We address this problem by replacing the max-function in (8) by a fully differentiable approximation to this function. The considered approximation is obtained from a fourth-order polynomial where parameters are calibrated to match the max-function in a relevant interval for the policy rate. We let $\widehat{\max}\{r_t, 0\}$ denote the approximated max-function which is displayed in Figure 1.[13]

To correct for potential pricing errors induced by the approximated max-function, we derive a novel bias correction to the perturbation solution from bond yields under perfect foresight.[14] To present our bias correction, let $\widetilde{y}_k(\mathbf{x}_t)$ denote bond yields under perfect foresight. For the shadow rate model, we have

$$\widetilde{y}_k(\mathbf{x}_t) = \frac{1}{k}\left\{\sum_{i=1}^{k} \max\left\{\alpha + \boldsymbol{\beta}'\mathbf{x}_{t+i}, 0\right\}\right\} \tag{18}$$

for $k = 1, 2, ..., K$, where

$$\mathbf{x}_{t+i} = \left(\boldsymbol{\Phi}^{-1}\left(\mathbf{I} - (\mathbf{I} - \boldsymbol{\Phi})^i\right)\right)\boldsymbol{\Phi}\boldsymbol{\mu} + (\mathbf{I} - \boldsymbol{\Phi})^i \mathbf{x}_t. \tag{19}$$

Let $\widetilde{y}_k^{app}(\mathbf{x}_t)$ denote bond yields under perfect foresight when using the approximated max-function $\widehat{\max}\{r_t, 0\}$. Its value in the the shadow rate model is given by

$$\widetilde{y}_k^{app}(\mathbf{x}_t) = \frac{1}{k}\left\{\sum_{i=1}^{k} \widehat{\max}\left\{\alpha + \boldsymbol{\beta}'\mathbf{x}_{t+i}, 0\right\}\right\} \tag{20}$$

[12] The generality of the considered class of DTSMs imply somewhat involved expressions for several third- and fourth-order derivatives in Appendix A. These terms greatly simplify when i) the pricing under $\mathbb{Q}$ as all derivatives of $M(\mathbf{x}_t, \mathbf{x}_{t+1})$ with respect to $\mathbf{x}_{t+1}$ are zero and/or ii) $\mathbf{h}(\mathbf{x}_t)$ is linear with $\mathbf{h}_{\mathbf{xx}} = \mathbf{0}$, $\mathbf{h}_{\mathbf{xxx}} = \mathbf{0}$, and $\mathbf{h}_{\mathbf{xxxx}} = \mathbf{0}$.

[13] Other functions may be applied to approximate the max-function, for instance $f(r) = r\exp(nr)/(\exp(nr) + 1)$ which is fully differentiable and converges to $\max\{r, 0\}$ for $n \longrightarrow \infty$. However, the perturbation approximation in (14) only uses a fourth-order Taylor-series expansion of $f(r)$. Although this particular fourth order polynomial may display high accuracy close to the deterministic steady state, its accuracy may easily deteriorate far from this point when n is large. It is for this reason that we prefer to specify a fourth-order polynomial and calibrate its parameters such that the approximated max-function displays satisfying behaviour in a relevant interval for the policy rate.

[14] A bias correction is sometimes also referred to as a control variate technique and widely used in finance to improve the accuracy of numerical approximations (see for instance Hull (2012)).

for $k = 1, 2, ..., K$ with $\mathbf{x}_{t+i}$ determined from (19). Hence, the bias induced by the approximated max-function under perfect foresight is given by

$$\widetilde{b}_k(\mathbf{x}_t) = \widetilde{y}_k(\mathbf{x}_t) - \widetilde{y}_k^{app}(\mathbf{x}_t) \tag{21}$$

for $k = 1, 2, ..., K$. For the perturbation solution in (14), we assume the bias $b_k(\mathbf{x}_t)$ due to the approximated max-function is affine in the bias under perfect foresight, i.e. $b_k(\mathbf{x}_t) = \gamma_{0,k} + \gamma_{1,k}\widetilde{b}_k(\mathbf{x}_t)$ where $\gamma_{0,k}$ and $\gamma_{1,k}$ are free parameters. The intercept $\gamma_{0,k}$ is included to correct potential biases in $p_{\sigma\sigma}^k$ and $p_{\sigma\sigma\sigma\sigma}^k$, whereas $\gamma_{1,k}\widetilde{b}_k(\mathbf{x}_t)$ captures biases depending on the level of the pricing factors $\mathbf{x}_t$. Hence, the considered bias-adjusted perturbation approximation to bond yields is

$$\begin{aligned} y_k^{BiasAdj}(\mathbf{x}_t) &= y_k^{app}(\mathbf{x}_t) + b_k(\mathbf{x}_t) \\ &= \gamma_{0,k} + y_k^{app}(\mathbf{x}_t) + \gamma_{1,k}\widetilde{b}_k(\mathbf{x}_t), \end{aligned} \tag{22}$$

where $y_{t,k}^{app}$ is given by (15). To calibrate $\gamma_{0,k}$ and $\gamma_{1,k}$ we rely on the Monte Carlo method to obtain the exact solution for bond yields $y_k(\mathbf{x}_t)$. The pricing errors $z_{t,k}$ implied by the bias-adjusted perturbation approximation are then given by

$$y_k(\mathbf{x}_t) = \gamma_{0,k} + y_k^{app}(\mathbf{x}_t) + \gamma_{1,k}\widetilde{b}_k(\mathbf{x}_t) + z_{t,k} \tag{23}$$

for $k = 1, 2, ..., K$ and $t = 1, 2, ..., T$, or

$$\mathbf{y}_k - \mathbf{y}_k^{app} = \mathbf{1}\gamma_{0,k} + \widetilde{\mathbf{b}}_k\gamma_{1,k} + \mathbf{z}_k, \tag{24}$$

where $\mathbf{y}_k$, $\mathbf{y}_k^{app}$, $\mathbf{1}$, $\widetilde{\mathbf{b}}_k$, and $\mathbf{z}_k$ have dimension $T \times 1$ for $k = 1, 2, ..., K$. We then suggest calibrating $\gamma_{0,k}$ and $\gamma_{1,k}$ by minimizing the squared pricing errors, which is equivalent to regressing $\mathbf{y}_k - \mathbf{y}_k^{app}$ on $\mathbf{X}_k \equiv \begin{bmatrix} \mathbf{1} & \widetilde{\mathbf{b}}_k \end{bmatrix}$, i.e.

$$\begin{bmatrix} \gamma_{0,k}^* \\ \gamma_{1,k}^* \end{bmatrix} = \left(\mathbf{X}_k'\mathbf{X}_k\right)^{-1}\mathbf{X}_k'\left(\mathbf{y}_k - \mathbf{y}_k^{app}\right) \tag{25}$$

for $k = 1, 2, ..., K$. This calibration of the bias correction is clearly model-dependent because the

considered shadow rate model and its parameters affect $y_{t,k}^{app}$ and $\widetilde{b}_{t,k}$. To account for this dependence we therefore suggest the following joint estimation and calibration procedure, where steps 2 to 5 may be iterated if desired:

Step 1: Let $\gamma_{0,k} = 0$ and $\gamma_{1,k} = 1$ for $k = 1, 2, ..., K$ and estimate the model.

Step 2: Simulate $\{\mathbf{x}_s\}_{s=1}^{T}$ from the estimated model.

Step 3: For $\{\mathbf{x}_s\}_{s=1}^{T}$, compute $\mathbf{y}_k$, $\mathbf{y}_k^{app}$ and $\widetilde{\mathbf{b}}_k$ for $k = 1, 2, ..., K$.

Step 4: Use (25) to compute $\left\{\gamma_{0,k}^*, \gamma_{1,k}^*\right\}_{k=1}^{K}$.

Step 5: Re-estimate the model given $\left\{\gamma_{0,k}^*, \gamma_{1,k}^*\right\}_{k=1}^{K}$ from step 4.

3.4 Expected future short rates and term premia

We define term premia as the difference between bond yields and average expected future short rates, i.e.

$$TP_k(\mathbf{x}_t) = y_k(\mathbf{x}_t) - \frac{1}{k}\sum_{i=0}^{k-1} \mathbb{E}_t^{\mathbb{P}}\left[y_1(\mathbf{x}_{t+i})\right] \tag{26}$$

for $k = 1, 2, ..., K$ as in Dai & Singleton (2002). Bond yields $y_k(\mathbf{x}_t)$ are approximated by (15) and we therefore only need to compute $\mathbb{E}_t^{\mathbb{P}}\left[y_1(\mathbf{x}_{t+i})\right]$. These conditional expectations are also straightforward to approximate by the perturbation method as shown in Appendix B. The formulas in Appendix B may be used to obtain conditional expectations of bond yields for any maturity, implying that our results are also useful when including survey data on future bond yields in DTSMs as suggested by Kim & Orphanides (2012). It is finally worth noticing that we do not need to bias correct bond yields or expected future short rates when computing term premia in shadow rate models, because these bias corrections are identical and therefore cancel out in (26).

4 The estimation procedure

All considered DTSMs are estimated using the sequential regression (SR) approach by Andreasen & Christensen (2013). Several reasons motivate our choice. First, QTSMs and shadow rate models introduce a non-linear filtering problem which is easily addressed by the SR approach as latent pricing

factors are estimated by a sequence of non-linear cross-section regressions. Second, the SR approach gives consistent and asymptotically normal estimates under weaker restrictions than typically considered for likelihood-based inference. The robust nature of the estimation approach is particularly attractive in our context because all the considered DTSMs differ only in their $\mathbb{Q}$ dynamics which may be estimated independently of the $\mathbb{P}$ dynamics in the SR approach. Hence, the models' abilities to match in-sample bond yields in the SR approach apply for *any* functional form of the market price of risk $\mathbf{f}(\mathbf{x}_t)$. Finally, the computational efficiency of the SR approach is very appealing, as it allows us to estimate three- and four-factor models without any difficulty.

We next present the SR approach and describe how the latent factors and model parameters are estimated in the models considered.

4.1 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.

$$y_{t,k} = g_k(\mathbf{x}_t; \boldsymbol{\theta}_1) + v_{t,k}. \tag{27}$$

The functional relationship between the pricing factors and bond yields is parameterized by $\boldsymbol{\theta}_1$, which corresponds to the risk-neutral parameters in no-arbitrage DTSMs. For the benchmark 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 $\boldsymbol{\theta}_1^{ATSM} \equiv \begin{bmatrix} \alpha & vec\left(diag(\boldsymbol{\Phi})\right)' & vech\left(\boldsymbol{\Sigma}\right)' \end{bmatrix}'$. The QTSM induces a slightly more complicated expression for bond yields because $g_k^{QTSM}\left(\mathbf{x}_t; \boldsymbol{\theta}_1^{QTSM}\right) \equiv -\frac{1}{k}\left(\tilde{A}_k + \tilde{\mathbf{B}}_k'\mathbf{x}_t + \mathbf{x}_t'\tilde{\mathbf{C}}_k\mathbf{x}_t\right)$ and $\boldsymbol{\theta}_1^{QTSM} \equiv \begin{bmatrix} \left(\boldsymbol{\theta}_1^{ATSM}\right)' & \boldsymbol{\mu}' & vec\left(\boldsymbol{\Psi}\right)' \end{bmatrix}'$. In the shadow rate model, $g_k^{SH}\left(\mathbf{x}_t; \boldsymbol{\theta}_1^{SH}\right) \equiv \gamma_{0,k} + \widetilde{y}_k^{app}(\mathbf{x}_t) + \gamma_{1,k}\widetilde{b}_k(\mathbf{x}_t)$ and therefore also non-linear in the pricing factors 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 may display heteroskedasticity and correlation in both the cross-section and the time series dimension.

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

$$\mathbf{x}_{t+1} = \mathbf{h}\left(\mathbf{x}_t, \boldsymbol{\epsilon}_{t+1}^{\mathbb{P}}; \boldsymbol{\theta}_1, \boldsymbol{\theta}_2\right). \tag{28}$$

The $\mathbf{h}$-function may depend on $\boldsymbol{\theta}_1$ containing the risk-neutral parameters and $\boldsymbol{\theta}_2$, which in our case are $\mathbf{f}_0$ and $\mathbf{f}_1$ specifying the market price of risk. All the DTSMs considered have a linear and unrestricted transition function which we represent by

$$\mathbf{x}_{t+1} = \mathbf{h}_0 + \mathbf{h}_{\mathbf{x}}\mathbf{x}_t + \boldsymbol{\varepsilon}_{t+1}^{\mathbb{P}}, \tag{29}$$

with $\mathbf{h}_0 \equiv \boldsymbol{\Phi}\boldsymbol{\mu} + \mathbf{f}_{\mathbf{0}}$, $\mathbf{h}_{\mathbf{x}} \equiv \mathbf{I} - \boldsymbol{\Phi} + \mathbf{f}_1$, and $\boldsymbol{\varepsilon}_{t+1}^{\mathbb{P}} \sim \mathcal{NID}\left(\mathbf{0}, \boldsymbol{\Sigma}\boldsymbol{\Sigma}'\right)$. Adopting this parametrization of the $\mathbf{h}$-function, we have $\boldsymbol{\theta}_2 \equiv \begin{bmatrix} \mathbf{h}_0' & vec\left(\mathbf{h}_{\mathbf{x}}\right)' & vech\left(\boldsymbol{\Sigma}\right)' \end{bmatrix}'$.

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.

4.1.1 The SR approach: Step 1

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

$$\hat{\mathbf{x}}_t\left(\boldsymbol{\theta}_1\right) = \arg\min_{\mathbf{x}_t \in \mathcal{X}_t} Q_t = \frac{1}{2n_{y,t}} \sum_{k=1}^{n_{y,t}} \left(y_{t,k} - g_k\left(\mathbf{x}_t; \boldsymbol{\theta}_1\right)\right)^2 \tag{30}$$

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 benchmark ATSM as g_k^{ATSM} is linear in the pricing factors. For the QTSM and the shadow rate model, the regressions in (30) are non-linear and solved using the Levenberg-Marquardt method with the pricing factors from the period time period $\hat{\mathbf{x}}_{2,t-1}\left(\boldsymbol{\theta}_1\right)$ serving as ideal starting values for $t = 2, 3, ..., T$.

The model parameters $\boldsymbol{\theta}_1$ are obtained by pooling all squared residuals from (30) and minimizing their sum with respect to $\boldsymbol{\theta}_1$, i.e.

$$\hat{\boldsymbol{\theta}}_1^{step1} = \arg\min_{\boldsymbol{\theta}_1 \in \Theta_1} Q_{1:T}^{step1} = \frac{1}{2N} \sum_{t=1}^{T} \sum_{k=1}^{n_{y,t}} \left(y_{t,k} - g_k\left(\hat{\mathbf{x}}_t\left(\boldsymbol{\theta}_1\right); \boldsymbol{\theta}_1\right)\right)^2, \tag{31}$$

where $N \equiv \sum_{t=1}^T n_{y,t}$. Given standard regularity conditions, Andreasen & Christensen (2013) show

consistency and asymptotic normality of $\hat{\boldsymbol{\theta}}_1^{step1}$, i.e.

$$\sqrt{N}\left(\hat{\boldsymbol{\theta}}_1^{step1} - \boldsymbol{\theta}_1^o\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \left(\mathbf{A}_o^{\boldsymbol{\theta}_1}\right)^{-1} \mathbf{B}_o^{\boldsymbol{\theta}_1} \left(\mathbf{A}_o^{\boldsymbol{\theta}_1}\right)^{-1}\right), \tag{32}$$

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 tend to infinity, implying $N \to \infty$. The expected value of the average Hessian matrix $\mathbf{A}_o^{\boldsymbol{\theta}_1}$ may be estimated consistently by

$$\hat{\mathbf{A}}^{\boldsymbol{\theta}_1} = \frac{1}{N} \sum_{t=1}^{T} \sum_{k=1}^{n_{y,t}} \left(\hat{\boldsymbol{\Psi}}_{t,k}^{\boldsymbol{\theta}_1}\right) \left(\hat{\boldsymbol{\Psi}}_{t,k}^{\boldsymbol{\theta}_1}\right)', \tag{33}$$

where

$$\boldsymbol{\Psi}_{t,k}^{\boldsymbol{\theta}_1}\left(\boldsymbol{\theta}_1\right) \equiv \frac{\partial \hat{\mathbf{x}}_{2,t}'\left(\boldsymbol{\theta}_1\right)}{\partial \boldsymbol{\theta}_1} \frac{\partial g_k\left(\hat{\mathbf{x}}_{2,t}\left(\boldsymbol{\theta}_1\right); \boldsymbol{\theta}_1\right)}{\partial \mathbf{x}_{2,t}\left(\boldsymbol{\theta}_1\right)} + \frac{\partial g_k\left(\hat{\mathbf{x}}_{2,t}\left(\boldsymbol{\theta}_1\right); \boldsymbol{\theta}_1\right)}{\partial \boldsymbol{\theta}_1} \tag{34}$$

and $\hat{\boldsymbol{\Psi}}_{t,k}^{\boldsymbol{\theta}_1} \equiv \boldsymbol{\Psi}_{t,k}^{\boldsymbol{\theta}_1}\left(\hat{\boldsymbol{\theta}}_1^{step1}\right)$. 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, cross-section correlation, and autocorrelation in $v_{t,k}$. The most general specification considered in Andreasen & Christensen (2013) is given by

$$\begin{aligned} \hat{\mathbf{B}}^{\boldsymbol{\theta}_1} \quad = \quad & \frac{1}{N} \sum_{t=1}^{T} \sum_{k=1}^{n_{y,t}} \sum_{k_T=-w_T}^{w_T} \sum_{k_D=-w_D}^{w_D} \left(1 - \frac{|k_T|}{1+w_T}\right)\left(1 - \frac{|k_D|}{1+w_D}\right) \\ & \times \left(\hat{\boldsymbol{\Psi}}_{t,k}^{\boldsymbol{\theta}_1}\right)\left(\hat{\boldsymbol{\Psi}}_{t+k_T,j+k_D}^{\boldsymbol{\theta}_1}\right)' \widehat{\tilde{v}}_{t,k} \widehat{\tilde{v}}_{t+k_T,k+k_D} \end{aligned} \tag{35}$$

where 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.

