



Centre Interuniversitaire sur le Risque,
les Politiques Économiques et l'Emploi

Cahier de recherche/Working Paper **14-04**

Nonlinear Kalman Filtering in Affine Term Structure Models

Peter Christoffersen

Christian Dorion

Kris Jacobs

Lotfi Karoui

Janvier/January 2014

Christoffersen: University of Toronto, CBS and CREATES

Dorion: HEC Montréal and CIRPÉE

Jacobs: University of Houston and Tilburg University

Corresponding author: C.T. Bauer College of Business, University of Houston

kjacobs@bauer.uh.edu

Karoui: Goldman, Sachs & Co.

Dorion thanks IFM2 for financial support. Christoffersen is supported by SSRHC and Bank of Canada. We would like to thank Luca Benzoni, Bob Kimmel, two anonymous referees, an associate editor, and the editor (Wei Xiong) for helpful comments. Any remaining inadequacies are ours alone. The views, if any, expressed in this paper are those of the authors, not the institutions they are affiliated to.

Abstract:

The extended Kalman filter, which linearizes the relationship between security prices and state variables, is widely used in fixed income applications. We investigate if the unscented Kalman filter should be used to capture nonlinearities, and compare the performance of the Kalman filter to that of the particle filter. We analyze the cross section of swap rates, which are mildly nonlinear in the states, and cap prices, which are highly nonlinear. When caps are used to filter the states, the unscented Kalman filter significantly outperforms its extended counterpart. The unscented Kalman filter also performs well when compared to the much more computationally intensive particle filter. These findings suggest that the unscented Kalman filter may prove to be a good approach for variety of problems in fixed income pricing.

Keywords: Kalman filtering, nonlinearity, term structure models, swaps, caps

JEL Classification: G12

1 Introduction

Multifactor affine term structure models (ATSMs) have become the standard in the literature on the valuation of fixed income securities, such as government bonds, corporate bonds, interest rate swaps, credit default swaps, and interest rate derivatives. Even though we have made significant progress in specifying these models, their implementation is still subject to substantial challenges.

One of the challenges is the proper identification of the parameters governing the dynamics of the risk premia (see Dai and Singleton (2002)). It has been recognized in the literature that the use of contracts that are nonlinear in the state variables, such as interest rate derivatives, can potentially help achieve such identification. Nonlinear contracts can also enhance the ability of affine models to capture time variation in excess returns and conditional volatility (see Bikbov and Chernov (2009) and Almeida, Graveline and Joslin (2011)).

Given the potentially valuable information content of nonlinear securities, efficient implementation of ATSMs for these securities is of paramount importance. One of the most popular techniques used in the literature, the extended Kalman filter (EKF), relies on a linearized version of the measurement equation, which links observed security prices to the models' state variables. Our paper is the first to extensively investigate the impact of this linearization. We find that when the states are extracted using securities that are mildly nonlinear in the state variables, such as swaps, the EKF is adequate. When extracting states using securities that are highly nonlinear in the state variables, such as interest rate derivatives, the approximation used by the EKF leads to significant noise and biases in the filtered state variables as well as the forecasts of security prices. We extensively analyze the properties of the unscented Kalman filter (UKF) when implementing affine term structure models for highly nonlinear securities, and we compare its performance with that of the EKF. The main advantage of the unscented Kalman filter is that it accounts for the non-linear relationship between the observed security prices and the underlying state variables.

We use an extensive Monte Carlo experiment that involves a cross-section of LIBOR and swap

rates as well as interest rate caps, to investigate the quality of the EKF and UKF filters as well as their in- and out-of-sample forecasting ability. We also benchmark the performance of the EKF and the UKF using the particle filter. The performance of both filters is roughly similar when interest rate caps are not included in the filtering exercise, and their performance compares well with that of the particle filter. When interest rate caps are included in the filtering exercise, the UKF robustly outperforms the EKF in two different exercises. First, for the purpose of filtering the unobserved state variables, the UKF outperforms the EKF, using the root-mean-square-error (RMSE) of the filtered state variables as a gauge for the performance of the filters. We also find that the UKF is numerically much more stable than the EKF, exhibiting a much lower dispersion of the RMSE across the Monte Carlo trajectories. Second, the improved precision of the UKF in filtering the state variables in this case translates into more accurate forecasts for LIBOR rates, swap rates, and cap prices. It is also critically important that the superior performance of the UKF comes at a reasonable computational cost. In our application, the computational time required for the unscented Kalman filter was about twice that required for the extended Kalman filter. Finally, when interest rate caps are included in the filtering exercise, the UKF also performs well when compared to the particle filter for most models and securities, but its relative performance worsens somewhat for highly nonlinear securities. Moreover, in our implementation, the UKF is approximately 500 times faster than the particle filter.

Throughout this paper we keep the structural parameters fixed at their true values. However, the poor results obtained when using the EKF to filter states when caps are used suggest that parameter estimation based on this technique would be highly unreliable, as the filter is unable to correctly fit rates and prices even when provided with the true model parameters. The dramatic improvements brought by the UKF suggest that it will also improve parameter estimation when derivative prices are used to estimate the parameters, which is of critical importance in the identification of the risk premium parameters.

Even though the use of the unscented Kalman filter has become popular in the engineering literature (see for instance Julier (2000) and Julier and Uhlmann (2004)), it has not been used extensively in the empirical asset pricing literature.¹ Our results suggest that the unscented Kalman filter may prove to be a good approach for tackling a number of problems in fixed income pricing where the relationship between the state vector and the observations is highly nonlinear. Given the modest computational burden, the UKF may also represent a viable alternative in mildly nonlinear fixed income problems, such as the estimation of term structure models using a cross section of coupon bonds or the estimation of quadratic term structure models.²

The paper proceeds as follows. Section 2 briefly discusses the pricing of LIBOR, swaps, and caps in affine term structure models. Section 3 discusses Kalman filtering in ATSMs, including the extended Kalman filter and the unscented Kalman filter, and also discusses our implementation of the particle filter. Section 4 reports the results of our Monte Carlo experiments. Section 5 discusses implications for parameter estimation, and Section 6 concludes.