4.1.2 The SR approach: Step 2

The parameters $\boldsymbol{\theta}_2$ in (29) are estimated 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. We follow Andreasen & Christensen (2013) and consider the moments

$$\mathbf{q}_T\left(\boldsymbol{\theta}_2\right) \equiv \frac{1}{T} \sum_{t=1}^{T} \mathbf{q}_t\left(\boldsymbol{\theta}_2\right) = \mathbf{0}, \tag{36}$$

where

$$\mathbf{q}_t(\boldsymbol{\theta}_2) \equiv \begin{bmatrix} \hat{\varepsilon}^{\mathbb{P}}_{t+1} \\ vec\left(\hat{\varepsilon}^{\mathbb{P}}_{t+1}\hat{\mathbf{x}}'_t - Cov\left(\mathbf{u}_{t+1},\mathbf{u}_t\right) + \mathbf{h}_{\mathbf{x}} Var\left(\mathbf{u}_t\right)\right) \\ vech\left(\begin{array}{c} \hat{\varepsilon}^{\mathbb{P}}_{t+1}\left(\hat{\varepsilon}^{\mathbb{P}}_{t+1}\right)' - Var\left(\hat{\varepsilon}^{\mathbb{P}}_{t+1}\right) - Var\left(\mathbf{u}_t\right) - \mathbf{h}_{\mathbf{x}} Var\left(\mathbf{u}_t\right)\mathbf{h}'_{\mathbf{x}} \\ +Cov\left(\mathbf{u}_{t+1},\mathbf{u}_t\right)\mathbf{h}'_{\mathbf{x}} + \mathbf{h}_{\mathbf{x}} Cov\left(\mathbf{u}_t,\mathbf{u}_{t+1}\right) \end{array}\right) \end{bmatrix} \tag{37}$$

and $\hat{\varepsilon}^{\mathbb{P}}_{t+1} \equiv \hat{\mathbf{x}}_{t+1} - \mathbf{h}_0 - \mathbf{h}_{\mathbf{x}}\hat{\mathbf{x}}_t$. Consistent estimators of $Var\left(\mathbf{u}_t\right)$, $Cov\left(\mathbf{u}_{t+1},\mathbf{u}_t\right)$, and $Cov\left(\mathbf{u}_t,\mathbf{u}_{t+1}\right)$ are provided in Andreasen & Christensen (2013) 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 considered models in the present paper have unrestricted $\mathbb{P}$ dynamics, and the moment conditions in (36) may then be solved in closed form. This solution is obtained by correcting all second moments for estimation uncertainty in $\{\hat{\mathbf{x}}_t\}_{t=1}^{T}$ and running the regression

$$\begin{aligned} \begin{bmatrix} \hat{\mathbf{h}}^{step2}_{\mathbf{x}} & \hat{\mathbf{h}}^{step2}_{0} \end{bmatrix} &= \left(\sum_{t=1}^{T-1}\begin{bmatrix} \hat{\mathbf{x}}_{t+1}\hat{\mathbf{x}}'_t - \widehat{Cov}\left(\mathbf{u}_{t+1},\mathbf{u}_t\right) & \hat{\mathbf{x}}_{t+1} \end{bmatrix}\right) \\ &\quad \times \left(\sum_{t=1}^{T-1}\begin{bmatrix} \hat{\mathbf{x}}_t\hat{\mathbf{x}}'_t - \widehat{Var}\left(\mathbf{u}_t\right) & \hat{\mathbf{x}}_t \\ \hat{\mathbf{x}}'_t & 1 \end{bmatrix}\right)^{-1} \end{aligned} \tag{38}$$

$$\begin{aligned} \widehat{Var}\left(\varepsilon^{\mathbb{P}}_{t+1}\right)^{step2} &= \frac{1}{T-1}\sum_{t=1}^{T-1}\left(\hat{\boldsymbol{\varepsilon}}_{t+1}\hat{\boldsymbol{\varepsilon}}'_{t+1} - \widehat{Var}\left(\mathbf{u}_t\right) - \hat{\mathbf{h}}_{\mathbf{x}}\widehat{Var}\left(\mathbf{u}_t\right)\hat{\mathbf{h}}'_{\mathbf{x}}\right. \\ &\qquad \left. + \widehat{Cov}\left(\mathbf{u}_{t+1},\mathbf{u}_t\right)\hat{\mathbf{h}}'_{\mathbf{x}} + \hat{\mathbf{h}}_{\mathbf{x}}\widehat{Cov}\left(\mathbf{u}_t,\mathbf{u}_{t+1}\right)\right), \end{aligned} \tag{39}$$

with $\hat{\boldsymbol{\Sigma}}^{step2}$ obtained from a Cholesky decomposition of $\widehat{Var}\left(\varepsilon^{\mathbb{P}}_{t+1}\right)^{step2}$. When T tends to infinity, Andreasen & Christensen (2013) show that the asymptotic distribution of $\boldsymbol{\theta}_2$ is

$$\sqrt{T}\left(\boldsymbol{\theta}^{step2}_2 - \boldsymbol{\theta}^{o}_2\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \left(\mathbf{R}\mathbf{S}^{-1}\mathbf{R}'\right)^{-1}\right), \tag{40}$$

where $\mathbf{R} \equiv \frac{\partial \mathbf{q}_T(\boldsymbol{\theta}^o_2)'}{\partial \boldsymbol{\theta}_2}$ and $\mathbf{S} \equiv \sum_{\nu=-\infty}^{\infty} E\left[\mathbf{q}_t\left(\boldsymbol{\theta}_2\right)\mathbf{q}_{t-\nu}\left(\boldsymbol{\theta}_2\right)'\right]$. We estimate $\mathbf{R}$ using numerical differentiation and $\mathbf{S}$ by the Newey-West estimator.

4.1.3 The SR approach: Step 3

The elements in $\boldsymbol{\Sigma}$ appear in $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ and are therefore estimated in both the first and second estimation step. These estimates may be combined in an optimal way to reduce any potential efficiency loss from the use of sequential identification as shown by Andreasen & Christensen (2013). We generally find that $\hat{\boldsymbol{\Sigma}}^{step1}$ is estimated very inaccurately compared to $\hat{\boldsymbol{\Sigma}}^{step2}$, meaning that the time series estimate $\hat{\boldsymbol{\Sigma}}^{step2}$ cannot be improved by adding cross-section information from $\hat{\boldsymbol{\Sigma}}^{step1}$.[15] Hence, the adopted estimate of $\boldsymbol{\Sigma}$ after the first two steps is given by $\hat{\boldsymbol{\Sigma}}^{step2}$.

Based on the more accurate estimate of $\boldsymbol{\Sigma}$ from the second step, it is natural to re-estimate the remaining elements in $\boldsymbol{\theta}_1$ when conditioned on $\hat{\boldsymbol{\Sigma}}^{step2}$. That is

$$\hat{\boldsymbol{\theta}}_{11}^{step3} = \arg \min_{\boldsymbol{\theta}_{11} \in \Theta_{11}} Q_{1:T}^{step3} = \frac{1}{2N} \sum_{t=1}^{T} \sum_{k=1}^{n_{y,t}} \left(y_{t,k} - g_k \left(\hat{\mathbf{x}}_t \left(\boldsymbol{\theta}_{11}, \hat{\boldsymbol{\Sigma}}^{step2} \right); \boldsymbol{\theta}_{11}, \hat{\boldsymbol{\Sigma}}^{step2} \right) \right)^2, \tag{41}$$

where $\boldsymbol{\theta}_{11}$ refers to all elements in $\boldsymbol{\theta}_1$ except $\boldsymbol{\Sigma}$. Andreasen & Christensen (2013) show consistency and asymptotic normality of $\hat{\boldsymbol{\theta}}_{11}^{step3}$ with

$$\widehat{Var}\left(\hat{\boldsymbol{\theta}}_{11}^{step3} \right) = \frac{\hat{\mathbf{V}}_{\boldsymbol{\theta}_{11}}^{step3} \left(\hat{\boldsymbol{\Sigma}}^{step2} \right)}{N} + \hat{\mathbf{K}} \widehat{Var} \left(\hat{\boldsymbol{\Sigma}}^{step2} \right) \hat{\mathbf{K}}'. \tag{42}$$

The first term $\hat{\mathbf{V}}_{\boldsymbol{\theta}_{11}}^{step3} \left(\hat{\boldsymbol{\Sigma}}^{step2} \right) / N$ is given by (32) when used on the subset of $\boldsymbol{\theta}_1$ corresponding to $\boldsymbol{\theta}_{11}$. The second term in (42) corrects for estimation uncertainty in $\hat{\boldsymbol{\Sigma}}^{step2}$ with $\mathbf{K} \equiv \partial \hat{\boldsymbol{\theta}}_{11}^{step3} (\boldsymbol{\Sigma}) / \partial vech (\boldsymbol{\Sigma})'$. We estimate $\mathbf{K}$ as suggested in Andreasen & Christensen (2013) and refer to their paper for further details.

Given the estimated pricing factors $\left\{ \hat{\mathbf{x}}_t \left(\boldsymbol{\theta}_{11}^{step3}, \hat{\boldsymbol{\Sigma}}^{step2} \right) \right\}_{t=1}^{T}$ from (41), we finally update our estimates of $\boldsymbol{\theta}_2$ using (38) and (39). These estimates are denoted $\hat{\mathbf{h}}_0^{step3}$, $\hat{\mathbf{h}}_{\mathbf{x}}^{step3}$, and $\hat{\boldsymbol{\Sigma}}^{step3}$ and we refer to them as $\boldsymbol{\theta}_2^{step3}$.

[15] Adopting the notation in Andreasen & Christensen (2013), we generally find $\boldsymbol{\Lambda} \approx \mathbf{0}$. The only exceptions are the two-factor QTSM and the four-factor ATSM where $\boldsymbol{\Lambda} \neq \mathbf{0}$. However, imposing $\boldsymbol{\Lambda} = \mathbf{0}$ gives nearly identical estimates to those obtained when $\boldsymbol{\Lambda}$ is determined optimally in both cases. To ensure similarity across all of the models considered, we therefore also let $\boldsymbol{\Lambda} = \mathbf{0}$ for the two-factor QTSM and the four-factor ATSM.

5 Empirical results: In-sample performance

This section estimates the benchmark ATSM, the QTSM, and the shadow rate model on post-war US data. For comparability with much of the existing literature on the ZLB, we first study models with two pricing factors before exploring the performance of three-factor models. We find several short-comings of these two- and three-factor models and we therefore also estimate models with four pricing factors. Our analysis is structured as follows. Section 5.1 presents the data, and our model estimates are discussed in Section 5.2. We examine the accuracy of the bias-adjusted perturbation approximation for the shadow rate model in Section 5.3. The two subsequent sections explore how well the models match various aspects of bond yields.

5.1 Data

We use start-of-month nominal bond yields in the US from July 1961 to May 2013 as provided by Gürkaynak, Sack & Wright (2007). The SR approach is constructed for a setting where many observables are available each time period, and we therefore include more bond yields for the estimation than typically used in the literature. Simulation results by Andreasen & Christensen (2013) suggest that about 15 bond yields are sufficient and that any efficiency loss of the SR approach compared to Maximum Likelihood may be small with 25 bond yields. Given our interest in the 10-year term structure, we include all bond yields in the 0.25-10 year maturity range when sampled at a quarterly frequency. That is, whenever possible, we include 40 bond yields having the maturities $\{0.25, 0.50, ..., 9.75, 10\}$.[16] Due to a lack of long-term Treasury notes before September 1971, bonds yields in the 7-10 year maturity range are not available before this date. We address this problem by explicitly accounting for missing values in the SR approach.

5.2 Model estimates

The estimation results for the two-factor models are reported in Table 1. The benchmark ATSM displays the usual properties with stationary and highly persistent factors under both the $\mathbb{Q}$ and $\mathbb{P}$ measure as $diag(\Phi) > \mathbf{0}$ and $\mathbf{h_x}$ has eigenvalues 0.9913 and 0.9646. The same properties hold for the

[16] These bond yields are computed using the estimated parametric form for the yield curve in Gürkaynak et al. (2007). Ongoing work explores the robustness of our results when using non-parametric estimation methods to extract the yield curves from coupon bonds.

pricing factors in the QTSM, where $\mathbf{\Psi}$ enforces the ZLB by having strictly positive eigenvalues (0.0359 and 1.9641). We also find that the two pricing factors in the QTSM are positively correlated (0.44) whereas they display negative correlation in the ATSM (-0.65). The estimates for the shadow rate with two pricing factors are generally very similar to those for the two-factor benchmark ATSM, but we also find some differences. For instance, the pricing factors in the shadow rate model are weakly positively correlated (0.15) and the conditional volatility of the second pricing factor is $\Sigma_{22} = 6.45 \times 10^{-4}$ and hence larger than the corresponding estimate in the ATSM of $\Sigma_{22} = 3.97 \times 10^{-4}$.[17]

< Table 1 about here >

Turning to three- and four-factor models in Tables 2 and 3, respectively, all models imply stationary and highly persistent pricing factors under both the $\mathbb{Q}$ and $\mathbb{P}$ measure. Our estimates imply that the ZLB in the three-factor QTSM is enforced by letting $\mathbf{\Psi}$ be positive definite, whereas $\mathbf{\Psi}$ is found to be positive semi-definite in the four-factor QTSM as $\mathbf{\Psi}$ has one eigenvalue equal to zero.[18] We also note that several estimates for the three- and four-factor benchmark ATSM differ substantially from the corresponding estimates in the shadow rate models. This finding indicates that one should be cautious of directly using parameters from the benchmark ATSM in the shadow rate model to explore the implications of the ZLB.

< Table 2 and 3 about here >

5.3 Calibration and accuracy of the perturbation approximation

For each of the considered shadow rate models, we calibrate the scaling of the bias correction from simulated time series of $2,000$ observations using preliminary estimates with $\gamma_{0,k} = 0$ and $\gamma_{1,k} = 1$ for $k = 1, 2, ..., K$. The 'true' solution to bond yields is here obtained by the Monte Carlo integration with $10,000$ draws. The calibrated values of $\left\{\gamma^*_{0,k}, \gamma^*_{1,k}\right\}_{k=1}^{K}$ for the shadow rate model with two, three, and four pricing factors are displayed in the left column of Figure 2. We generally find that the $\gamma^*_{1,k}$ parameters start at one and then decrease with maturity. Most calibrated intercepts $\gamma^*_{0,k}$ are

[17] For the robust standard errors in Table 1 we use a bandwidth of 5 in the Newey-West estimate of $\mathbf{S}$ in (40) and $w_D = w_T = 10$ in (35). These standard errors are broadly similar when using slightly smaller or larger bandwidths.

[18] This implies that the estimates are on the boundary of the domain for the parameters and that the standard errors in Table 3 only serve as an approximation. Ongoing work aims to compute more accurate standard errors using a bootstrap procedure.

very close to zero, except for bond yields with long maturities in the four-factor model. Given these calibrated values of $\left\{\gamma_{0,k}^*, \gamma_{1,k}^*\right\}_{k=1}^K$, the three shadow rate models are then re-estimated to obtain the results reported in Tables 1-3.

We next explore the accuracy of the fourth-order perturbation approximation when using the estimated pricing factors and model parameters. The charts in the right column of Figure 2 show the root mean squared pricing errors (RMSEs) for bond yields by maturity, where the 'true' solution to bond yields is computed using Monte Carlo integration with $10,000$ draws. The RMSEs without the bias correction are marked with black lines and found to be between 15 and 30 annualized basis points for all the shadow rate models considered. Adding the unscaled bias correction ($\gamma_{0,k} = 0$ and $\gamma_{1,k} = 1$) to the fourth-order perturbation approximation substantially reduces the RMSEs as they fall to between 5 and 15 annualized basis points. Using the optimal scaling of the bias correction $\left\{\gamma_{0,k}^*, \gamma_{1,k}^*\right\}_{k=1}^K$ further reduces the RMSEs to about 5 annualized basis points, except for bond yields with short maturities in the two-factor models where the RMSEs are around 10 annualized basis points.