2 Affine Term Structure Models

In this section, we define the risk-neutral dynamics in ATSMs, a pricing kernel and the pricing formulas for LIBOR rates, swap rates, and cap prices. We follow the literature on term structure models and assume that the swap and LIBOR contracts as well as the interest rate caps are default-free. See Dai and Singleton (2000), Collin-Dufresne and Solnik (2001), Feldhutter and Lando (2008), and Andersen and Benzoni (2010) for further discussion.

¹See Carr and Wu (2007), Bakshi, Carr and Wu (2008), and Li (2013) for applications to equity options. van Binsbergen and Koijen (2012) use the unscented Kalman filter to estimate present-value models.

²See Fontaine and Garcia (2012) for a recent application of the unscented Kalman filter to the estimation of term structure models for coupon bonds. See Chen, Cheng, Fabozzi, and Liu (2008) for an application of the unscented Kalman filter to the estimation of quadratic term structure models.

2.1 Risk-Neutral Dynamics

Affine term structure models (ATSMs) assume that the short rate is given by $r_t = \delta_0 + \delta_1' x_t$, and the state vector x_t follows an affine diffusion under the risk-neutral measure Q

$$dx_t = \tilde{\kappa} (\tilde{\theta} - x_t) dt + \Sigma \sqrt{S_t} d\tilde{W}_t, \quad (1)$$

where \tilde{W}_t is an N -dimensional vector of independent standard Q -Brownian motions, $\tilde{\kappa}$ and Σ are $N \times N$ matrices and S_t is a diagonal matrix with a i th diagonal element given by

$$[S_t]_{ii} = \alpha_i + \beta_i' x_t. \quad (2)$$

Following Duffie and Kan (1996), we write

$$\begin{aligned} \psi^Q(u, t, \tau) &= E_t^Q \left[e^{-\int_t^{t+\tau} r_s ds} e^{u' x_t} \right] \\ &= \exp \{ A_u(\tau) - B_u'(\tau) x_t \}, \end{aligned} \quad (3)$$

where τ is the time to maturity, and $A_u(\tau)$ and $B_u(\tau)$ satisfy the following Ricatti ODEs

$$\frac{dA_u(\tau)}{d\tau} = -\tilde{\theta}' \tilde{\kappa} B_u(\tau) + \frac{1}{2} \sum_{i=1}^N [\Sigma B_u(\tau)]_i^2 \alpha_i - \delta_0 \quad (4)$$

and

$$\frac{dB_u(\tau)}{d\tau} = -\tilde{\kappa} B_u(\tau) + \frac{1}{2} \sum_{i=1}^N [\Sigma B_u(\tau)]_i^2 \beta_i + \delta_1. \quad (5)$$

Equations (4) and (5) can be solved numerically with initial conditions $A_u(0) = 0$ and $B_u(0) = -u$.

The resulting zero-coupon bond price is exponentially affine in the state vector

$$P(t, \tau) = \psi^Q(0, t, \tau) = \exp \{ A_0(\tau) - B_0'(\tau) x_t \}. \quad (6)$$

2.2 Pricing Kernel

The model is fully specified once the dynamics of the state price are known. The dynamic of the pricing kernel π_t is assumed to be of the form

$$\frac{d\pi_t}{\pi_t} = -r_t dt - \Lambda_t' dW_t, \quad (7)$$

where W_t is a N -dimensional vector of independent standard P -Brownian motions and Λ_t denotes the market price of risk. The dynamics of the state vector under the actual measure P can be obtained by subtracting $\Sigma\sqrt{S_t}\Lambda_t$ from the drift of equation (1).

The market price of risk Λ_t does not depend on the maturity of the bond and is a function of the current value of the state vector x_t . We study completely affine models which specify the market price of risk as follows

$$\Lambda_t = \sqrt{S_t}\lambda_0. \quad (8)$$

See Cheridito, Filipović and Kimmel (2007), Duffee (2002), and Duarte (2004) for alternative specifications of the market price of risk.

2.3 LIBOR and Swap Rates

In ATSMs, the time- t LIBOR rate on a loan maturing at $t + \tau$ is given by

$$\begin{aligned} L(t, \tau) &= \frac{1 - P(t, \tau)}{\tau P(t, \tau)} \\ &= \tau^{-1} [\exp(-A_0(\tau) + B'_0(\tau)x_t) - 1]. \end{aligned} \quad (9)$$

while the fair rate at time t on a swap contract with semi-annual payments up to maturity $t + \tau$ can be written as

$$\begin{aligned} SR(t, \tau) &= \frac{1 - P(t, \tau)}{0.5 \times \sum_{j=1}^{2\tau} P(t, 0.5j)} \\ &= \frac{1 - \exp(A_0(\tau) - B'_0(\tau)x_t)}{0.5 \times \sum_{j=1}^{2\tau} \exp(A_0(0.5j) - B'_0(0.5j)x_t)}. \end{aligned} \quad (10)$$

As mentioned earlier, $A_0(\tau)$ and $B_0(\tau)$ can be obtained numerically from equations (4) and (5).

2.4 Cap Prices

Computing cap prices is more computationally intensive. Given the current latent state x_0 , the value of an at-the-money cap C_L on the 3-month LIBOR rate $L(t, 0.25)$ with strike $\bar{R} = L(0, 0.25)$

and maturity in T years is

$$C_L(0, T, \bar{R}) = \sum_{j=2}^{T/0.25} E^Q \left[e^{-\int_0^{T_j} r_s ds} 0.25 \left(L(T_{j-1}, 0.25) - \bar{R} \right)^+ \right] = \sum_{j=2}^{T/0.25} c_L(0, T_j, \bar{R}), \quad (11)$$

where $T_j = 0.25j$. The cap price is thus the sum of the value of caplets $c_L(0, T_j, \bar{R})$ with strike \bar{R} and maturity T_j .

The payoff $\Pi_{T_{j-1}}$ of caplet $c_L(0, T_j, \bar{R})$ is known at time T_{j-1} but paid at T_j . It is given by

$$\begin{aligned} \Pi_{T_{j-1}} &= 0.25 \left(L(T_{j-1}, 0.25) - \bar{R} \right)^+ \\ &= 0.25 \left(\frac{1 - P(T_{j-1}, 0.25)}{0.25 P(T_{j-1}, 0.25)} - \bar{R} \right)^+ \\ &= \frac{1 + 0.25\bar{R}}{P(T_{j-1}, 0.25)} \left(\frac{1}{1 + 0.25\bar{R}} - P(T_{j-1}, 0.25) \right)^+. \end{aligned} \quad (12)$$

Since the discounted value of the caplet is a martingale under the risk-neutral measure, we have

for $K = \frac{1}{1 + 0.25\bar{R}}$

$$\begin{aligned} c_L(0, T_j, \bar{R}) &= E^Q \left[e^{-\int_0^{T_j} r_s ds} \Pi_{T_{j-1}} \right] \\ &= \frac{1}{K} E^Q \left[e^{-\int_0^{T_{j-1}} r_s ds} \left(K - P(T_{j-1}, 0.25) \right)^+ \right] \\ &= \frac{1}{K} \mathcal{P}(0, T_{j-1}, T_j, K) \end{aligned} \quad (13)$$

Equation (13) represents the time-0 value of $1/K$ puts with maturity T_{j-1} and strike K on a zero-coupon bond maturing in T_j years. Duffie, Pan, and Singleton (2000) show that the price of such a put option is given by

$$\begin{aligned} \mathcal{P}(0, T_{j-1}, T_j, K) &= E^Q \left[e^{-\int_0^{T_{j-1}} r_s ds} \left(K - \exp \{ A_0(0.25) - B'_0(0.25)x_{T_{j-1}} \} \right)^+ \right] \\ &= e^{A_0(0.25)} E^Q \left[e^{-\int_0^{T_{j-1}} r_s ds} \left(e^{-A_0(0.25)} K - \exp \{ -B'_0(0.25)x_{T_{j-1}} \} \right)^+ \right] \\ &= e^{A_0(0.25)} \left[c G_{0,d}(\log c, 0, T_{j-1}) - G_{d,d}(\log c, 0, T_{j-1}) \right], \end{aligned} \quad (14)$$

where $c = e^{-A_0(0.25)} K$, $d = -B'_0(0.25)$, and

$$G_{a,b}(y, 0, T_{j-1}) = \frac{1}{2} \psi^Q(a, 0, T_{j-1}) - \frac{1}{\pi} \int_0^\infty \frac{1}{\varkappa} \text{Im} [\psi^Q(a + i\varkappa b, 0, T_{j-1}) e^{-i\varkappa y}] d\varkappa \quad (15)$$

In general, the integral in (15) can only be solved numerically. Note that this requires solving the Ricatti ODEs for $A_u(\tau)$ and $B_u(\tau)$ in (4) and (5) at each point $u = a + ixb$.

Empirical studies of cap pricing and hedging can be found in Li and Zhao (2006) and Jarrow, Li and Zhao (2007).

At this point we have specified the state variable dynamics and the pricing kernel. We have also provided model-based formulas for LIBOR and swaps rates as well as cap prices. We are now ready to discuss methods for filtering the unobserved state variables using observed rates and prices.

3 Kalman Filtering the State Vector

Consider the following general nonlinear state-space system

$$x_{t+1} = F(x_t, \epsilon_{t+1}), \quad (16)$$

and

$$y_t = G(x_t) + u_t \quad (17)$$

where y_t is a D -dimensional vector of observables, ϵ_{t+1} is state noise and u_t is observation noise that has zero mean and a covariance matrix denoted by R . In term structure applications, the transition function F is specified by the dynamic of the state vector and the measurement function G is specified by the pricing function of the fixed income securities being studied. In our application, the transition function F follows from the affine state vector dynamic in (1), y_t are the LIBOR, swap rates, and cap prices observed weekly for different maturities, and the function G is given by the pricing functions in (9), (10), and (11).

The transition equation (16) reflects the discrete time evolution of the state variables, whereas the measurement equation provides the mapping between the unobserved state vector and the observed variables. If x_t is an affine diffusion process, a discrete expression of its dynamics is only available for Gaussian processes. When the state vector is not Gaussian, one can obtain an

approximate transition equation by exploiting the existence of the two first conditional moments in closed-form and replacing the original state vector with a new Gaussian state vector with identical two first conditional moments. While this approximation results in inconsistent estimates, Monte Carlo evidence shows that its impact is negligible in practice (see Duan and Simonato (1999) and de Jong (2000)).

3.1 The Extended Kalman Filter

The most popular implementation of the Kalman filter in economics and finance is the extended Kalman filter (EKF).³ For the models we are interested in, the conditional expectation of the state vector is an affine function of the state.⁴ Using (1) and an Euler discretization, the transition equation (16) can therefore be rewritten as follows

$$x_{t+1} = F(x_t, \epsilon_{t+1}) = a + bx_t + \epsilon_{t+1}, \quad (18)$$

where $\epsilon_{t+1|t} \sim \mathcal{N}(0, v(x_t))$ and $v(x_t)$ is the conditional covariance matrix of the state vector.