< Figure 2 about here >

Further evidence on the satisfying performance of the bias-adjusted fourth-order perturbation approximation using $\left\{\gamma_{0,k}^*, \gamma_{1,k}^*\right\}_{k=1}^K$ is provided in Figure 3, showing bond yields and pricing errors for a selected number of maturities in the estimated three-factor model.[19] This figure shows that the RMSEs are low when interest rates are far from zero but also when they approach the ZLB during 2002-2004 and after 2008. As a result, the maximal pricing errors rarely exceed ± 10 annualized basis points and this shows that our approximation achieves high accuracy for bond yields throughout the sample.

< Figure 3 about here >

Another notable advantage of the suggested perturbation approximation is its computational speed. It only takes 0.6 seconds to obtain all bond yields in the two-factor model. The computational requirement only increase gradually with the number of pricing factors, as we are able to solve the three- and four-factor shadow rate models in just 1.7 and 3.7 seconds, respectively.[20]

[19] Similar plots for the two- and four-factor shadow rate models are available on request.

[20] The computations are done in MATLAB 2012a using an Intel(R) Core(TM) i5-320M CPU 2 2.50Ghz.

These findings lead us to the conclusion that the bias-adjusted perturbation approximation is highly accurate when applied to bond yields in shadow rate models. We also find that this method is computationally very appealling because its execution time remains highly tractable even in models with four pricing factors.

5.4 Goodness of fit for bond yields

We next study the in-sample fit by looking at the objective functions for the models considered. The first part of Table 4 reports $\tilde{Q}_{1:T}^{step1} \equiv 100\sqrt{Q_{1:T}^{step1}/2}$ which measures the standard deviation of all residuals in annualized basis points for the first step in the SR approach. With two pricing factors, the QTSM clearly provides the best fit with $\tilde{Q}_{1:T}^{step1} = 10.15$, whereas the shadow rate model and the benchmark ATSM have $\tilde{Q}_{1:T}^{step1} = 11.37$ and $\tilde{Q}_{1:T}^{step1} = 11.50$, respectively. The QTSM also delivers the best fit with three pricing factors where $\tilde{Q}_{1:T}^{step1} = 4.81$, whereas the worst fit is seen in the shadow rate model $\left(\tilde{Q}_{1:T}^{step1} = 5.34\right)$. Turning to four-factor models, we find that the QTSM and the shadow rate model obtain the same fit of bond yields with $\tilde{Q}_{1:T}^{step1} = 2.03$, whereas the benchmark ATSM has $\tilde{Q}_{1:T}^{step1} = 2.40$. It is important to note that these results are obtained without using the law of motion for the pricing factors under $\mathbb{P}$, and our results therefore hold for *any* functional form of the market price of risk $\mathbf{f}(\mathbf{x}_t)$.

The second part of Table 4 shows the scaled objective functions from the third step in the SR approach, i.e. $\tilde{Q}_{1:T}^{step3} \equiv 100\sqrt{Q_{1:T}^{step3}/2}$, where $\mathbf{\Sigma}$ is estimated from the time-series dimension instead of the cross-section dimension as in the first step. Here, $\tilde{Q}_{1:T}^{step3}$ is only marginally larger than $\tilde{Q}_{1:T}^{step1}$ for all models, meaning that the in-sample fit of bond yields is almost unaffected by the alternative estimate of $\mathbf{\Sigma}$. It is therefore reasonable to believe that the dependence on the $\mathbb{P}$ dynamics through $\mathbf{\Sigma}$ is minimal in our case and that results in the third step of the SR approach largely remain robust to the chosen functional form of $\mathbf{f}(\mathbf{x}_t)$.

< Table 4 about here >

A more careful examination of the in-sample fit is provided in Figure 4, where charts in the first column show recursively computed objective functions using the estimates in Tables 1 to 3, i.e. $\left\{\tilde{Q}_{1:t}^{step3}\right\}_{t=1}^{T}$. These charts suggest that the in-sample fit generally deteriorates during the 1970's and improve afterwards, and that the relative performance of the three models is fairly stable throughout

the sample. The same type of plots are provided in the second column in Figure 4 but only when the objective functions are computed from January 1998 to May 2013 to explore the fit when bond yields are close to the ZLB. We emphasize that these plots are computed using the estimates from the full sample, i.e. those in Tables 1 to 3. In this time period, the QTSM clearly delivers the best in-sample fit regardless of the number of pricing factors, whereas the shadow rate model only outperforms the benchmark ATSM with four pricing factors.

Another way to explore the in-sample fit of bond yields is provided in the final column in Figure 4, showing the standard deviation of the residuals by maturity, i.e. $\sigma_k = 100\sqrt{\frac{1}{T}\sum_{t=1}^{T} v_{t,k}^2}$ for $k = 1, 2, ..., K$ with expressed σ_k in annualized basis points. All two-factor models clearly struggle to match bond yields at the short and long end of the term structure as $\sigma_{0.25y} = 40$ and $\sigma_{10y} = 15$. Including a third pricing factor substantially reduces the residuals at the long end of the term structure ($\sigma_{10y} = 6$), but the residuals remain high at the short end with $\sigma_{0.25y} > 15$ in all models. One way to address this shortcoming of three-factor models is to include a fourth pricing factor as the standard deviation of the residuals at the short end then fall below 7 annualized basis points. In other words, a fourth pricing factors is required to properly match short-term bond yields.

< Figure 4 about here >

Based on these findings we conclude that a quadratic policy rate is the best way to enforce the ZLB in two- and three-factor models when measured by the in-sample fit. In four-factor models, a quadratic policy rate and a shadow rate specification deliver broadly the same fit of bond yields, and we are therefore unable to rank the two mechanisms for enforcing the ZLB in this case.

5.5 Matching key moments for bond yields

The QTSM with n_x pricing factors has $n_x (n_x + 1)/2 - 1$ additional parameters compared to the two other models, and the quadratic model is therefore likely to do well when measured by the in-sample fit. Issues related to overfitting may be partially addressed by also evaluating the considered models by their ability to match moments not directly included in the estimation. The first set of moments we explore are the unconditional first and second moments of bond yields in Figure 5. All models match the upward sloping unconditional yield curve fairly well, except for the three-factor QTSM. We also note that most models provide relatively low estimates of the unconditional mean in bond yields.

The decreasing pattern in the unconditional volatility of bond yields is broadly matched by nearly all models, except for the three-factor QTSM.

< Figure 5 about here >

Another set of moments typically considered to assess the performance of DTSMs are derived in Dai & Singleton (2002). They start by considering the regressions for excess returns

$$y_{t+1,k-1} - y_{t,k} = \delta_k + \phi_k \frac{y_{t,k} - r_t}{k-1} + u_{t,k} \tag{43}$$

where $u_{t,k}$ denotes the residual. The expectation hypothesis implies $\delta_k = 0$ and $\phi_k = 1$ for $k = 1, 2, ..., K$, but this prediction is clearly rejected for US data, as ϕ_k is negative and decreases with maturity. The ability of DTSMs to reproduce the observed pattern in $\{\phi_k\}_{k=1}^{K}$ is referred to as LPY(i) by Dai & Singleton (2002) and tests whether the models are able to capture the $\mathbb{P}$ dynamics of bond yields. The charts in the first column of Figure 6 examine the performance of the considered DTSMs along this dimension, where the corresponding model-moments of ϕ_k are computed from simulated time series of 500,000 observations. We first note that the benchmark ATSM does extremely well along this dimension, even with just two pricing factors. The QTSMs and the shadow rate models struggle to match the downward sloping pattern in ϕ_k with two and three pricing factors. Much better performance is obtained in the four-factor models, as they match the negative and decreasing pattern in ϕ_k.

The second set of moment conditions studied in Dai & Singleton (2002) are derived by modifying the regressions in (43) as follows

$$y_{t+1,k-1} - y_{t,k} - \frac{e_{t,k}}{k-1} = \delta_k^{\mathbb{Q}} + \phi_k^{\mathbb{Q}} \frac{y_{t,k} - r_t}{k-1} + u_{t,k}^{\mathbb{Q}} \tag{44}$$

where $e_{t,k} \equiv \mathbb{E}_t^{\mathbb{P}} \left[\log\left(P_{t+1,k-1}\right) / \log\left(P_{t,k}\right) - r_t \right]$ denotes the excess holding period return and $u_{t,k}^{\mathbb{Q}}$ is the residual. If the risk premia adjustment in $e_{t,k}$ is correctly specified, then we recover the expectation hypothesis with $\phi_k^{\mathbb{Q}} = 1$ for all k as shown by Dai & Singleton (2002). The ability of DTSMs to generate this implication is referred to as LPY(ii) by Dai & Singleton (2002) and tests if the models are correctly specified under the $\mathbb{Q}$ measure. Given that all of the models considered in the present paper only differ in their $\mathbb{Q}$ distribution, we view LPY(ii) as a very informative test to discriminate

between the models. Charts in the second column of Figure 6 study the ability of the models to reproduce $\phi_k^{\mathbb{Q}} = 1$ for all k. We interestingly find that both the benchmark ATSM and the QTSM struggle with two and three pricing factors as the regression loadings differ substantially from one. Slightly better performance is observed for these models with four pricing factors, but substantial deviations remain. In sharp contrast to the performance of these models, all the shadow rate models do extremely well along this dimension as they nearly reproduce a regression coefficient of one. This suggests that the shadow rate model is much better at matching the $\mathbb{Q}$ dynamics than the benchmark ATSM and the QTSM.[21]

< Figure 6 about here >

We finally plot the 10-year term premium in Figure 7 for all estimated models. For the two-factor models, we observe some differences in reported term premia which is expected given the relative large residuals in these models. For three- and four-factor models, we observe that the benchmark ATSM and the shadow rate model provide broadly similar estimates of term premia, except after 2008 where bond yields approach the ZLB. We also note that the estimated term premia in the three- and four-factor QTSMs are somewhat higher than those obtained from the three- and four-factor shadow rate models.

< Figure 7 about here >

These findings lead us to the following conclusions. In the QTSM, four pricing factors are needed to match the unconditional first and second moments whereas only three factors are required in the shadow rate model. We also find that both models rely on four pricing factors to match the dynamics of bond yields under the $\mathbb{P}$ measure and hence pass the LPY(i) test. Importantly, only shadow rate models pass the LPY(ii) test, suggesting that the $\mathbb{Q}$ dynamics are best captured by a shadow rate specification.

[21] Ongoing work explores if the inability of the benchmark ATSM and the QTSM to pass the LPY(ii) test is caused by the low interest rates after 2000.

6 Empirical results: performance out-of-sample

Another commonly used method to correct for potential overfitting is to conduct a forecasting exercise out-of-sample. This is the topic of the current section where we evaluate the forecasting performance of the benchmark ATSM, the QTSM, and the shadow rate model with three pricing factors from January 2009 to May 2013.[22] We focus on this period because the Federal Open Market Committee has set a target range of 0-0.25% for the effective Federal Funds Rate, meaning that policy rate has been at its effective ZLB. The forecasting study is carried out by estimating all three-factor models recursively every month to forecasts bond yields up to 12 months ahead. Given that the last 12 months of data is reserved for evaluating the final forecasts, each model is estimated a total of 41 times which is easily done due to the computational efficiency of the SR approach relative to commonly used estimation alternatives.

Figure 8 shows root mean squared prediction error (RMSPE) statistics from the three models at forecast horizons of 1, 3, 6, and 12 months, alongside RMSPE statistics under the assumption that bond yields of all maturities follow random walks. Three main results emerge. First, the one-month ahead forecasting performance of all models is similar, with none of the models outperforming a random walk for any maturity. Second, the forecasting performance of the benchmark ATSM for short-term bond yields becomes substantially worse at the 12-month horizon, which is qualitatively in line with the findings by Pooter, Ravazzolo & van Dijk (2010). In contrast, the forecasting performance of the QTSM and, in particular, the shadow rate model does not deteriorate with the forecast horizon for short-term bond yields. At a 12-month horizon the shadow rate model even performs significantly better than a random walk at forecasting bond yields with 3-month and 1-year maturities.[23] The reason why the QTSM and shadow rate model perform better in this respect is likely to be because they can generate a lower degree of mean reversion in bond yields than the benchmark ATSM when yields are close to the ZLB. For example, in May 2012 the 12-month ahead forecasts of the 3-month rate are 0.60% in the ATSM, 0.10% in the QTSM, and 0.16% in the shadow rate model.

Third, the relative performance of the QTSM is worse for long-term bond yields, particularly at longer forecasting horizons. It seems plausible that for bond yields not currently close to the lower

[22] Ongoing work explores the forecasting performance of the considered two- and four-factor models and on longer time periods.

[23] Outperformance at the 5% significance level is based on the Harvey, Leybourne & Newbold (1997) small sample modification of the Diebold & Mariano (1995) test.

bound, the need to estimate more parameters in the QTSM outweighs any benefits from including quadratic terms in the policy rate. Even the shadow rate model does not come close to beating a random walk for long-term bond yields, which seems broadly consistent with the findings in Pooter et al. (2010).

< Figure 8 about here >

In conclusion, the shadow rate model appears to offer the best forecasting performance in terms of RMSPE, where it out-performs the benchmark ATSM at short maturity and the QTSM at longer maturities.

7 Conclusion

This paper studies whether DTSMs for US nominal bond yields should enforce the ZLB through a quadratic policy rate or a shadow rate. The question is addressed by estimating QTSMs and shadow rate models with two, three, and four latent pricing factors using the SR approach. Bond yields in the shadow rate models are efficiently computed by a fourth-order perturbation approximation, extended with a novel bias correction for interest rates under perfect foresight. When measured in terms of in-sample fit, we generally find that QTSMs outperform shadow rate models, except with four pricing factors where both models deliver the same fit of bond yields. However, some of the good performance for the QTSMs are likely related to some degree of overfitting, as these models perform worse than shadow rate models on moments not directly incorporated in the estimation. This includes the ability of the models to match unconditional first and second moments for bond yields, and importantly pass the LPY(i) and LPY(ii) tests of Dai & Singleton (2002). Furthermore, the shadow rate models also seem to outperform QTSMs in a proper out-of-sample forecasting exercise. Our work therefore suggests that DTSMs for US bond yields should enforce the ZLB by adopting a shadow rate specification instead of a quadratic policy rate, possibly using four pricing factors.

A Computing bond prices by the perturbation method

A.1 Notation

We let the indices α and γ relate to elements of $\mathbf{x}_t$, while ϕ corresponds to elements of $\boldsymbol{\epsilon}_t$. Subscripts on these indices capture the sequence in which derivatives are taken. For example, α_1 corresponds to the first time a function is differentiated with respect to $\mathbf{x}_t$, while α_2 is used when differentiating with respect to $\mathbf{x}_t$ the second time, and so on. As typically done in the literature, we adopt the tensor notation. Hence, $[p^k_{\mathbf{x}}]_{\gamma_1}$ denotes the γ_1-th element of the $1 \times n_x$ vector of derivatives of p^k with respect to $\mathbf{x}$. Similarly, the derivative of $\mathbf{h}$ with respect to $\mathbf{x}$ is an $n_x \times n_x$ matrix and $[\mathbf{h}_{\mathbf{x}}]^{\gamma_1}_{\alpha_1}$ is the element of this matrix located at the intersection of the γ_1-th row and the α_1-th column. This notation facilitates writing summation as

$$\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_1} [\mathbf{h}_{\mathbf{x}}]^{\gamma_1}_{\alpha_1} = \textstyle\sum_{\gamma_1=1}^{n_x} \dfrac{\partial\, p^{k-1}}{\partial\, \mathbf{x}_{\gamma_1}} \dfrac{\partial\, \mathbf{h}^{\gamma_1}}{\partial \mathbf{x}_{\alpha_1}},$$

and

$$\left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_1\gamma_2} [\mathbf{h}_{\mathbf{x}}]^{\gamma_2}_{\alpha_2} [\mathbf{h}_{\mathbf{x}}]^{\gamma_1}_{\alpha_1} = \textstyle\sum_{\gamma_1=1}^{n_x} \sum_{\gamma_2=1}^{n_x} \dfrac{\partial^2\, p^{k-1}}{\partial\, \mathbf{x}_{\gamma_1} \partial\, \mathbf{x}_{\gamma_2}} \dfrac{\partial\, \mathbf{h}^{\gamma_2}}{\partial \mathbf{x}_{\alpha_2}} \dfrac{\partial\, \mathbf{h}^{\gamma_1}}{\partial \mathbf{x}_{\alpha_1}},$$

where, for instance, $\mathbf{h}^{\gamma_1}$ denotes the γ_1-th function of mapping $\mathbf{h}$ and $\mathbf{x}_{\alpha_1}$ is the α_1-th element of vector $\mathbf{x}$. Here, and throughout, we omit the function arguments as all functions are evaluated at the deterministic steady state.