Given that y_t is observed and assuming that it is a Gaussian random variable, the Kalman filter recursively provides the optimal minimum MSE estimate of the state vector. The prediction step consists of

$$x_{t|t-1} = a + bx_{t-1|t-1}, \quad (19)$$

$$P_{xx(t|t-1)} = bP_{xx(t-1|t-1)}b' + v(x_{t-1|t-1}) \quad (20)$$

$$K_t = P_{xy(t|t-1)}P_{yy(t|t-1)}^{-1}, \quad (21)$$

and

$$y_{t|t-1} = E_{t-1}[G(x_t)]. \quad (22)$$

³For early applications of the extended Kalman filter in finance see Chen and Scott (1995), Duan and Simonato (1999), Duffee (1999), and Lund (1997).

⁴It is straightforward to derive analytical expressions for the first two conditional moments. See the internet appendix available at <http://papers.ssrn.com/abstract=2322760>.

Updating is done using

$$x_{t|t} = x_{t|t-1} + K_t (y_t - y_{t|t-1}), \quad (23)$$

and

$$P_{xx(t|t)} = P_{xx(t|t-1)} - K_t P_{yy(t|t-1)} K_t', \quad (24)$$

When G in (22) is a linear function, e.g. if the observations are zero-coupon yields, then the covariance matrices $P_{xy(t|t-1)}$ and $P_{yy(t|t-1)}$ can be computed exactly and the only approximation is therefore induced by the Gaussian transformation of the state vector used in (18). When the relationship between the state vector and the observation is nonlinear, as is the case when swap contracts, coupon bonds, or interest rate options are used, then $G(x_t)$ needs to be well approximated in order to obtain good estimates of the covariance matrices $P_{xy(t|t-1)}$ and $P_{yy(t|t-1)}$. The most important difference between the unscented and extended Kalman filter is their treatment of nonlinearities in the measurement equation. The extended Kalman filter (EKF) relies on a first order Taylor expansion of the measurement equation around the predicted state $x_{t|t-1}$. The measurement equation is then rewritten as follows

$$y_t = G(x_{t|t-1}) + J_t (x_t - x_{t|t-1}) + u_t, \quad (25)$$

where

$$J_t = \left. \frac{\partial G}{\partial x_t} \right|_{x_t=x_{t|t-1}}$$

denotes the Jacobian matrix of the nonlinear function $G(x_{t|t-1})$ computed at $x_{t|t-1}$.

The covariance matrices $P_{xy(t|t-1)}$ and $P_{yy(t|t-1)}$ are then computed as

$$P_{xy(t|t-1)} = P_{xx(t|t-1)} J_t, \quad (26)$$

and

$$P_{yy(t|t-1)} = J_t P_{xx(t|t-1)} J_t' + R. \quad (27)$$

The estimate of the state vector is then updated using (23), (24), and

$$K_t = P_{xx(t|t-1)} J_t P_{yy(t|t-1)}^{-1}. \quad (28)$$

3.2 The Unscented Kalman Filter

Unlike the extended Kalman filter, the unscented Kalman filter uses the exact nonlinear function $G(x_t)$ and does not linearize the measurement equation. Rather than approximating $G(x_t)$, the unscented Kalman filter approximates the conditional distribution of x_t using the scaled unscented transformation (Julier (2000)), which can be defined as a method for computing the statistics of a nonlinear transformation of a random variable. Julier and Uhlmann (2004) prove that such an approximation is accurate to the third order for Gaussian states and to the second order for non-Gaussian states. It must also be noted that the approximation does not require computation of the Jacobian or Hessian matrices and that the computational burden associated with the unscented Kalman is comparable to that of the extended Kalman filter. In our application below, the computation time for the unscented Kalman filter is on average twice that of the extended Kalman filter.

Consider the random variable x with mean μ_x and covariance matrix P_{xx} , and the nonlinear transformation $y = G(x)$. The basic idea behind the scaled unscented transformation is to generate a set of points, called sigma points, with the first two sample moments equal to μ_x and P_{xx} . The nonlinear transformation is then applied at each sigma point. In particular, the N -dimensional random variable is approximated by a set of $2N + 1$ weighted points given by

$$\mathcal{X}_0 = \mu_x, \tag{29}$$

$$\mathcal{X}_i = \mu_x + \left(\sqrt{(N + \xi) P_{xx}} \right)_i, \text{ for } i = 1, \dots, N \tag{30}$$

$$\mathcal{X}_i = \mu_x - \left(\sqrt{(N + \xi) P_{xx}} \right)_i, \text{ for } i = N + 1, \dots, 2N \tag{31}$$

with weights

$$W_0^m = \frac{\xi}{(N + \xi)}, \quad W_0^c = \frac{\xi}{(N + \xi)} + (1 - \rho^2 + \gamma) \tag{32}$$

$$W_i^m = W_i^c = \frac{1}{2(N + \xi)}, \text{ for } i = 1, \dots, 2N, \tag{33}$$

where $\xi = \rho^2 (N + \chi) - N$, and where $\left(\sqrt{(N + \xi) P_{xx}}\right)_i$ is the i th column of the matrix square root of $(N + \xi) P_{xx}$. The scaling parameter $\rho > 0$ is intended to minimize higher order effects and can be made arbitrary small. The restriction $\chi > 0$ guarantees the positivity of the covariance matrix. The parameter $\gamma \geq 0$ can capture higher order moments of the state distribution; it is equal to two for the Gaussian distribution. The nonlinear transformation is applied to the sigma points (29)-(31), thus generating

$$\mathcal{Y}_i = G(\mathcal{X}_i), \text{ for } i = 0, \dots, 2N.$$

The unscented Kalman filter relies on the unscented transformation to approximate the covariance matrices $P_{xy(t|t-1)}$ and $P_{yy(t|t-1)}$. The state vector is augmented with the state white-noise $\xi_t = \epsilon_t / \sqrt{v(x_t)}$.⁵ With N state variables this results in the $N_a = 2N$ dimensional vector

$$x_t^a = [x_t' \ \xi_t']'. \quad (34)$$

The unscented transformation is applied to this augmented vector in order to compute the sigma points.