The recursions for bond price derivatives that differ from zero are stated below. Here, we rely on derivatives of the $\mathbf{h}$-function and the M-function. We adopt the notation that $\mathbf{h}_{\mathbf{x}^n}$ refers to derivatives of the $\mathbf{h}$-function with respect to $\mathbf{x}_t$ taken n times. Similarly, $M_{\mathbf{z}^j\mathbf{y}^n}$ refers to the derivative of M with respect to $\mathbf{z}$ taken j times and with respect to $\mathbf{y}$ taken n times, for $\mathbf{z}, \mathbf{y} \in \{\mathbf{x}_t, \mathbf{x}_{t+1}\}$. All these derivatives are evaluated at the deterministic steady state. Given that the $\mathbf{h}$-function and the M-function are known, all required derivatives with respect to these functions follow by simple differentiation. In evaluating the expressions below, it is useful to recall the limits for the adopted indices

$$\alpha_1, \alpha_2, \alpha_3, \alpha_4 = 1, 2, ...n_x$$

$$\gamma_1, \gamma_2, \gamma_3, \gamma_4 = 1, 2, ...n_x$$

$$\phi_1, \phi_2, \phi_3, \phi_4 = 1, 2, ..., n_\epsilon$$

where n_x denotes the number of pricing factors in $\mathbf{x}_t$ and n_ϵ refers to the number of innovations. Finally, the validity of the stated bond price derivatives have been verified in relation to *Dynare++* on a number of test examples.

A.2 Bond price derivatives: first-order terms

$$\left[p^k_{\mathbf{x}}\right]_{\alpha_1} = \left[p^1_{\mathbf{x}}\right]_{\alpha_1} + \left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_1} [\mathbf{h}_{\mathbf{x}}]^{\gamma_1}_{\alpha_1}$$

A.3 Bond price derivatives: second-order terms

$$\left[p^k_{\mathbf{xx}}\right]_{\alpha_1\alpha_2} = \left[p^1_{\mathbf{xx}}\right]_{\alpha_1\alpha_2} + \left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_1\gamma_2} [\mathbf{h}_{\mathbf{x}}]^{\gamma_2}_{\alpha_2} [\mathbf{h}_{\mathbf{x}}]^{\gamma_1}_{\alpha_1} + \left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_1} [\mathbf{h}_{\mathbf{xx}}]^{\gamma_1}_{\alpha_1\alpha_2}$$

$$
\begin{aligned}
\left[p^k_{\sigma\sigma}\right] &= \left[p^1_{\sigma\sigma}\right] + \left[p^{k-1}_{\sigma\sigma}\right] + \left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_1\gamma_2} [\boldsymbol{\eta}]^{\gamma_2}_{\phi_2} [\boldsymbol{\eta}]^{\gamma_1}_{\phi_1} [\mathbf{I}]^{\phi_1}_{\phi_2} \\
&\quad + \left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_2} [\boldsymbol{\eta}]^{\gamma_2}_{\phi_2} \left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_1} [\boldsymbol{\eta}]^{\gamma_1}_{\phi_1} [\mathbf{I}]^{\phi_1}_{\phi_2} + 2M^{-1} \left[M_{\mathbf{x}_{t+1}}\right]_{\gamma_1} [\boldsymbol{\eta}]^{\gamma_1}_{\phi_1} \left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_2} [\boldsymbol{\eta}]^{\gamma_2}_{\phi_2} [\mathbf{I}]^{\phi_2}_{\phi_1}
\end{aligned}
$$

A.4 Bond price derivatives: third-order terms

$$
\begin{aligned}
\left[p^k_{\mathbf{xxx}}\right]_{\alpha_1\alpha_2\alpha_3} &= \left[p^1_{\mathbf{xxx}}\right]_{\alpha_1\alpha_2\alpha_3} + \left[p^{k-1}_{\mathbf{xxx}}\right]_{\gamma_1\gamma_2\gamma_3} [\mathbf{h_x}]^{\gamma_3}_{\alpha_3} [\mathbf{h_x}]^{\gamma_2}_{\alpha_2} [\mathbf{h_x}]^{\gamma_1}_{\alpha_1} \\
&\quad + \left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_1\gamma_2} [\mathbf{h_{xx}}]^{\gamma_2}_{\alpha_2\alpha_3} [\mathbf{h_x}]^{\gamma_1}_{\alpha_1} + \left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_1\gamma_2} [\mathbf{h_x}]^{\gamma_2}_{\alpha_2} [\mathbf{h_{xx}}]^{\gamma_1}_{\alpha_1\alpha_3} \\
&\quad + \left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_1\gamma_3} [\mathbf{h_x}]^{\gamma_3}_{\alpha_3} [\mathbf{h_{xx}}]^{\gamma_1}_{\alpha_1\alpha_2} + \left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_1} [\mathbf{h_{xxx}}]^{\gamma_1}_{\alpha_1\alpha_2\alpha_3}
\end{aligned}
$$

$$
\begin{aligned}
\left[p^k_{\sigma\sigma\mathbf{x}}\right]_{\alpha_3} &= \left[p^1_{\sigma\sigma\mathbf{x}}\right]_{\alpha_3} - 2\left[p^1_{\mathbf{x}}\right]_{\alpha_3} \left[M_{\mathbf{x}_{t+1}}\right]_{\gamma_1} [\boldsymbol{\eta}]^{\gamma_1}_{\phi_1} M^{-1} \left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_2} [\boldsymbol{\eta}]^{\gamma_2}_{\phi_2} [\mathbf{I}]^{\phi_2}_{\phi_1} \\
&\quad +2\left(\left[M_{\mathbf{x}_{t+1}\mathbf{x}_{t+1}}\right]_{\gamma_1\gamma_3} [\mathbf{h_x}]^{\gamma_3}_{\alpha_3} + \left[M_{\mathbf{x}_{t+1}\mathbf{x}_t}\right]_{\gamma_1\alpha_3}\right) [\boldsymbol{\eta}]^{\gamma_1}_{\phi_1} M^{-1} \left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_2} [\boldsymbol{\eta}]^{\gamma_2}_{\phi_2} [\mathbf{I}]^{\phi_2}_{\phi_1} \\
&\quad +2\left[M_{\mathbf{x}_{t+1}}\right]_{\gamma_1} [\boldsymbol{\eta}]^{\gamma_1}_{\phi_1} M^{-1} \left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_2\gamma_3} [\mathbf{h_x}]^{\gamma_3}_{\alpha_3} [\boldsymbol{\eta}]^{\gamma_2}_{\phi_2} [\mathbf{I}]^{\phi_2}_{\phi_1} \\
&\quad + \left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_2\gamma_3} [\mathbf{h_x}]^{\gamma_3}_{\alpha_3} [\boldsymbol{\eta}]^{\gamma_2}_{\phi_2} \left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_1} [\boldsymbol{\eta}]^{\gamma_1}_{\phi_1} [\mathbf{I}]^{\phi_1}_{\phi_2} \\
&\quad + \left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_2} [\boldsymbol{\eta}]^{\gamma_2}_{\phi_2} \left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_1\gamma_3} [\mathbf{h_x}]^{\gamma_3}_{\alpha_3} [\boldsymbol{\eta}]^{\gamma_1}_{\phi_1} [\mathbf{I}]^{\phi_1}_{\phi_2} \\
&\quad + \left[p^{k-1}_{\mathbf{xxx}}\right]_{\gamma_1\gamma_2\gamma_3} [\mathbf{h_x}]^{\gamma_3}_{\alpha_3} [\boldsymbol{\eta}]^{\gamma_2}_{\phi_2} [\boldsymbol{\eta}]^{\gamma_1}_{\phi_1} [\mathbf{I}]^{\phi_1}_{\phi_2} + \left[p^{k-1}_{\sigma\sigma\mathbf{x}}\right]_{\gamma_3} [\mathbf{h_x}]^{\gamma_3}_{\alpha_3}
\end{aligned}
$$

A.5 Bond price derivatives: fourth-order terms

$$
\begin{aligned}
\left[p^k_{\mathbf{xxxx}}\right]_{\alpha_1\alpha_2\alpha_3\alpha_4} &= \left[p^1_{\mathbf{xxxx}}\right]_{\alpha_1\alpha_2\alpha_3\alpha_4} + \left[p^{k-1}_{\mathbf{xxxx}}\right]_{\gamma_1\gamma_2\gamma_3\gamma_4} [\mathbf{h_x}]^{\gamma_4}_{\alpha_4} [\mathbf{h_x}]^{\gamma_3}_{\alpha_3} [\mathbf{h_x}]^{\gamma_2}_{\alpha_2} [\mathbf{h_x}]^{\gamma_1}_{\alpha_1} \\
&\quad + \left[p^{k-1}_{\mathbf{xxx}}\right]_{\gamma_1\gamma_2\gamma_3} \left([\mathbf{h_{xx}}]^{\gamma_3}_{\alpha_3\alpha_4} [\mathbf{h_x}]^{\gamma_2}_{\alpha_2} [\mathbf{h_x}]^{\gamma_1}_{\alpha_1} + [\mathbf{h_x}]^{\gamma_3}_{\alpha_3} [\mathbf{h_{xx}}]^{\gamma_2}_{\alpha_2\alpha_4} [\mathbf{h_x}]^{\gamma_1}_{\alpha_1}\right) \\
&\quad + \left[p^{k-1}_{\mathbf{xxx}}\right]_{\gamma_1\gamma_2\gamma_3} [\mathbf{h_x}]^{\gamma_3}_{\alpha_3} [\mathbf{h_x}]^{\gamma_2}_{\alpha_2} [\mathbf{h_{xx}}]^{\gamma_1}_{\alpha_1\alpha_4} \\
&\quad + \left[p^{k-1}_{\mathbf{xxx}}\right]_{\gamma_1\gamma_2\gamma_4} [\mathbf{h_x}]^{\gamma_4}_{\alpha_4} [\mathbf{h_{xx}}]^{\gamma_2}_{\alpha_2\alpha_3} [\mathbf{h_x}]^{\gamma_1}_{\alpha_1} \\
&\quad + \left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_1\gamma_2} \left([\mathbf{h_{xxx}}]^{\gamma_2}_{\alpha_2\alpha_3\alpha_4} [\mathbf{h_x}]^{\gamma_1}_{\alpha_1} + [\mathbf{h_{xx}}]^{\gamma_2}_{\alpha_2\alpha_3} [\mathbf{h_{xx}}]^{\gamma_1}_{\alpha_1\alpha_4}\right) \\
&\quad + \left[p^{k-1}_{\mathbf{xxx}}\right]_{\gamma_1\gamma_2\gamma_4} [\mathbf{h_x}]^{\gamma_4}_{\alpha_4} [\mathbf{h_x}]^{\gamma_2}_{\alpha_2} [\mathbf{h_{xx}}]^{\gamma_1}_{\alpha_1\alpha_3} \\
&\quad + \left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_1\gamma_2} \left([\mathbf{h_{xx}}]^{\gamma_2}_{\alpha_2\alpha_4} [\mathbf{h_{xx}}]^{\gamma_1}_{\alpha_1\alpha_3} + [\mathbf{h_x}]^{\gamma_2}_{\alpha_2} [\mathbf{h_{xxx}}]^{\gamma_1}_{\alpha_1\alpha_3\alpha_4}\right) \\
&\quad + \left[p^{k-1}_{\mathbf{xxx}}\right]_{\gamma_1\gamma_3\gamma_4} [\mathbf{h_x}]^{\gamma_4}_{\alpha_4} [\mathbf{h_x}]^{\gamma_3}_{\alpha_3} [\mathbf{h_{xx}}]^{\gamma_1}_{\alpha_1\alpha_2} \\
&\quad + \left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_1\gamma_3} \left([\mathbf{h_{xx}}]^{\gamma_3}_{\alpha_3\alpha_4} [\mathbf{h_{xx}}]^{\gamma_1}_{\alpha_1\alpha_2} + [\mathbf{h_x}]^{\gamma_3}_{\alpha_3} [\mathbf{h_{xxx}}]^{\gamma_1}_{\alpha_1\alpha_2\alpha_4}\right) \\
&\quad + \left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_1\gamma_4} [\mathbf{h_x}]^{\gamma_4}_{\alpha_4} [\mathbf{h_{xxx}}]^{\gamma_1}_{\alpha_1\alpha_2\alpha_3} + \left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_1} [\mathbf{h_{xxxx}}]^{\gamma_1}_{\alpha_1\alpha_2\alpha_3\alpha_4}
\end{aligned}
$$

$$
\begin{aligned}
\left[p^k_{\sigma\sigma\mathbf{xx}}\right]_{\alpha_3\alpha_4} &= -\left[p^k_{\mathbf{x}}\right]_{\alpha_4} \left[p^k_{\mathbf{x}}\right]_{\alpha_3} \left[p^k_{\sigma\sigma}\right] - \left[p^k_{\mathbf{xx}}\right]_{\alpha_3\alpha_4} \left[p^k_{\sigma\sigma}\right] - \left[p^k_{\mathbf{x}}\right]_{\alpha_3} \left[p^k_{\sigma\sigma\mathbf{x}}\right]_{\alpha_4} - \left[p^k_{\mathbf{x}}\right]_{\alpha_4} \left[p^k_{\sigma\sigma\mathbf{x}}\right]_{\alpha_3} \\
&\quad + \left[p^1_{\sigma\sigma\mathbf{xx}}\right]_{\alpha_3\alpha_4} + \left[p^1_{\sigma\sigma\mathbf{x}}\right]_{\alpha_3} \left[p^1_{\mathbf{x}}\right]_{\alpha_4} + \left[p^1_{\sigma\sigma\mathbf{x}}\right]_{\alpha_4} \left[p^1_{\mathbf{x}}\right]_{\alpha_3} + \left[p^1_{\sigma\sigma}\right] \left[p^1_{\mathbf{x}}\right]_{\alpha_4} \left[p^1_{\mathbf{x}}\right]_{\alpha_3} + \left[p^1_{\sigma\sigma}\right] \left[p^1_{\mathbf{xx}}\right]_{\alpha_3\alpha_4} \\
&\quad + \left(\left[p^1_{\sigma\sigma\mathbf{x}}\right]_{\alpha_3} + \left[p^1_{\sigma\sigma}\right] \left[p^1_{\mathbf{x}}\right]_{\alpha_3}\right) \left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_4} [\mathbf{h_x}]^{\gamma_4}_{\alpha_4}
\end{aligned}
$$