As shown by equations (30) and (31), implementation of the unscented Kalman filter requires computation of the square root of the variance-covariance matrix of the augmented state. There is no guarantee that the variance-covariance matrix will be positive definite. Positive definiteness of the variance-covariance matrix is also not guaranteed with the extended Kalman filter which in turn can affect its numerical stability. In the unscented case, a more stable algorithm is provided by the square-root unscented Kalman filter proposed by van der Merwe and Wan (2002). The basic intuition behind the square-root implementation of the unscented Kalman filter is to propagate and update the square-root of the variance-covariance matrix rather than the variance-covariance

⁵The state vector could further be augmented with the measurement noise u_t . However, the models we consider have additive measurement errors that can easily be dealt with by simply including the covariance matrix R of the error term when updating the conditional covariance matrix of y . This reduces the number of sigma points by $2D$ and considerably speeds up computation. We are grateful to an anonymous referee for pointing this out.

matrix itself.⁶

If we denote the square-root matrix of P by S , the square-root implementation of the unscented Kalman filter can be summarized by the following algorithm:

0. Initialize the algorithm at $t = 0$ using unconditional moments

$$x_{0|0}^a = [E[x_t] \quad 0]' \quad S_{xx(0|0)} = \sqrt{\text{var}[x_t]} \quad S_{\xi\xi(0|0)} = I$$

For $t = 1, \dots, T$:

1. Compute the $2N_a + 1$ sigma points:

$$\mathcal{X}_{t-1|t-1}^a = \begin{bmatrix} x_{t-1|t-1}^a & x_{t-1|t-1}^a \pm \sqrt{(N_a + \xi)S_{t-1|t-1}^a} \end{bmatrix},$$

where $S_{t-1|t-1}^a$ is a block-diagonal matrix with $S_{xx(t-1|t-1)}$ and $S_{\xi\xi(t-1|t-1)}$ on the diagonal.

2. Prediction step:

$$\begin{aligned} \mathcal{X}_{t|t-1}^x &= a + b\mathcal{X}_{t-1|t-1}^x + \sqrt{v(\mathcal{X}_{t-1|t-1}^x)'} \mathcal{X}_{t-1|t-1}^\xi, & x_{t|t-1} &= \sum_{i=0}^{2N} W_i^m \mathcal{X}_{i,t|t-1}^x, \\ S_{xx(t|t-1)} &= \text{CHOLUPDATE} \left\{ \text{QR} \left\{ \sqrt{W_1^c} \left(\mathcal{X}_{1:2N,t|t-1}^x - x_{t|t-1} \right) \sqrt{v(x_{t|t-1})} \right\}, \mathcal{X}_{0,t|t-1}^x - x_{t|t-1}, W_0^c \right\}, \\ \mathcal{Y}_{t|t-1} &= G \left(\mathcal{X}_{t|t-1}^x \right) & \text{and} & \quad y_{t|t-1} = \sum_{i=0}^{2N} W_i^m \mathcal{Y}_{i,t|t-1}, \end{aligned}$$

where (i) $\text{QR}\{A\}$ returns the Q matrix from the ‘QR’ orthogonal-triangular decomposition of A , and (ii) $S_{cu} = \text{CHOLUPDATE}\{S, z, \pm\nu\}$ is a rank-1 update (or downdate) to Cholesky factorization: Assuming that S is the Cholesky factor of P , then S_{cu} is the Cholesky factor of $P \pm \nu z z'$. If z is a matrix, the update (or downdate) is performed using the columns of z sequentially.

⁶While the square-root implementation of the unscented Kalman filter is numerically more stable, its computational complexity is similar to that of the original unscented Kalman.

3. Measurement update:

$$P_{xy(t|t-1)} = \sum_{i=0}^{2N} W_i^c \left[\mathcal{X}_{i,t|t-1}^x - x_{t|t-1} \right] \left[\mathcal{Y}_{i,t|t-1} - y_{t|t-1} \right]'$$

$$S_{yy(t|t-1)} = \text{CHOLUPDATE} \left\{ \text{QR} \left\{ \sqrt{W_1^c} (\mathcal{Y}_{1:2N,t|t-1} - y_{t|t-1}) \sqrt{R} \right\}, \mathcal{Y}_{0,t|t-1} - y_{t|t-1}, W_0^c \right\}$$

where R is the variance of the measurement error. Then

$$K_t = \left(P_{xy(t|t-1)} / S'_{yy(t|t-1)} \right) / S_{yy(t|t-1)}, \quad x_{t|t} = x_{t|t-1} + K_t (y_t - y_{t|t-1}),$$

$$S_{xx(t|t)} = \text{CHOLUPDATE} \left\{ S_{xx(t|t-1)}, K_t S_{yy(t|t-1)}, -1 \right\},$$

where ‘/’ denotes the efficient least-squares solution to the problem $Ax = b$. The estimate of the state vector is then updated using equations (23) and (24).

3.3 The Particle Filter

To benchmark the performance of the UKF and the EKF, we also investigate the performance of the particle filter (PF). Our implementation of the PF follows Li (2011) who proposes a sequential particle filter with a proposal density that is generated by the unscented Kalman filter.⁷ The algorithm is as follows:

0. Initialize at $t = 0$. Draw a set of particles $\{x_0^{(i)}, i = 1, \dots, N_p\}$ from the prior $p(x_0)$ and give each particle a weight of one.