$$
\begin{aligned}
&+\left(\left[p^1_{\sigma\sigma\mathbf{x}}\right]_{\alpha_4}+\left[p^1_{\sigma\sigma}\right]\left[p^1_{\mathbf{x}}\right]_{\alpha_4}\right)\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_3}_{\alpha_3}\\
&+\left[p^1_{\sigma\sigma}\right]\left(\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_3}_{\alpha_3}+\left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_3\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_3}_{\alpha_3}+\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_3}\left[\mathbf{h}_{\mathbf{xx}}\right]^{\gamma_3}_{\alpha_3\alpha_4}\right)\\
&+2M^{-1}\left(\left[M_{\mathbf{x}_{t+1}\mathbf{x}_{t+1}\mathbf{x}_{t+1}}\right]_{\gamma_1\gamma_3\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}+\left[M_{\mathbf{x}_{t+1}\mathbf{x}_{t+1}\mathbf{x}_t}\right]_{\gamma_1\gamma_3\alpha_4}\right)\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_3}_{\alpha_3}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_2}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[\mathbf{I}\right]^{\phi_2}_{\phi_1}\\
&+2M^{-1}\left[M_{\mathbf{x}_{t+1}\mathbf{x}_{t+1}}\right]_{\gamma_1\gamma_3}\left[\mathbf{h}_{\mathbf{xx}}\right]^{\gamma_3}_{\alpha_3\alpha_4}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_2}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[\mathbf{I}\right]^{\phi_2}_{\phi_1}\\
&+2M^{-1}\left(\left[M_{\mathbf{x}_{t+1}\mathbf{x}_t\mathbf{x}_{t+1}}\right]_{\gamma_1 a_3\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}+\left[M_{\mathbf{x}_{t+1}\mathbf{x}_t\mathbf{x}_t}\right]_{\gamma_1 a_3\alpha_4}\right)\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_2}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[\mathbf{I}\right]^{\phi_2}_{\phi_1}\\
&+2M^{-1}\left(\left[M_{\mathbf{x}_{t+1}\mathbf{x}_{t+1}}\right]_{\gamma_1\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_3}_{\alpha_3}+\left[M_{\mathbf{x}_{t+1}\mathbf{x}_t}\right]_{\gamma_1 a_3}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\right)\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_2}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[\mathbf{I}\right]^{\phi_2}_{\phi_1}\\
&+2M^{-1}\left(\left[M_{\mathbf{x}_{t+1}\mathbf{x}_{t+1}}\right]_{\gamma_1\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_3}_{\alpha_3}+\left[M_{\mathbf{x}_{t+1}\mathbf{x}_t}\right]_{\gamma_1 a_3}\right)\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_2\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[\mathbf{I}\right]^{\phi_2}_{\phi_1}\\
&+2M^{-1}\left(\left[M_{\mathbf{x}_{t+1}\mathbf{x}_{t+1}}\right]_{\gamma_1\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}+\left[M_{\mathbf{x}_{t+1}\mathbf{x}_t}\right]_{\gamma_1 a_4}\right)\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_3}_{\alpha_3}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_2}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[\mathbf{I}\right]^{\phi_2}_{\phi_1}\\
&+2M^{-1}\left[M_{\mathbf{x}_{t+1}}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_3}_{\alpha_3}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_2}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[\mathbf{I}\right]^{\phi_2}_{\phi_1}\\
&+2M^{-1}\left[M_{\mathbf{x}_{t+1}}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_3\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_3}_{\alpha_3}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_2}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[\mathbf{I}\right]^{\phi_2}_{\phi_1}\\
&+2M^{-1}\left[M_{\mathbf{x}_{t+1}}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_3}\left[\mathbf{h}_{\mathbf{xx}}\right]^{\gamma_3}_{\alpha_3\alpha_4}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_2}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[\mathbf{I}\right]^{\phi_2}_{\phi_1}\\
&+2M^{-1}\left[M_{\mathbf{x}_{t+1}}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_3}_{\alpha_3}\left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_2\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[\mathbf{I}\right]^{\phi_2}_{\phi_1}\\
&+2M^{-1}\left(\left[M_{\mathbf{x}_{t+1}\mathbf{x}_{t+1}}\right]_{\gamma_1\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}+\left[M_{\mathbf{x}_{t+1}\mathbf{x}_t}\right]_{\gamma_1 a_4}\right)\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_2\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_3}_{\alpha_3}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[\mathbf{I}\right]^{\phi_2}_{\phi_1}\\
&+2M^{-1}\left[M_{\mathbf{x}_{t+1}}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}\left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_2\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_3}_{\alpha_3}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[\mathbf{I}\right]^{\phi_2}_{\phi_1}\\
&+2M^{-1}\left[M_{\mathbf{x}_{t+1}}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[p^{k-1}_{\mathbf{xxx}}\right]_{\gamma_2\gamma_3\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_3}_{\alpha_3}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[\mathbf{I}\right]^{\phi_2}_{\phi_1}\\
&+2M^{-1}\left[M_{\mathbf{x}_{t+1}}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_2\gamma_3}\left[\mathbf{h}_{\mathbf{xx}}\right]^{\gamma_3}_{\alpha_3\alpha_4}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[\mathbf{I}\right]^{\phi_2}_{\phi_1}\\
&+\left(\left[p^1_{\mathbf{x}}\right]_{\alpha_3}\left[p^1_{\mathbf{x}}\right]_{\alpha_4}+\left[p^1_{\mathbf{xx}}\right]_{\alpha_3\alpha_4}\right)\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_2}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[\mathbf{I}\right]^{\phi_1}_{\phi_2}\\
&+\left[p^1_{\mathbf{x}}\right]_{\alpha_3}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_2}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[\mathbf{I}\right]^{\phi_1}_{\phi_2}\\
&+2\left[p^1_{\mathbf{x}}\right]_{\alpha_3}\left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_2\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[\mathbf{I}\right]^{\phi_1}_{\phi_2}\\
&+\left[p^1_{\mathbf{x}}\right]_{\alpha_4}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_3}_{\alpha_3}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_2}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[\mathbf{I}\right]^{\phi_1}_{\phi_2}\\
&+\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_3}_{\alpha_3}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_2}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[\mathbf{I}\right]^{\phi_1}_{\phi_2}\\
&+\left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_3\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_3}_{\alpha_3}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_2}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[\mathbf{I}\right]^{\phi_1}_{\phi_2}\\
&+\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_3}\left[\mathbf{h}_{\mathbf{xx}}\right]^{\gamma_3}_{\alpha_3\alpha_4}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_2}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[\mathbf{I}\right]^{\phi_1}_{\phi_2}\\
&+2\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_3}_{\alpha_3}\left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_2\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[\mathbf{I}\right]^{\phi_1}_{\phi_2}\\
&+2\left[p^1_{\mathbf{x}}\right]_{\alpha_4}\left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_2\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_3}_{\alpha_3}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[\mathbf{I}\right]^{\phi_1}_{\phi_2}\\
&+2\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}\left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_2\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_3}_{\alpha_3}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[\mathbf{I}\right]^{\phi_1}_{\phi_2}\\
&+2\left[p^{k-1}_{\mathbf{xxx}}\right]_{\gamma_2\gamma_3\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_3}_{\alpha_3}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[\mathbf{I}\right]^{\phi_1}_{\phi_2}\\
&+2\left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_2\gamma_3}\left[\mathbf{h}_{\mathbf{xx}}\right]^{\gamma_3}_{\alpha_3\alpha_4}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[\mathbf{I}\right]^{\phi_1}_{\phi_2}+2\left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_2\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_3}_{\alpha_3}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_1\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[\mathbf{I}\right]^{\phi_1}_{\phi_2}\\
&+\left(\left[p^1_{\mathbf{x}}\right]_{\alpha_3}\left[p^1_{\mathbf{x}}\right]_{\alpha_4}+\left[p^1_{\mathbf{xx}}\right]_{\alpha_3\alpha_4}\right)\left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_1\gamma_2}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[\mathbf{I}\right]^{\phi_1}_{\phi_2}\\
&+\left[p^1_{\mathbf{x}}\right]_{\alpha_3}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}\left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_1\gamma_2}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[\mathbf{I}\right]^{\phi_1}_{\phi_2}\\
&+\left[p^1_{\mathbf{x}}\right]_{\alpha_3}\left[p^{k-1}_{\mathbf{xxx}}\right]_{\gamma_1\gamma_2\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[\mathbf{I}\right]^{\phi_1}_{\phi_2}+\left[p^1_{\mathbf{x}}\right]_{\alpha_4}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_3}_{\alpha_3}\left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_1\gamma_2}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[\mathbf{I}\right]^{\phi_1}_{\phi_2}\\
&+\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}\left[p^{k-1}_{\mathbf{x}}\right]_{\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_3}_{\alpha_3}\left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_1\gamma_2}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[\mathbf{I}\right]^{\phi_1}_{\phi_2}\\
&+\left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_3\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_4}_{\alpha_4}\left[\mathbf{h}_{\mathbf{x}}\right]^{\gamma_3}_{\alpha_3}\left[p^{k-1}_{\mathbf{xx}}\right]_{\gamma_1\gamma_2}\left[\boldsymbol{\eta}\right]^{\gamma_2}_{\phi_2}\left[\boldsymbol{\eta}\right]^{\gamma_1}_{\phi_1}\left[\mathbf{I}\right]^{\phi_1}_{\phi_2}
\end{aligned}
$$

$$
\begin{aligned}
&+\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_3}\left[\mathbf{h}_{\mathbf{xx}}\right]_{\alpha_3\alpha_4}^{\gamma_3}\left[p_{\mathbf{xx}}^{k-1}\right]_{\gamma_1\gamma_2}\left[\boldsymbol{\eta}\right]_{\phi_2}^{\gamma_2}\left[\boldsymbol{\eta}\right]_{\phi_1}^{\gamma_1}\left[\mathbf{I}\right]_{\phi_2}^{\phi_1}+\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]_{\alpha_3}^{\gamma_3}\left[p_{\mathbf{xxx}}^{k-1}\right]_{\gamma_1\gamma_2\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]_{\alpha_4}^{\gamma_4}\left[\boldsymbol{\eta}\right]_{\phi_2}^{\gamma_2}\left[\boldsymbol{\eta}\right]_{\phi_1}^{\gamma_1}\left[\mathbf{I}\right]_{\phi_2}^{\phi_1}\\
&+\left[p_{\mathbf{x}}^{1}\right]_{\alpha_4}\left[p_{\mathbf{xxx}}^{k-1}\right]_{\gamma_1\gamma_2\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]_{\alpha_3}^{\gamma_3}\left[\boldsymbol{\eta}\right]_{\phi_2}^{\gamma_2}\left[\boldsymbol{\eta}\right]_{\phi_1}^{\gamma_1}\left[\mathbf{I}\right]_{\phi_2}^{\phi_1}+\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]_{\alpha_4}^{\gamma_4}\left[p_{\mathbf{xxx}}^{k-1}\right]_{\gamma_1\gamma_2\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]_{\alpha_3}^{\gamma_3}\left[\boldsymbol{\eta}\right]_{\phi_2}^{\gamma_2}\left[\boldsymbol{\eta}\right]_{\phi_1}^{\gamma_1}\left[\mathbf{I}\right]_{\phi_2}^{\phi_1}\\
&+\left[p_{\mathbf{xxxx}}^{k-1}\right]_{\gamma_1\gamma_2\gamma_3\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]_{\alpha_4}^{\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]_{\alpha_3}^{\gamma_3}\left[\boldsymbol{\eta}\right]_{\phi_2}^{\gamma_2}\left[\boldsymbol{\eta}\right]_{\phi_1}^{\gamma_1}\left[\mathbf{I}\right]_{\phi_2}^{\phi_1}+\left[p_{\mathbf{xxx}}^{k-1}\right]_{\gamma_1\gamma_2\gamma_3}\left[\mathbf{h}_{\mathbf{xx}}\right]_{\alpha_3\alpha_4}^{\gamma_3}\left[\boldsymbol{\eta}\right]_{\phi_2}^{\gamma_2}\left[\boldsymbol{\eta}\right]_{\phi_1}^{\gamma_1}\left[\mathbf{I}\right]_{\phi_2}^{\phi_1}\\
&+\left(\left[p_{\mathbf{x}}^{1}\right]_{\alpha_3}\left[p_{\mathbf{x}}^{1}\right]_{\alpha_4}+\left[p_{\mathbf{xx}}^{1}\right]_{\alpha_3\alpha_4}\right)\left[p_{\sigma\sigma}^{k-1}\right]+\left[p_{\mathbf{x}}^{1}\right]_{\alpha_3}\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]_{\alpha_4}^{\gamma_4}\left[p_{\sigma\sigma}^{k-1}\right]+\left[p_{\mathbf{x}}^{1}\right]_{\alpha_3}\left[p_{\sigma\sigma\mathbf{x}}^{k-1}\right]_{\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]_{\alpha_4}^{\gamma_4}\\
&+\left[p_{\mathbf{x}}^{1}\right]_{\alpha_4}\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]_{\alpha_3}^{\gamma_3}\left[p_{\sigma\sigma}^{k-1}\right]+\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]_{\alpha_4}^{\gamma_4}\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]_{\alpha_3}^{\gamma_3}\left[p_{\sigma\sigma}^{k-1}\right]+\left[p_{\mathbf{xx}}^{k-1}\right]_{\gamma_3\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]_{\alpha_4}^{\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]_{\alpha_3}^{\gamma_3}\left[p_{\sigma\sigma}^{k-1}\right]\\
&+\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_3}\left[\mathbf{h}_{\mathbf{xx}}\right]_{\alpha_3\alpha_4}^{\gamma_3}\left[p_{\sigma\sigma}^{k-1}\right]+\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]_{\alpha_3}^{\gamma_3}\left[p_{\sigma\sigma\mathbf{x}}^{k-1}\right]_{\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]_{\alpha_4}^{\gamma_4}+\left[p_{\mathbf{x}}^{1}\right]_{\alpha_4}\left[p_{\sigma\sigma\mathbf{x}}^{k-1}\right]_{\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]_{\alpha_3}^{\gamma_3}\\
&+\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]_{\alpha_4}^{\gamma_4}\left[p_{\sigma\sigma\mathbf{x}}^{k-1}\right]_{\gamma_3}\left[\mathbf{h}_{\mathbf{x}}\right]_{\alpha_3}^{\gamma_3}+\left[p_{\sigma\sigma\mathbf{xx}}^{k-1}\right]_{\gamma_3\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]_{\alpha_4}^{\gamma_4}\left[\mathbf{h}_{\mathbf{x}}\right]_{\alpha_3}^{\gamma_3}+\left[p_{\sigma\sigma\mathbf{x}}^{k-1}\right]_{\gamma_3}\left[\mathbf{h}_{\mathbf{xx}}\right]_{\alpha_3\alpha_4}^{\gamma_3}
\end{aligned}
$$

Finally, let $m^4\left(\boldsymbol{\epsilon}_{t+1}\right)$ denotes the fourth moment of $\boldsymbol{\epsilon}_{t+1}\left(\phi_1,1\right)$ for $\phi_1=1,2,...,n_e$. Hence, $\mathbf{m}^4\left(\boldsymbol{\epsilon}_{t+1}\right)$ is an $n_e\times n_e\times n_e\times n_e$ matrix.

$$
\begin{aligned}
&\left[p_{\sigma\sigma\sigma\sigma}^{k}\right]=-3\left[p_{\sigma\sigma}^{k}\right]\left[p_{\sigma\sigma}^{k}\right]+\left[p_{\sigma\sigma\sigma\sigma}^{1}\right]+3\left[p_{\sigma\sigma}^{1}\right]\left[p_{\sigma\sigma}^{1}\right]\\
&\quad+6M^{-1}\left[M_{\mathbf{x}_{t+1}\mathbf{x}_{t+1}}\right]_{\gamma_1\gamma_4}\left[\boldsymbol{\eta}\right]_{\phi_4}^{\gamma_4}\left[\boldsymbol{\eta}\right]_{\phi_1}^{\gamma_1}\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_3}\left[\boldsymbol{\eta}\right]_{\phi_3}^{\gamma_3}\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_2}\left[\boldsymbol{\eta}\right]_{\phi_2}^{\gamma_2}\left[\mathbf{m}^4\right]_{\phi_1\phi_3\phi_4}^{\phi_2}\\
&\quad+4M^{-1}\left[M_{\mathbf{x}_{t+1}\mathbf{x}_{t+1}\mathbf{x}_{t+1}}\right]_{\gamma_1\gamma_2\gamma_4}\left[\boldsymbol{\eta}\right]_{\phi_4}^{\gamma_4}\left[\boldsymbol{\eta}\right]_{\phi_2}^{\gamma_2}\left[\boldsymbol{\eta}\right]_{\phi_1}^{\gamma_1}\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_3}\left[\boldsymbol{\eta}\right]_{\phi_3}^{\gamma_3}\left[\mathbf{m}^4\right]_{\phi_1\phi_2\phi_4}^{\phi_3}\\
&\quad+6M^{-1}\left[M_{\mathbf{x}_{t+1}\mathbf{x}_{t+1}}\right]_{\gamma_1\gamma_2}\left[\boldsymbol{\eta}\right]_{\phi_2}^{\gamma_2}\left[\boldsymbol{\eta}\right]_{\phi_1}^{\gamma_1}\left[p_{\mathbf{xx}}^{k-1}\right]_{\gamma_3\gamma_4}\left[\boldsymbol{\eta}\right]_{\phi_4}^{\gamma_4}\left[\boldsymbol{\eta}\right]_{\phi_3}^{\gamma_3}\left[\mathbf{m}^4\right]_{\phi_1\phi_2\phi_4}^{\phi_3}+6\left[p_{\sigma\sigma}^{1}\right]\left[p_{\sigma\sigma}^{k-1}\right]\\
&\quad+4M^{-1}\left[M_{\mathbf{x}_{t+1}}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]_{\phi_1}^{\gamma_1}\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_4}\left[\boldsymbol{\eta}\right]_{\phi_4}^{\gamma_4}\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_3}\left[\boldsymbol{\eta}\right]_{\phi_3}^{\gamma_3}\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_2}\left[\boldsymbol{\eta}\right]_{\phi_2}^{\gamma_2}\left[\mathbf{m}^4\right]_{\phi_2\phi_3\phi_4}^{\phi_1}\\
&\quad+12M^{-1}\left[M_{\mathbf{x}_{t+1}}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]_{\phi_1}^{\gamma_1}\left[p_{\mathbf{xx}}^{k-1}\right]_{\gamma_3\gamma_4}\left[\boldsymbol{\eta}\right]_{\phi_4}^{\gamma_4}\left[\boldsymbol{\eta}\right]_{\phi_3}^{\gamma_3}\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_2}\left[\boldsymbol{\eta}\right]_{\phi_2}^{\gamma_2}\left[\mathbf{m}^4\right]_{\phi_2\phi_3\phi_4}^{\phi_1}\\
&\quad+12M^{-1}\left[M_{\mathbf{x}_{t+1}}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]_{\phi_1}^{\gamma_1}\left[p_{\sigma\sigma}^{k-1}\right]\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_2}\left[\boldsymbol{\eta}\right]_{\phi_2}^{\gamma_2}\left[\mathbf{I}\right]_{\phi_1}^{\phi_2}\\
&\quad+4M_t^{-1}\left[M_{\mathbf{x}_{t+1}}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]_{\phi_1}^{\gamma_1}\left[p_{\mathbf{xxx}}^{k-1}\right]_{\gamma_2\gamma_3\gamma_4}\left[\boldsymbol{\eta}\right]_{\phi_4}^{\gamma_4}\left[\boldsymbol{\eta}\right]_{\phi_3}^{\gamma_3}\left[\boldsymbol{\eta}\right]_{\phi_2}^{\gamma_2}\left[\mathbf{m}^4\right]_{\phi_2\phi_3\phi_4}^{\phi_1}\\
&\quad+12M^{-1}\left[M_{\mathbf{x}_{t+1}}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]_{\phi_1}^{\gamma_1}\left[p_{\mathbf{x}\boldsymbol{\sigma}\boldsymbol{\sigma}}^{k-1}\right]_{\gamma_2}\left[\boldsymbol{\eta}\right]_{\phi_2}^{\gamma_2}\left[\mathbf{I}\right]_{\phi_1}^{\phi_2}\\
&\quad+\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_4}\left[\boldsymbol{\eta}\right]_{\phi_4}^{\gamma_4}\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_3}\left[\boldsymbol{\eta}\right]_{\phi_3}^{\gamma_3}\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_2}\left[\boldsymbol{\eta}\right]_{\phi_2}^{\gamma_2}\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]_{\phi_1}^{\gamma_1}\left[\mathbf{m}^4\right]_{\phi_2\phi_3\phi_4}^{\phi_1}\\
&\quad+6\left[p_{\mathbf{xx}}^{k-1}\right]_{\gamma_3\gamma_4}\left[\boldsymbol{\eta}\right]_{\phi_4}^{\gamma_4}\left[\boldsymbol{\eta}\right]_{\phi_3}^{\gamma_3}\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_2}\left[\boldsymbol{\eta}\right]_{\phi_2}^{\gamma_2}\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]_{\phi_1}^{\gamma_1}\left[\mathbf{m}^4\right]_{\phi_2\phi_3\phi_4}^{\phi_1}\\
&\quad+4\left[p_{\mathbf{xxx}}^{k-1}\right]_{\gamma_2\gamma_3\gamma_4}\left[\boldsymbol{\eta}\right]_{\phi_4}^{\gamma_4}\left[\boldsymbol{\eta}\right]_{\phi_3}^{\gamma_3}\left[\boldsymbol{\eta}\right]_{\phi_2}^{\gamma_2}\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]_{\phi_1}^{\gamma_1}\left[\mathbf{m}^4\right]_{\phi_2\phi_3\phi_4}^{\phi_1}\\
&\quad+3\left[p_{\mathbf{xx}}^{k-1}\right]_{\gamma_2\gamma_3}\left[\boldsymbol{\eta}\right]_{\phi_3}^{\gamma_3}\left[\boldsymbol{\eta}\right]_{\phi_2}^{\gamma_2}\left[p_{\mathbf{xx}}^{k-1}\right]_{\gamma_1\gamma_4}\left[\boldsymbol{\eta}\right]_{\phi_4}^{\gamma_4}\left[\boldsymbol{\eta}\right]_{\phi_1}^{\gamma_1}\left[\mathbf{m}^4\right]_{\phi_2\phi_3\phi_4}^{\phi_1}\\
&\quad+12\left[p_{\mathbf{x}\boldsymbol{\sigma}\boldsymbol{\sigma}}^{k-1}\right]_{\gamma_2}\left[\boldsymbol{\eta}\right]_{\phi_2}^{\gamma_2}\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]_{\phi_1}^{\gamma_1}\left[\mathbf{I}\right]_{\phi_2}^{\phi_1}+6\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_3}\left[\boldsymbol{\eta}\right]_{\phi_3}^{\gamma_3}\left[p_{\boldsymbol{\sigma}\boldsymbol{\sigma}}^{k-1}\right]\left[p_{\mathbf{x}}^{k-1}\right]_{\gamma_1}\left[\boldsymbol{\eta}\right]_{\phi_1}^{\gamma_1}\left[\mathbf{I}\right]_{\phi_3}^{\phi_1}\\
&\quad+6\left[p_{\boldsymbol{\sigma}\boldsymbol{\sigma}}^{k-1}\right]\left[p_{\mathbf{xx}}^{k-1}\right]_{\gamma_1\gamma_4}\left[\boldsymbol{\eta}\right]_{\phi_4}^{\gamma_4}\left[\boldsymbol{\eta}\right]_{\phi_1}^{\gamma_1}\left[\mathbf{I}\right]_{\phi_4}^{\phi_1}+\left[p_{\mathbf{xxxx}}^{k-1}\right]_{\gamma_1\gamma_2\gamma_3\gamma_4}\left[\boldsymbol{\eta}\right]_{\phi_4}^{\gamma_4}\left[\boldsymbol{\eta}\right]_{\phi_3}^{\gamma_3}\left[\boldsymbol{\eta}\right]_{\phi_2}^{\gamma_2}\left[\boldsymbol{\eta}\right]_{\phi_1}^{\gamma_1}\left[\mathbf{m}^4\right]_{\phi_2\phi_3\phi_4}^{\phi_1}\\
&\quad+6\left[p_{\mathbf{xx}\boldsymbol{\sigma}\boldsymbol{\sigma}}^{k-1}\right]_{\gamma_1\gamma_2}\left[\boldsymbol{\eta}\right]_{\phi_2}^{\gamma_2}\left[\boldsymbol{\eta}\right]_{\phi_1}^{\gamma_1}\left[\mathbf{I}\right]_{\phi_2}^{\phi_1}+3\left[p_{\boldsymbol{\sigma}\boldsymbol{\sigma}}^{k-1}\right]\left[p_{\sigma\sigma}^{k-1}\right]+\left[p_{\sigma\sigma\sigma\sigma}^{k-1}\right]
\end{aligned}
$$