For $t = 1, \dots, T$:

1. Apply the square-root UKF update to x_{t-1} , as discussed in the previous section, obtain a (global) proposal distribution $\tilde{\pi}$ with mean \bar{x}_t and variance $S'_t S_t$.⁸

⁷We are grateful to an anonymous referee for suggestions that considerably reduced the computational demands of this algorithm in our framework.

⁸Li’s (2011) algorithm is developed in order to address a general non-Gaussian model. In this context, the UKF

2. Sample $\{x_t^{(i)}, i = 1, \dots, N_p\}$ from $\tilde{\pi}(x_t^{(i)}|y_{1:t}) = \mathcal{N}(\bar{x}_t, S'_t S_t)$. Note that this can be done efficiently by drawing three-dimensional standard normal variables $\epsilon_t^{(i)}$ and setting $x_t^{(i)} = \bar{x}_t + S'_t \epsilon_t^{(i)}$.

3. Update the weight of each particle according to

$$w_t^{(i)} = w_{t-1}^{(i)} \frac{p(y_t|x_t^{(i)})p(x_t^{(i)}, x_{t-1}^{(i)})}{\tilde{\pi}(x_t^{(i)}|y_{1:t})} \quad (35)$$

and normalize the weights, i.e. $\tilde{w}_t^{(i)} = w_t^{(i)} / \sum_{i=1}^{N_p} w_t^{(i)}$.

4. Resample the weights (Liu and Chen (1998))

- Retain $N'_i = \lfloor N_p \tilde{w}_t^{(i)} \rfloor$ copies of $x_t^{(i)}$;
- Uniformly sample the remaining $N''_i = N_p - \sum_{i=1}^{N_p} N'_i$ with new weights $(N_p \tilde{w}_t^{(i)} - N'_i) / N''_i$;
- Reset the weights to $1/N_p$.

4 Monte Carlo Analysis

We are now ready to conduct a Monte Carlo study comparing the UKF and the EKF to the benchmark PF for the purpose of filtering states and forecasting fixed income security rates and prices. We want to focus on the numerical stability of each filtering algorithm and the potential biases caused by a linear approximation of LIBOR, swap rates, and especially cap prices, which are very nonlinear in the states. By design, the comparison below is not affected by the issue of parameter estimation as we keep the parameters fixed at their true values throughout.

update must be applied to each particle. In our framework, given conditional normality, the UKF provides a good global proposal. This requires pricing each instrument only $(2N_a + 1) + N_p$ times at a given time step, rather than $(2N_a + 1) \times N_p$, where $(2N_a + 1)$ is the number of sigma points used in the UKF and N_p is the number of particles used in the PF.

4.1 Monte Carlo Design

Table 1 summarizes the parameter values used in the Monte Carlo experiment. Empty entries represent parameters that have to be set to zero for the purpose of model identification. Grey-shaded entries represent parameters that are set to zero to obtain closed-form solutions for the Ricatti ODEs. Our application involves pricing the cap contracts a large number of times for each model, and pricing each caplet requires solving the Ricatti ODEs at every integration point in (15). Numerically solving the Ricatti ODEs in this setup is therefore prohibitively expensive. We reduce the computational burden by following Ait-Sahalia and Kimmel (2010) and restrict some of the model parameters in order to obtain closed-form solutions for the Ricatti ODEs. These restrictions are such that the models are specified with $N - M$ correlated Gaussian processes that are uncorrelated with M square-root processes, which themselves are uncorrelated.⁹

With some exceptions, the parameters in Table 1 are from Table 8 in Ait-Sahalia and Kimmel (2010). The exceptions, all motivated by numerical considerations, are as follows. The δ_0 parameter is set to 3% for all models. Ait-Sahalia and Kimmel (2010) find zero values for the κ_{11} parameter in the $A_1(3)$ and $A_2(3)$ models. To enhance numerical stability in the matrix inversions we have shifted these parameter values by one standard error. Ait-Sahalia and Kimmel (2010) also have zero values for the mean θ_j of some of the volatility factors. We set these parameters equal to 5%, which enhances the stability of the simulations. Note that these changes result in parameters estimates that are well inside the estimated confidence intervals in Ait-Sahalia and Kimmel (2010) when using conventional significance levels.

Our choice of parameters does not materially affect our findings below. In each case, we have chosen conservative parameterizations that reduce the nonlinearities of bond and cap prices in the state variables compared with previous parameterizations. As a result, for most realistic parame-

⁹Restrictions on the parameters in ATSMs are also imposed in Balduzzi, Das, Foresi and Sundaram (1996), Chen (1996), and Dai and Singleton (2000).

terizations, the benefits of the proposed UKF method will generally be more substantial than what we document.

For each canonical model, we simulate 500 samples of LIBOR rates with maturities of 3 and 6 months, swap rates with maturities of 1, 2, 5, 7, and 10 years, and at-the-money interest rate caps on 3-month LIBOR with maturities of 1 and 5 years. Each sample contains 260 weekly observations.¹⁰ We use an Euler discretization of the stochastic differential equation of the state vector and divide the week into 100 time steps. Weekly observations are then extracted by taking the 100th observation within each week.

For each simulated sample, the filtered values of the unobserved state variables and the rate or price implied by each filter are compared to the simulated data under various scenarios. The unscented Kalman filter is implemented with rescaling parameters $\rho = 1$, $\chi = 0$, and $\gamma = 2$.¹¹

Figure 1 shows the unconditional term structure of interest rates implied by the parameters in Table 1 for each of the four models. Note that the parameters generate quite different term structures across models, and note in particular that the implied term structure for the $A_3(3)$ model is fairly flat.