B Computing conditional expectations by the perturbation method

Suppose we have a perturbation approximation to the variable r_t and want to compute its conditional expectations, i.e. $r_{1,t}\equiv\mathbb{E}_t\left[r_{t+1}\right]$, $r_{2,t}\equiv\mathbb{E}_t\left[r_{t+2}\right]$, $r_{3,t}\equiv\mathbb{E}_t\left[r_{t+3}\right]$, etc. The law of iterated expectations implies $r_{2,t}\equiv\mathbb{E}_t\left[r_{t+2}\right]=\mathbb{E}_t\left[E_{t+1}\left[r_{t+2}\right]\right]=\mathbb{E}_t\left[r_{1,t+1}\right]$ and so on. We therefore only need a formula to compute $r_{1,t}\equiv E_t\left[r_{t+1}\right]$ as all remaining expectations follow by iterating this formula. Hence, consider the problem

$$
b\left(\mathbf{x}_t,\sigma\right)=\mathbb{E}_t\left[r\left(\mathbf{x}_{t+1},\sigma\right)\right], \tag{45}
$$

and observe that

$$
F\left(\mathbf{x}_t,\sigma\right)\equiv\mathbb{E}_t\left[-b\left(\mathbf{x}_t,\sigma\right)+r\left(\mathbf{h}\left(\mathbf{x}_t\right)+\sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1},\sigma\right)\right]=0, \tag{46}
$$

given (10). The conditional expectation $\mathbb{E}_t$ is evaluated under the same measure as the innovations $\boldsymbol{\epsilon}_{t+1}$. The expression in (46) must hold for all values of $(\mathbf{x}_t, \sigma)$ and for all possible derivatives of $F(\mathbf{x}_t, \sigma)$. Using these conditions, it is possible to compute all derivatives of b with respect to $(\mathbf{x}_t, \sigma)$ around the deterministic steady state, given derivatives of $\mathbf{h}(\mathbf{x}_t)$ and $r(\mathbf{x}_{t+1}, \sigma)$ around the same point. In presenting the formulas below, we adopt the same notation as in Appendix A.

B.1 Conditional expectations: first-order terms

$$[b_{\mathbf{x}}]_{\alpha_1} = [r_{\mathbf{x}}]_{\gamma_1} [\mathbf{h}_{\mathbf{x}}]^{\gamma_1}_{\alpha_1}$$

B.2 Conditional expectations: second-order terms

$$[b_{\mathbf{xx}}]_{\alpha_1\alpha_2} = [r_{\mathbf{xx}}]_{\gamma_1\gamma_2} [\mathbf{h}_{\mathbf{x}}]^{\gamma_2}_{\alpha_2} [\mathbf{h}_{\mathbf{x}}]^{\gamma_1}_{\alpha_1} + [r_{\mathbf{x}}]_{\gamma_1} [\mathbf{h}_{\mathbf{xx}}]^{\gamma_1}_{\alpha_1\alpha_2}$$

$$[b_{\sigma\sigma}] = [r_{\mathbf{xx}}]_{\gamma_1\gamma_2} [\boldsymbol{\eta}]^{\gamma_1}_{\phi_1} [\boldsymbol{\eta}]^{\gamma_2}_{\phi_2} [\mathbf{I}]^{\phi_2}_{\phi_1} + [r_{\sigma\sigma}]$$

B.3 Conditional expectations: third-order terms

$$\begin{aligned}
[b_{\mathbf{xxx}}]_{\alpha_1\alpha_2\alpha_3} &= [r_{\mathbf{xxx}}]_{\gamma_1\gamma_2\gamma_3} [\mathbf{h}_{\mathbf{x}}]^{\gamma_3}_{\alpha_3} [\mathbf{h}_{\mathbf{x}}]^{\gamma_2}_{\alpha_2} [\mathbf{h}_{\mathbf{x}}]^{\gamma_1}_{\alpha_1} \\
&\quad + [r_{\mathbf{xx}}]_{\gamma_1\gamma_2} [\mathbf{h}_{\mathbf{xx}}]^{\gamma_2}_{\alpha_2\alpha_3} [\mathbf{h}_{\mathbf{x}}]^{\gamma_1}_{\alpha_1} + [r_{\mathbf{xx}}]_{\gamma_1\gamma_2} [\mathbf{h}_{\mathbf{x}}]^{\gamma_2}_{\alpha_2} [\mathbf{h}_{\mathbf{xx}}]^{\gamma_1}_{\alpha_1\alpha_3} \\
&\quad + [r_{\mathbf{xx}}]_{\gamma_1\gamma_3} [\mathbf{h}_{\mathbf{x}}]^{\gamma_3}_{\alpha_3} [\mathbf{h}_{\mathbf{xx}}]^{\gamma_1}_{\alpha_1\alpha_2} + [r_{\mathbf{x}}]_{\gamma_1} [\mathbf{h}_{\mathbf{xxx}}]^{\gamma_1}_{\alpha_1\alpha_2\alpha_3}
\end{aligned}$$

$$[b_{\sigma\sigma\mathbf{x}}]_{\alpha_3} = [r_{\mathbf{xxx}}]_{\gamma_1\gamma_2\gamma_3} [\mathbf{h}_{\mathbf{x}}]^{\gamma_3}_{\alpha_3} [\boldsymbol{\eta}]^{\gamma_1}_{\phi_1} [\boldsymbol{\eta}]^{\gamma_2}_{\phi_2} [\mathbf{I}]^{\phi_2}_{\phi_1} + [r_{\sigma\sigma\mathbf{x}}]_{\gamma_3} [\mathbf{h}_{\mathbf{x}}]^{\gamma_3}_{\alpha_3}$$

B.4 Conditional expectations: fourth-order terms

$$\begin{aligned}
[b_{\mathbf{xxxx}}]_{\alpha_1\alpha_2\alpha_3\alpha_4} &= [r_{\mathbf{xxxx}}]_{\gamma_1\gamma_2\gamma_3\gamma_4} [\mathbf{h}_{\mathbf{x}}]^{\gamma_4}_{\alpha_4} [\mathbf{h}_{\mathbf{x}}]^{\gamma_3}_{\alpha_3} [\mathbf{h}_{\mathbf{x}}]^{\gamma_2}_{\alpha_2} [\mathbf{h}_{\mathbf{x}}]^{\gamma_1}_{\alpha_1} \\
&\quad + [r_{\mathbf{xxx}}]_{\gamma_1\gamma_2\gamma_3} [\mathbf{h}_{\mathbf{xx}}]^{\gamma_3}_{\alpha_3\alpha_4} [\mathbf{h}_{\mathbf{x}}]^{\gamma_2}_{\alpha_2} [\mathbf{h}_{\mathbf{x}}]^{\gamma_1}_{\alpha_1} + [r_{\mathbf{xxx}}]_{\gamma_1\gamma_2\gamma_3} [\mathbf{h}_{\mathbf{x}}]^{\gamma_3}_{\alpha_3} [\mathbf{h}_{\mathbf{xx}}]^{\gamma_2}_{\alpha_2\alpha_4} [\mathbf{h}_{\mathbf{x}}]^{\gamma_1}_{\alpha_1} \\
&\quad + [r_{\mathbf{xxx}}]_{\gamma_1\gamma_2\gamma_3} [\mathbf{h}_{\mathbf{x}}]^{\gamma_3}_{\alpha_3} [\mathbf{h}_{\mathbf{x}}]^{\gamma_2}_{\alpha_2} [\mathbf{h}_{\mathbf{xx}}]^{\gamma_1}_{\alpha_1\alpha_4} + [r_{\mathbf{xxx}}]_{\gamma_1\gamma_2\gamma_4} [\mathbf{h}_{\mathbf{x}}]^{\gamma_4}_{\alpha_4} [\mathbf{h}_{\mathbf{xx}}]^{\gamma_2}_{\alpha_2\alpha_3} [\mathbf{h}_{\mathbf{x}}]^{\gamma_1}_{\alpha_1} \\
&\quad + [r_{\mathbf{xx}}]_{\gamma_1\gamma_2} [\mathbf{h}_{\mathbf{xxx}}]^{\gamma_2}_{\alpha_2\alpha_3\alpha_4} [\mathbf{h}_{\mathbf{x}}]^{\gamma_1}_{\alpha_1} + [r_{\mathbf{xx}}]_{\gamma_1\gamma_2} [\mathbf{h}_{\mathbf{xx}}]^{\gamma_2}_{\alpha_2\alpha_3} [\mathbf{h}_{\mathbf{xx}}]^{\gamma_1}_{\alpha_1\alpha_4} \\
&\quad + [r_{\mathbf{xxx}}]_{\gamma_1\gamma_2\gamma_4} [\mathbf{h}_{\mathbf{x}}]^{\gamma_4}_{\alpha_4} [\mathbf{h}_{\mathbf{x}}]^{\gamma_2}_{\alpha_2} [\mathbf{h}_{\mathbf{xx}}]^{\gamma_1}_{\alpha_1\alpha_3} + [r_{\mathbf{xx}}]_{\gamma_1\gamma_2} [\mathbf{h}_{\mathbf{xx}}]^{\gamma_2}_{\alpha_2\alpha_4} [\mathbf{h}_{\mathbf{xx}}]^{\gamma_1}_{\alpha_1\alpha_3} \\
&\quad + [r_{\mathbf{xx}}]_{\gamma_1\gamma_2} [\mathbf{h}_{\mathbf{x}}]^{\gamma_2}_{\alpha_2} [\mathbf{h}_{\mathbf{xxx}}]^{\gamma_1}_{\alpha_1\alpha_3\alpha_4} + [r_{\mathbf{xxx}}]_{\gamma_1\gamma_3\gamma_4} [\mathbf{h}_{\mathbf{x}}]^{\gamma_4}_{\alpha_4} [\mathbf{h}_{\mathbf{x}}]^{\gamma_3}_{\alpha_3} [\mathbf{h}_{\mathbf{xx}}]^{\gamma_1}_{\alpha_1\alpha_2} \\
&\quad + [r_{\mathbf{xx}}]_{\gamma_1\gamma_3} [\mathbf{h}_{\mathbf{xx}}]^{\gamma_3}_{\alpha_3\alpha_4} [\mathbf{h}_{\mathbf{xx}}]^{\gamma_1}_{\alpha_1\alpha_2} + [r_{\mathbf{xx}}]_{\gamma_1\gamma_3} [\mathbf{h}_{\mathbf{x}}]^{\gamma_3}_{\alpha_3} [\mathbf{h}_{\mathbf{xxx}}]^{\gamma_1}_{\alpha_1\alpha_2\alpha_4} \\
&\quad + [r_{\mathbf{xx}}]_{\gamma_1\gamma_4} [\mathbf{h}_{\mathbf{x}}]^{\gamma_4}_{\alpha_4} [\mathbf{h}_{\mathbf{xxx}}]^{\gamma_1}_{\alpha_1\alpha_2\alpha_3} + [r_{\mathbf{x}}]_{\gamma_1} [\mathbf{h}_{\mathbf{xxxx}}]^{\gamma_1}_{\alpha_1\alpha_2\alpha_3\alpha_4}
\end{aligned}$$

$$\begin{aligned}
[b_{\sigma\sigma\mathbf{xx}}]_{\alpha_3\alpha_4} &= [r_{\mathbf{xxxx}}]_{\gamma_1\gamma_2\gamma_3\gamma_4} [\mathbf{h}_{\mathbf{x}}]^{\gamma_4}_{\alpha_4} [\mathbf{h}_{\mathbf{x}}]^{\gamma_3}_{\alpha_3} [\boldsymbol{\eta}]^{\gamma_1}_{\phi_1} [\boldsymbol{\eta}]^{\gamma_2}_{\phi_2} [\mathbf{I}]^{\phi_2}_{\phi_1} \\
&\quad + [r_{\mathbf{xxx}}]_{\gamma_1\gamma_2\gamma_3} [\mathbf{h}_{\mathbf{xx}}]^{\gamma_3}_{\alpha_3\alpha_4} [\boldsymbol{\eta}]^{\gamma_1}_{\phi_1} [\boldsymbol{\eta}]^{\gamma_2}_{\phi_2} [\mathbf{I}]^{\phi_2}_{\phi_1} \\
&\quad + [r_{\sigma\sigma\mathbf{xx}}]_{\gamma_3\gamma_4} [\mathbf{h}_{\mathbf{x}}]^{\gamma_4}_{\alpha_4} [\mathbf{h}_{\mathbf{x}}]^{\gamma_3}_{\alpha_3} + [r_{\sigma\sigma\mathbf{x}}]_{\gamma_3} [\mathbf{h}_{\mathbf{xx}}]^{\gamma_3}_{\alpha_3\alpha_4}
\end{aligned}$$

$$\begin{aligned}
[b_{\sigma\sigma\sigma\sigma}] &= [r_{\mathbf{xxxx}}]_{\gamma_1\gamma_2\gamma_3\gamma_4} [\boldsymbol{\eta}]^{\gamma_4}_{\phi_4} [\boldsymbol{\eta}]^{\gamma_3}_{\phi_3} [\boldsymbol{\eta}]^{\gamma_1}_{\phi_1} [\boldsymbol{\eta}]^{\gamma_2}_{\phi_2} \left[\mathbf{m}^4(\boldsymbol{\epsilon}_{t+1})\right]^{\phi_1}_{\phi_2\phi_3\phi_4} \\
&\quad + 6\,[r_{\mathbf{xx}\sigma\sigma}]_{\gamma_1\gamma_2} [\boldsymbol{\eta}]^{\gamma_1}_{\phi_1} [\boldsymbol{\eta}]^{\gamma_2}_{\phi_2} [\mathbf{I}]^{\phi_2}_{\phi_1} + [r_{\sigma\sigma\sigma\sigma}]
\end{aligned}$$

References

Ahn, D.-H., Dittmar, R. F. & Gallant, A. R. (2002), 'Quadratic term structure models: Theory and evidence', *The Review of Financial Studies* **15**(1), 243–288.

Andreasen, M. M. (2012), 'On the effects of rare disasters and uncertainty shocks for risk premia in non-linear DSGE models', *Review of Economic Dynamics* **15**, 295–316.

Andreasen, M. M. (Forthcoming), 'Non-linear DSGE models and the central difference kalman filter', *Journal of Applied Econometrics* .

Andreasen, M. M. & Christensen, B. J. (2013), 'The SR approach: A new estimation procedure for non-linear and non-gaussian dynamic term structure models', *Working Paper* .

Andreasen, M. M., Fernandez-Villaverde, J. & Rubio-Ramirez, J. F. (2013), 'The Pruned State Space System for Non-Linear DSGE Models: Theory and Empirical Applications to Estimation', *NBER Working Paper* .

Andreasen, M. M. & Zabczyk, P. (2010), 'An efficient method of computing higher order bond price perturbation approximations', *Working Paper* .

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

Aruoba, S. B., Fernandez-Villaverde, J. & Rubio-Ramirez, J. F. (2006), 'Comparing solution methods for dynamic equilibrium economies', *Journal of Economic Dynamics and Control* **30**, 2477–2508.

Bauer, M. D. & Rudebusch, G. D. (2013), 'Monetary policy expectations at the zero lower bound', *Working Paper* .

Black, F. (1995), 'Interest rates as options', *The Journal of Finance* **50**(5), 1371–1376.

Christensen, J. H. E. & Rudebusch, G. D. (2013), 'Estimating shadow-rate term structure models with near-zero yields', *Federal Reserve Bank of San Francisco: Working Paper 2013-07* .

Clarida, R., Gali, J. & Gertler, M. (2000), 'Monetary policy rules and macroeconomic stability: Evidence and some theory', *The Quarterly Journal of Economics* **115**, 147–180.

Cox, J. C., Ingersoll, J. E. & Ross, S. A. (1985), 'A theory of the term structure of interest rates', *Econometrica* **53**(2), 385–407.

Dai, Q. & Singleton, K. J. (2000), 'Specification analysis of affine term structure models', *Journal of Finance* **55**, 1946–1978.

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

de Jong, F. (2000), 'Time series and cross-section information in affine term-structure models', *Journal of Business and Economic Statistics* **18**, 300–314.

Diebold, F. X. & Mariano, R. S. (1995), 'Comparing predictive accuracy', *Journal of Business and Economic Statistics* **13**, 253–263.

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

Duan, J.-C. & Simonato, J.-G. (1999), 'Estimating and testing exponential-affine term structure models by Kalman filter', *Review of Quantitative Finance and Accounting* **13**, 111–135.

Gorovoi, V. & Linetsky, V. (2004), 'Black's model of interest rates as options, eigenfunction expansions and japanese interest rates', *Mathematical Finance* **14**(1), 49–78.

Gürkaynak, R., Sack, B. & Wright, J. (2007), 'The U.S. treasury yield curve: 1961 to the present', *Journal of Monetary Economics* **54**, 2291–2304.

Harvey, D. I., Leybourne, S. J. & Newbold, P. (1997), 'Testing the equality of prediction mean squared errors', *International Journal of Forecasting* **13**, 281–291.

Hull, J. C. (2012), *Options, Futures, and Other Derivatives*, eighth edn, Pearson.

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

Ichiue, H. & Ueno, Y. (2013), 'Estimating term premia at the zero bound: An analysis of japanese, us, and uk yields', *Bank of Japan Working Paper Series* **13-E-8**.

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

Judd, K. L. & Guu, S.-M. (1997), 'Asymptotic methods for aggregate growth models', *Journal of Economic Dynamics and Control* **21**, 1025–1042.

Kamenik, O. (2005), 'Solving SDGE models: A new algorithm for the sylvester equation', *Computational Economics* **25**, 167–187.

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

Kim, D. H. & Singleton, K. J. (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* .

Leippold, M. & Wu, L. (2002), 'Asset pricing under the quadratic class', *The Journal of Financial and Quantitative Analysis* **37**(2), 271–295.

Pooter, M. D., Ravazzolo, F. & van Dijk, D. (2010), 'Term stucture forecasting using macro factors and forecast combination', *Board of Governors of the Federal Reserve System, Internationl Finance Discusion Paper* **993**.

Realdon, M. (2006), 'Quadratic term structure models in discrete time', *Finance Research Letters* **3**, 277–289.

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

Schmitt-Grohé, S. & Uribe, M. (2004), 'Solving dynamic general equilibrium models using a second-order approximation to the policy function', *Journal of Economic Dynamics and Control* **28**, 755–775.

Swanson, E., Anderson, G. & Levin, A. (2005), 'Higher-order perturbation solutions to dynamic, discrete-time rational expectations models', *Working Paper* .

Table 1: Estimation results: Two-factor models

Robust standard errors for elements in $\hat{\boldsymbol{\theta}}_{11}^{step3}$ are computed using (42) and (35) with $w_D = 10$ and $w_T = 10$. For elements in $\hat{\boldsymbol{\theta}}_{2}^{step3}$, robust standard errors are computed using (40) with $\mathbf{S}$ obtained by the Newey-West estimator with a bandwidth of 5. The estimates for the shadow rate model are obtained using a fourth-order perturbation approximation where the scaling of the bias correction is calibrated as described in Section 3.3 using a simulated time series of 2,000 observations. Estimates with one or two stars denote significance at the 5 percent and 1 percent level, respectively.

	ATSM		QTSM		Shadow rate	
	Estimate	SE	Estimate	SE	Estimate	SE
α	0.0082^{**}	0.0012	-	-	0.0093	0.0202
Ψ_{21}	-	-	0.9641^{**}	0.1775	-	-
Φ_{11}	0.0064^{**}	0.0011	0.0037	0.0056	0.0034	0.0163
Φ_{22}	0.0348^{**}	0.0010	0.0420^{*}	0.0203	0.0420	0.0626
μ_1	-	-	0.0686	0.0652	-	-
μ_2	-	-	0.0227^{*}	0.0112	-	-
$h_0(1,1)$	2.75×10^{-5}	3.57×10^{-5}	4.65×10^{-5}	3.98×10^{-4}	-1.24×10^{-5}	3.42×10^{-5}
$h_0(2,1)$	-7.07×10^{-5}	5.47×10^{-5}	2.53×10^{-5}	6.99×10^{-4}	-3.46×10^{-5}	5.41×10^{-5}
$h_x(1,1)$	1.0056^{**}	0.0080	0.9866^{**}	0.0066	0.9918^{**}	0.0064
$h_x(1,2)$	0.0370^{**}	0.0121	0.0330^{**}	0.0105	0.0226^{**}	0.0086
$h_x(2,1)$	-0.0159	0.0116	0.0242	0.0141	0.0062	0.0107
$h_x(2,2)$	0.9503^{**}	0.0173	0.9241^{**}	0.0228	0.9544^{**}	0.0194
Σ_{11}	$4.80^{**} \times 10^{-4}$	2.34×10^{-5}	0.0032^{**}	2.04×10^{-4}	$4.23^{**} \times 10^{-4}$	2.01×10^{-5}
Σ_{21}	$-3.67^{**} \times 10^{-4}$	2.88×10^{-5}	-0.0045^{**}	8.77×10^{-4}	$-3.56^{**} \times 10^{-4}$	5.11×10^{-5}
Σ_{22}	$3.97^{**} \times 10^{-4}$	5.64×10^{-5}	0.0058^{**}	6.70×10^{-4}	$6.45^{**} \times 10^{-4}$	8.69×10^{-5}

Table 2: Estimation results: Three-factor models

Robust standard errors for elements in $\hat{\boldsymbol{\theta}}_{11}^{step3}$ are computed using (42) and (35) with $w_D = 10$ and $w_T = 10$. For elements in $\hat{\boldsymbol{\theta}}_{2}^{step3}$, robust standard errors are computed using (40) with $\mathbf{S}$ obtained by the Newey-West estimator with a bandwidth of 5. The estimates for the shadow rate model are obtained using a fourth-order perturbation approximation where the scaling of the bias correction is calibrated as described in Section 3.3 using a simulated time series of 2,000 observations. Estimates with one or two stars denote significance at the 5 percent and 1 percent level, respectively.

	ATSM		QTSM		Shadow rate	
	Estimate	SE	Estimate	SE	Estimate	SE
α	0.0113^{**}	0.0016	-	-	0.0078^{**}	0.0007
Ψ_{12}	-	-	0.8513^{**}	0.0846	-	-
Ψ_{13}	-	-	0.8768	0.4848	-	-
Ψ_{23}	-	-	0.9858^{**}	0.1486	-	-
Φ_{11}	0.0026	0.0018	0.0010	0.0022	0.0098^{**}	0.0013
Φ_{22}	0.0318^{**}	0.0025	0.0372^{**}	0.0064	0.1280^{**}	0.0006
Φ_{33}	0.0996^{**}	0.0020	0.1349	0.1173	0.0092^{**}	0.0006
μ_1	-	-	0.0935	0.1738	-	-
μ_2	-	-	0.0422^{**}	0.0150	-	-
μ_3	-	-	0.0012	0.0082	-	-
$h_0(1,1)$	-3.45×10^{-5}	5.48×10^{-5}	-1.53×10^{-4}	3.54×10^{-4}	-0.0017	0.0010
$h_0(2,1)$	1.10×10^{-4}	1.12×10^{-4}	-0.0018	0.0011	-2.67×10^{-5}	5.63×10^{-5}
$h_0(3,1)$	-2.54×10^{-4}	1.33×10^{-4}	0.0044^{**}	0.0014	0.0017	0.0010
$h_x(1,1)$	0.9927^{**}	0.0075	0.9821^{**}	0.0081	0.3427	0.2587
$h_x(1,2)$	0.0313^{**}	0.0082	0.0184^{**}	0.0058	0.3914	0.6854
$h_x(1,3)$	0.0166	0.0148	0.0072	0.0135	-0.6019^{*}	0.2547
$h_x(2,1)$	0.0189	0.0152	0.0461^{**}	0.0170	-0.0026	0.0133
$h_x(2,2)$	0.9520^{**}	0.0247	1.0053^{**}	0.0193	0.8745^{**}	0.0294
$h_x(2,3)$	0.0535	0.0408	0.0774	0.0403	-0.0039	0.0132
$h_x(3,1)$	-0.0361^{*}	0.0173	-0.0695^{**}	0.0201	0.6497^{*}	0.2529
$h_x(3,2)$	0.0016	0.0235	-0.0531^{**}	0.0259	-0.3483	0.6690
$h_x(3,3)$	0.8581^{**}	0.0374	0.8111^{**}	0.0454	1.5949^{**}	0.2490
Σ_{11}	$3.72^{**} \times 10^{-4}$	2.33×10^{-5}	0.0025^{**}	1.77×10^{-4}	0.0160^{**}	0.0011
Σ_{21}	$-4.30^{**} \times 10^{-4}$	7.09×10^{-5}	-0.0025^{**}	6.18×10^{-4}	$-3.19^{**} \times 10^{-4}$	4.11×10^{-5}
Σ_{22}	$7.26^{**} \times 10^{-4}$	5.13×10^{-5}	0.0061^{**}	5.22×10^{-4}	$5.23^{**} \times 10^{-4}$	3.36×10^{-5}
Σ_{31}	$1.99^{**} \times 10^{-4}$	5.91×10^{-5}	6.36×10^{-4}	5.40×10^{-4}	-0.0158^{**}	0.0011
Σ_{32}	$-6.12^{**} \times 10^{-4}$	3.67×10^{-5}	-0.0061^{**}	6.47×10^{-4}	$-2.28^{**} \times 10^{-4}$	3.37×10^{-5}
Σ_{33}	$4.38^{**} \times 10^{-4}$	4.78×10^{-5}	0.0036^{**}	3.15×10^{-4}	$3.90^{**} \times 10^{-4}$	3.75×10^{-5}

Table 3: Estimation results: Four-factor models

Robust standard errors for elements in $\hat{\boldsymbol{\theta}}_{11}^{step3}$ are computed using (42) and (35) with $w_D = 10$ and $w_T = 10$. For elements in $\hat{\boldsymbol{\theta}}_{2}^{step3}$, robust standard errors are computed using (40) with $\mathbf{S}$ obtained by the Newey-West estimator with a bandwidth of 5. The estimates for the shadow rate model are obtained using a fourth-order perturbation approximation where the scaling of the bias correction is calibrated as described in Section 3.3 using a simulated time series of 2,000 observations. Estimates with one or two stars denote significance at the 5 percent and 1 percent level, respectively.

	ATSM		QTSM		Shadow rate	
	Estimate	SE	Estimate	SE	Estimate	SE
α	0.0088	0.0105	-	-	0.0134	0.0214
Ψ_{12}	-	-	0.8744^{**}	0.0335	-	-
Ψ_{13}	-	-	0.7860^{**}	0.0550	-	-
Ψ_{14}	-	-	0.8890^{**}	0.0227	-	-
Ψ_{23}	-	-	0.9800^{**}	0.0643	-	-
Ψ_{24}	-	-	0.9993^{**}	0.0127	-	-
Ψ_{34}	-	-	0.9735^{**}	0.0817	-	-
Φ_{11}	0.0087^{**}	0.0028	0.0007	0.0011	0.0360	0.0529
Φ_{22}	0.0757^{**}	0.0030	0.0380^{**}	0.0040	0.6022	0.9565
Φ_{33}	0.4415^{**}	0.0027	0.5665^{**}	0.0606	0.0619	0.1029
Φ_{44}	0.0086^{*}	0.0037	0.0644^{**}	0.0122	0.0015	0.0021
μ_1	-	-	0.1204	0.0897	-	-
μ_2	-	-	0.0191	0.0221	-	-
μ_3	-	-	0.0218^{**}	0.0032	-	-
μ_4	-	-	0.0063	0.0186	-	-
$h_0(1,1)$	-0.0122	0.0063	-0.0005	0.0006	1.60×10^{-4}	2.53×10^{-4}
$h_0(2,1)$	5.13×10^{-4}	7.49×10^{-5}	0.0027	0.0025	-4.32×10^{-4}	2.30×10^{-4}
$h_0(3,1)$	-1.81×10^{-4}	9.72×10^{-5}	0.0101^{**}	0.0021	-1.88×10^{-4}	2.76×10^{-4}
$h_0(4,1)$	0.0121	0.0062	-0.0041	0.0026	-3.55×10^{-5}	6.44×10^{-5}
$h_x(1,1)$	-5.0208^{**}	1.3957	0.9844^{**}	0.0080	0.9942^{**}	0.0449
$h_x(1,2)$	1.1644	2.8339	0.0162^{**}	0.00567	-0.2261^{**}	0.0667
$h_x(1,3)$	-15.06^{**}	3.3068	0.0292	0.0154	0.0952	0.0566
$h_x(1,4)$	-5.9616^{**}	1.3913	0.0085	0.0088	0.0162	0.0291
$h_x(2,1)$	0.0384^{**}	0.0149	0.0490	0.0331	-0.0024	0.0257
$h_x(2,2)$	0.9100^{**}	0.0327	1.0165^{**}	0.0363	0.5857^{**}	0.0565
$h_x(2,3)$	0.1664^{**}	0.0431	-0.2174^{**}	0.0720	0.0111	0.0347
$h_x(2,4)$	0.0382^{**}	0.0148	0.0973^{*}	0.0493	-0.0586^{*}	0.0258
$h_x(3,1)$	-0.0494^{**}	0.0165	-0.0412	0.0226	-0.0442	0.0469
$h_x(3,2)$	-0.0188	0.0348	-0.0174	0.0224	0.2780^{**}	0.0732
$h_x(3,3)$	0.6017^{**}	0.0566	0.61733^{**}	0.0627	0.8344^{**}	0.0587
$h_x(3,4)$	-0.0495^{**}	0.0166	-0.0138	0.0281	-0.0125	0.0314
$h_x(4,1)$	5.9955^{**}	1.3876	-0.0438	0.0359	0.0249^{**}	0.0076
$h_x(4,2)$	-1.1177	2.8174	-0.0405	0.0386	0.0513^{**}	0.0139
$h_x(4,3)$	15.02^{**}	3.2861	0.2902^{**}	0.0795	0.0160	0.0101
$h_x(4,4)$	6.9364^{**}	1.3831	0.8654^{**}	0.0491	0.9946^{**}	0.0072
Σ_{11}	0.0884^{**}	0.0061	0.0025^{**}	1.77×10^{-4}	0.0018^{**}	1.12×10^{-4}
Σ_{21}	$-5.97^{**} \times 10^{-4}$	6.16×10^{-5}	-0.0056^{**}	0.0011	$4.33^{**} \times 10^{-4}$	1.14×10^{-4}
Σ_{22}	$6.47^{**} \times 10^{-4}$	4.45×10^{-5}	0.0125^{**}	7.05×10^{-4}	$9.14^{**} \times 10^{-4}$	8.41×10^{-5}
Σ_{31}	$3.43^{**} \times 10^{-4}$	9.63×10^{-5}	-0.0016^{**}	5.68×10^{-4}	-0.0018^{**}	1.10×10^{-4}
Σ_{32}	$-5.15^{**} \times 10^{-4}$	7.47×10^{-5}	0.0020^{**}	7.31×10^{-4}	$-3.12^{**} \times 10^{-4}$	3.75×10^{-5}
Σ_{33}	$6.31^{**} \times 10^{-4}$	4.81×10^{-5}	0.0065^{**}	7.46×10^{-4}	$4.80^{**} \times 10^{-4}$	4.16×10^{-5}
Σ_{41}	-0.0880^{**}	0.0060	0.0045^{**}	0.0010	$-1.89^{**} \times 10^{-4}$	2.84×10^{-5}
Σ_{42}	$-1.89^{**} \times 10^{-4}$	3.85×10^{-5}	-0.0130^{**}	7.83×10^{-4}	-2.07×10^{-5}	1.66×10^{-5}
Σ_{43}	$-1.14^{**} \times 10^{-4}$	2.37×10^{-5}	-0.0026^{**}	2.42×10^{-4}	$-1.02^{**} \times 10^{-4}$	2.93×10^{-5}
Σ_{44}	$3.71^{**} \times 10^{-4}$	4.64×10^{-5}	0.0027^{**}	2.50×10^{-4}	$2.99^{**} \times 10^{-4}$	1.73×10^{-5}

Table 4: In-sample fit: The objective functions

This table reports $100\sqrt{Q_{1:T}^{step1}/2}$ and $100\sqrt{Q_{1:T}^{step3}/2}$ for the first and third step in the SR approach. Figures in bold highlight the best in-sample fit for a given estimation step.