¹⁰To investigate if sample length affects our conclusions, we subsampled, for the $A_1(3)$ model, 10 paths (k) with median ratios $RMSE_k^{EKF}/RMSE_k^{UKF}$ (using caps) and extended them to 520 and 1300 weeks instead of 260. Our conclusions were not affected by this change.

¹¹We set $\gamma = 2$ because it implies a Gaussian state. This assumption therefore induces a bias in our implementation of the UKF which is identical to the EKF bias. This ensures that our comparison is focused on the implications of nonlinearities in the measurement equation. Setting $\chi = 0$ is the standard for state estimation (van der Merwe and Wan (2002), Julier and Uhlmann (2004)). The parameter ρ , $0 \leq \rho \leq 1$, controls the size of the sigma-point distribution; setting it to 1 reflects a diffuse prior on the distribution of the (volatility) states. See Dunik, Simandel and Straka (2012) for further discussion of the scaling parameters.

4.2 State Vector Extraction

We begin by assessing the ability of each filter to accurately extract the unobserved path of the state variables. For each simulated time series, we compare the filtered state variables implied by each method to the actual state observations. As a gauge for the goodness of fit, we calculate the root mean squared error (RMSE) over 260 weeks for each Monte Carlo sample as follows

$$RMSE_k^F(i) = \sqrt{\frac{1}{260} \sum_{t=1}^{260} (x_{i,k,t} - x_{i,k,t}^F)^2}$$

where $x_{i,k,t}$ denotes the true but unobserved state variable i in sample k at time t . The filtered state variable is denoted by $x_{i,k,t}^F$ where F stands for the EKF, UKF or PF. For each model we compute the RMSE for each state variable and on each of 500 Monte Carlo samples.

Tables 2 and 3 provide the mean, median, and standard deviation of the RMSE for each state variable across the 500 simulated paths. For example, for the mean RMSE we compute

$$Mean(RMSE_k^F(i)) = \frac{1}{500} \sum_{k=1}^{500} RMSE_k^F(i).$$

For the EKF and UKF we also report for each statistic its ratio with that from the PF.

Panel A presents results for the $A_0(3)$ model, Panel B for the $A_1(3)$ model, Panel C for the $A_2(3)$ model, and Panel D for the $A_3(3)$ model. Table 2 provides results for the case where state filtering is done without caps, while Table 3 additionally uses cap prices to extract the state variables. Our prior is that because the cap prices are more nonlinear in the states, the relative performance of the UKF versus the EKF should be better in Table 3 compared with Table 2.

Table 2, which does not use caps to filter states, yields two important conclusions. First, the performance of the EKF and the UKF is very similar. Second, the performance of both filters is very similar to the performance of the PF, which requires more than 400 times the computations in the EKF, and more than 200 times the computations in the UKF.¹² Evidently, when caps are not

¹²These numbers are based on the 3000-particle PF used when filtering based on rates only. The computational burden of pricing caplets can easily be 1000 times that of computing a LIBOR rate, since the computation involves

used to filter states, the nonlinearities are not very pronounced and the performance of the three filters is similar.

Table 3 provides results for the case where caps are included in filtering. The UKF now significantly outperforms the EKF in every dimension. The mean and median RMSE are smaller for all models, often by a large amount. The mean and median RMSE for the UKF also compare fairly well with that of the PF. The standard deviation of the RMSE is also uniformly larger for the EKF compared to the UKF, in some cases dramatically so. This is a clear indication that the UKF is numerically more stable. The UKF is also very stable when using our PF implementation as a benchmark.

Comparing Tables 2 and 3, it is clear that for the EKF, the mean and median state RMSEs are dramatically larger when caps are included in filtering. As discussed earlier, the EKF performs a first order Taylor expansion around the predicted state variables. The quality of the EKF filtering thus crucially depends on the numerical gradient used in this first order approximation. The increased RMSEs demonstrate that when the highly nonlinear caps are used, the gradient offers a very poor approximation of the impact that variations in states have on the measurement equation. The UKF's median RMSEs in Table 3 are typically, but not always, smaller than those in Table 2. When judged by the mean RMSEs, the picture is mixed. These results on the UKF's mean and median RMSEs suggest that state extraction does not work well for some outliers when caps are used to extract states.

Figures 2 and 3 provide further insight in these results by showing scatter plots of the state RMSEs of the EKF and UKF against the RMSEs of the particle filter. For each model, each row of panels shows results for a different state variable. The two left-most columns of panels in each

numerical integrals. For the 3000-particle PF, this leads to a very high computational burden. Fortunately, the observed cap prices are very informative on the states (see Li, 2011), and we can use fewer particles, namely 300. The resulting PF still requires more than 40 times the computations in the EKF and more than 20 times the computations in the UKF.

figure show the case where caps are excluded from the filtering algorithm, while the two right-side columns are obtained by also including caps in the cross-section of observed securities. The right-side panels of Figures 2 and 3 clearly illustrate the superiority of the UKF in dealing with securities that are highly nonlinear in the state variables. Importantly the EKF is also numerically much less stable, as evidenced by the high number of outliers in the scatter plots.

Overall the results on the UKF in Tables 2-3 and Figures 2-3 are quite striking, and indicate that the UKF is usually able to incorporate the additional information contained in caps to extract the underlying states more precisely. The EKF, however, actually suffers from the additional information in caps because of the linearization. This initial Monte Carlo exercise therefore leaves little doubt that the UKF is much superior in filtering the state variables than the EKF when highly nonlinear securities are used. The UKF's outperformance is of course model-dependent, because different models and different model parameterizations imply different degrees of nonlinearity. The UKF also performs well when compared with the PF. When securities are only mildly nonlinear in the state variables as in Table 2, the performance of the EKF is adequate.