	Step 1			Step 3		
	ATSM	QTSM	Shadow rate	ATSM	QTSM	Shadow rate
Two factors	11.50	**10.15**	11.37	11.59	**10.40**	11.42
Three factors	5.14	**4.81**	5.34	5.15	**4.83**	5.53
Four factors	2.40	**2.03**	**2.03**	2.40	**2.03**	2.04

Figure 1: The approximated max-function

The approximated max-function is given by $\widehat{\max}\{r, 0\} = a_0 + a_1 r + a_2 r^2 + a_3 r^3 + a_4 r^4$, where $a_0 = 0.00072$, $a_1 = 0.53$, $a_2 = 65.80$, $a_3 = -2661.14$, and $a_4 = 10000$. This specification is obtained by calibrating the approximated max-function to the true max-function on the displayed interval for annualized interest rates, i.e. 0.10 corresponds to an annual rate of 10 percent.

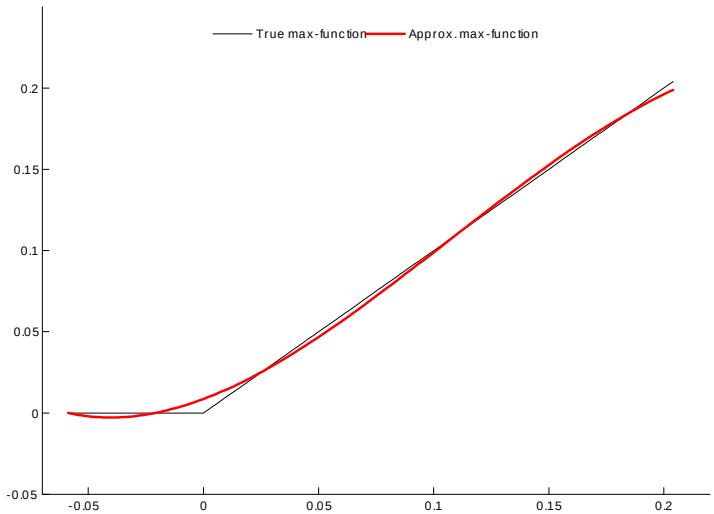


Figure 2: Accuracy of the perturbation approximation

The calibrated values of γ_0^* and γ_1^* are computed from simulated time series for the pricing factors of 2,000 observations using preliminary model estimates obtained with $\gamma_0 = 0$ and $\gamma_1 = 1$. The RMSEs by maturity are computed using the estimated pricing factors on US data. In both cases, the true solution for bond yields are obtained by Monte Carlo integration using 10,000 draws.

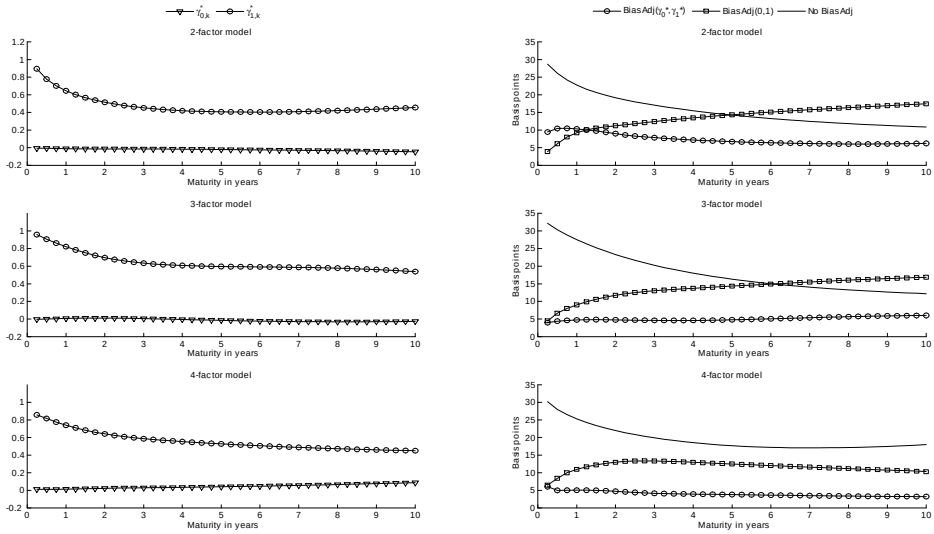


Figure 3: The three-factor shadow rate model: Bond yields and pricing errors

Bond yields and the pricing errors are computed using the pricing factors estimated from US data for the three-factor shadow rate model. In both cases, the true solution for bond yields are approximated by Monte Carlo integration using 10,000 draws.

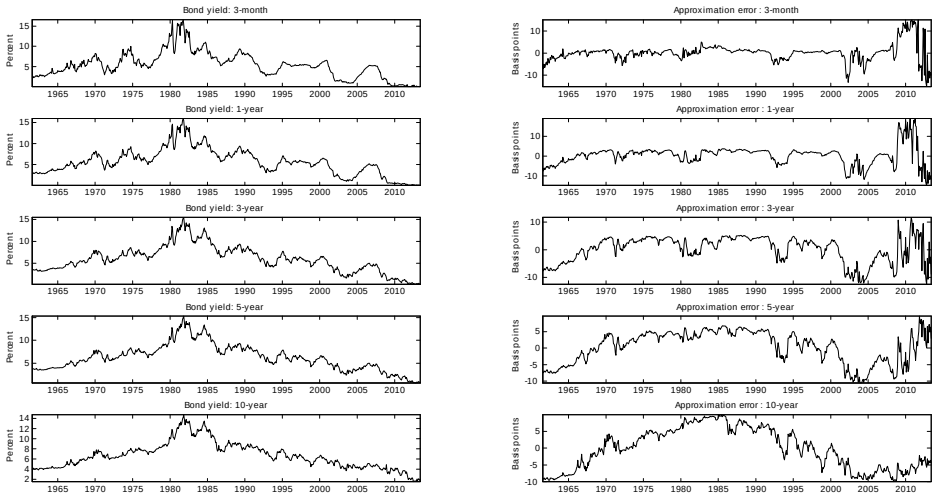


Figure 4: The in-sample fit

Charts in the first column report $100\sqrt{Q_{1:T}^{step3}/2}$. Charts in the second column report $100\sqrt{Q_{1998:T}^{step3}/2}$ and the final column reports σ_k. All charts are computed using the estimates obtained from the full sample.

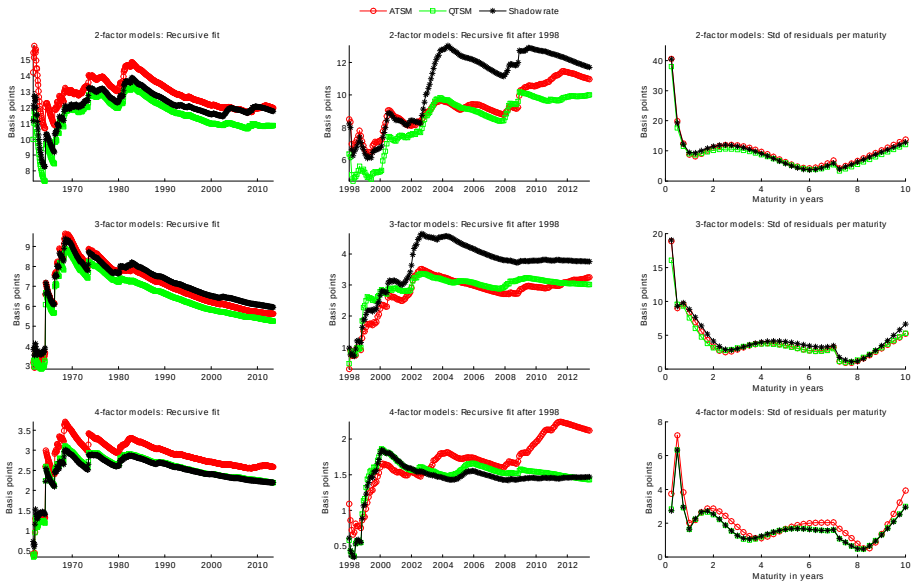


Figure 5: Unconditional means and standard deviations

All model-based moments are obtained from simulated time series of 500,000 observations using the estimates from Tables 1 to 3.

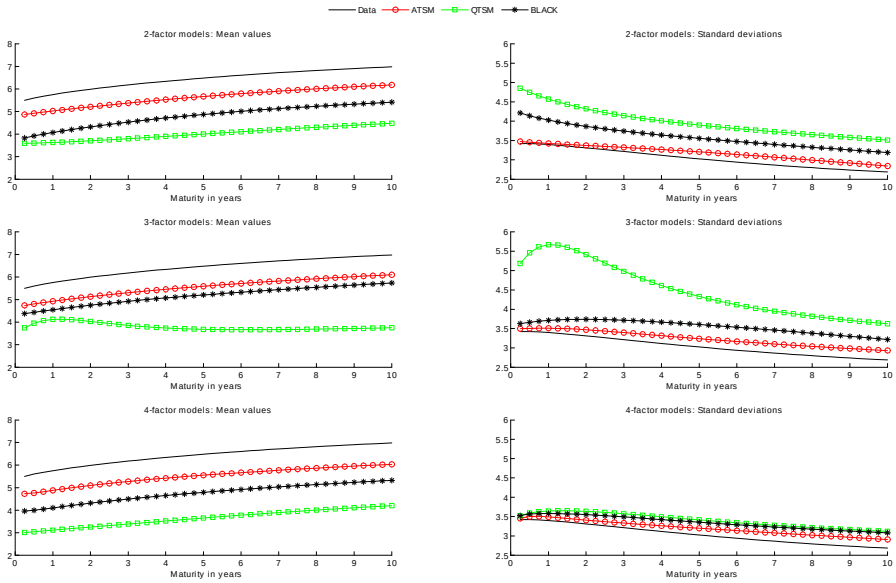


Figure 6: LPY tests for the P and Q dynamics

Charts in the first column show $\widehat{\phi}_k$ and charts in the second column report $\widehat{\phi}_k^{\mathbb{Q}}$. All model-based moments are obtained from simulated time series of 500,000 observations using the estimates from Tables 1 to 3.

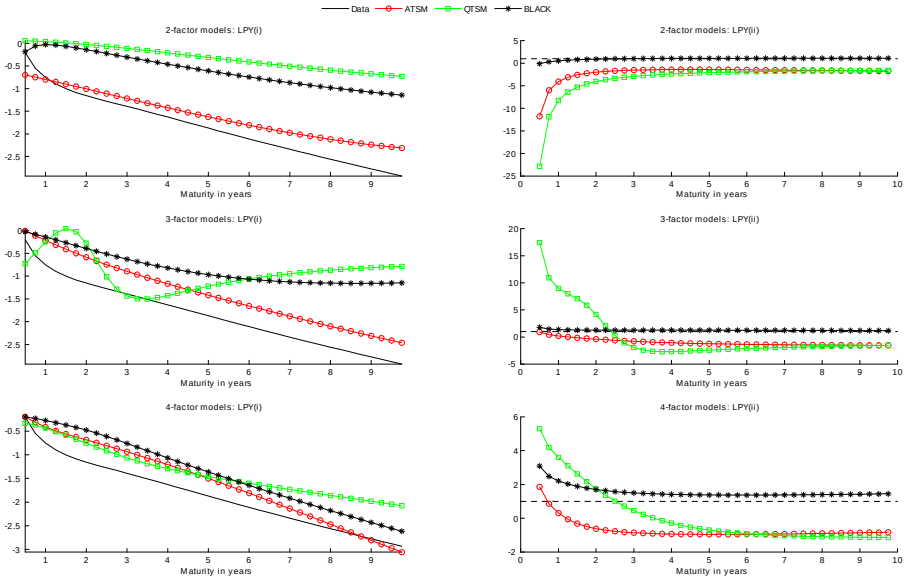


Figure 7: The 10-year term premia and the shadow rate

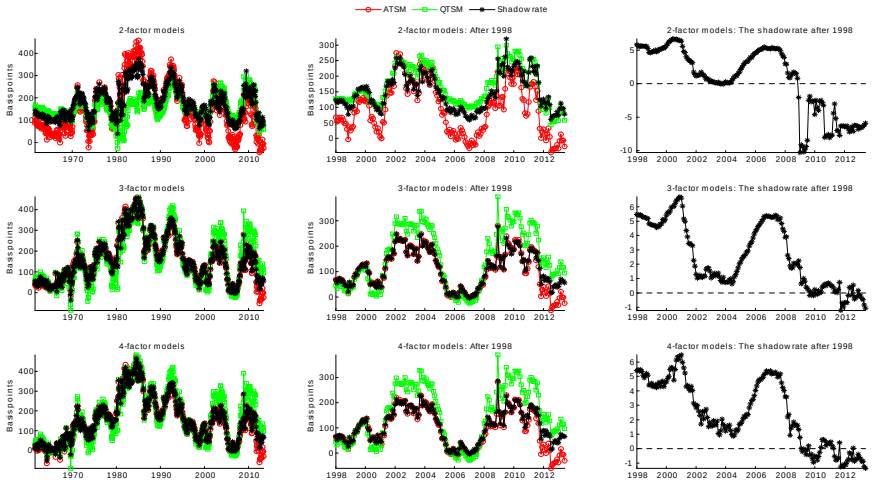
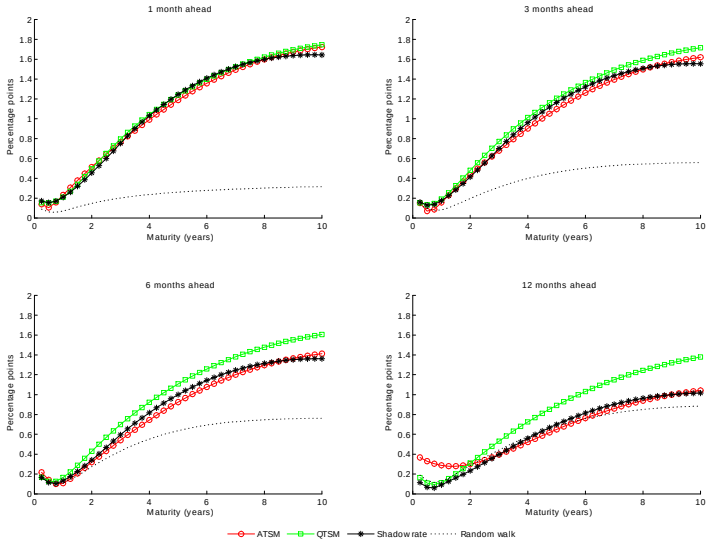


Figure 8: Out-of-sample forecasting: Three-factor models

This figure reports the root mean squared prediction errors (RMSPEs) for out-of-sample forecasts in three-factor models from January 2009 to May 2013. The RMSPEs are computed by estimating the three-factor models recursively and constructing 12-month ahead forecasts.