Our first main conclusion is that the UKF offers a clear improvement over the EKF in extracting states, in particular when caps are used to filter the states, and that the UKF performs well compared to the (optimal) PF, at a significantly lower computational cost.

4.3 Implications for Rates and Prices

In order to assess the economic implications of the two filtering methods, we now investigate the filters' ability to match observed LIBOR and swap rates as well as cap prices. We compare the fitted LIBOR, swap rates, and cap prices implied by states from each filter to the actual rates and prices computed from the true states. We also provide the fitted LIBOR, swap rates, and cap prices provided by the PF as a benchmark.

Tables 4 and 5 compare the security prices implied by the filtered states to the simulated true

values. We provide RMSEs as well as Bias. For each security the RMSE for filter F is computed as

$$RMSE^F = \frac{1}{500} \sum_{k=1}^{500} RMSE_k^F = \frac{1}{500} \sum_{k=1}^{500} \sqrt{\frac{1}{260} \sum_{t=1}^{260} (y_{k,t} - y_{k,t}^F)^2}$$

where $y_{k,t}$ is the true price or rate in sample k at week t , $y_{k,t}^F$ is the value obtained using a filtered state vector and F stands for the UKF, EKF, or PF. Bias is defined by

$$Bias^F = \frac{1}{130,000} \sum_{k=1}^{500} \sum_{t=1}^{260} (y_{k,t} - y_{k,t}^F)$$

All estimates are in basis points. Table 4 provides results when cap prices are not used in the filtering step, while Table 5 provides results when cap prices are included to filter the states.

Table 4 indicates that the UKF usually provides a lower RMSE than the EKF, even if the UKF bias is sometimes larger. The lower RMSE is robust across securities but not across models: the UKF underperforms the EKF for the $A_2(3)$ model. The differences in RMSE between the UKF and the EKF are small, which is to be expected given the very similar performance in extracting the states documented in Table 2. Table 4 also confirms the finding from Table 2 that the relative performance of the UKF compared to the EKF is indeed the worst in the case of the $A_2(3)$ model. Finally, Table 4 indicates that the performance of both the EKF and the UKF is satisfactory when benchmarked against the PF.

Table 5 indicates that the UKF dramatically outperforms the EKF when caps are also included in state filtering. Compared to Table 4, the RMSE implied by the UKF is substantially smaller for cap prices, and somewhat larger overall for LIBOR and swap rates. This result is not surprising because the states filtered on all securities represent a compromise between fitting rates and cap prices. Interestingly, the performance of the EKF relative to Table 4 deteriorates for the majority of securities and models. In terms of bias, the UKF generally outperforms the EKF in Table 5.

The degree of outperformance of the UKF over the EKF in Table 5 is model-dependent and security-dependent. The RMSE ratio is the largest for LIBOR. The EKF's RMSE for 3-month LIBOR is roughly eight times higher in the case of the $A_1(3)$ model, and more than three times

higher in the case of the $A_0(3)$ model. For the 6-month LIBOR, the corresponding ratios are approximately five and three. The improvements offered by the UKF are less dramatic in the case of the $A_3(3)$ model where the nonlinearities are less pronounced.

While the UKF substantially outperforms the EKF when highly nonlinear caps are used for extracting states, Table 5 indicates that it does not perform as well as the PF for pricing the caps. The differences are small for the 1-yr cap, but for the 5-yr cap the RMSE for the UKF is on average twice as large. The UKF's performance in fitting the more linear LIBOR and swap rates is better than our implementation of the PF, again suggesting a trade-off between pricing the more and less linear securities.

Our second main conclusion is that when caps are used to extract states, the UKF's improvement over the EKF in extracting states carries over to improvements in securities pricing. Furthermore, while adding nonlinear securities generally improves the performance of the nonlinear UKF filter in state vector extraction, the economic benefits are not evenly distributed across securities. However, the benefits are clear for the pricing of highly nonlinear securities. Overall the UKF also performs well compared to the PF.

4.4 Dynamic Implications: Rate and Price Forecasts

Dynamic term structure models are used not only for the valuation of securities at present but also to forecast future rates and prices (see for example Backus, Foresi, Mozumdar and Wu (2001) and Egorov, Hong and Li (2006)). The usefulness of the model for this purpose depends crucially on the accuracy of the state vector filter.

Table 6 summarizes the relative performance of the EKF and UKF for predicting LIBOR, swap rates, and cap prices for various forecasting horizons in each of our four models. For each security we compute the forecast RMSE for each horizon, h , defined by

$$RMSE_k^F(h) = \sqrt{\frac{1}{260-h} \sum_{t=1}^{260-h} \left(y_{k,t+h} - y_{k,t+h|t}^F \right)^2}$$

where $y_{k,t+h|t}^F$ is the price or rate of the security computed using the filter-dependent h -week ahead state vector forecast, $x_{k,t+h|t}^F$. In Table 6, Panel A reports the EKF to UKF ratios for $\text{Mean}(RMSE_k^F(h))$, Panel B reports ratios for $\text{Median}(RMSE_k^F(h))$, and Panel C reports ratios for $\text{Stdev}(RMSE_k^F(h))$. The moments are computed for $h = 1, 4, \text{ and } 12$ week horizons across the 500 samples, which are denoted by k . Forecasts are based on states filtered using LIBOR, swap rates, and cap prices.¹³

Table 6 confirms the conclusions from the contemporaneous fit in Tables 4-5: when considering forecasts based on states filtered using LIBOR, swap rates, and caps, the UKF clearly outperforms the EKF. This confirms the EKF's problems in handling securities that are highly nonlinear in the states. The magnitude of the improvement is smallest in the case of the $A_2(3)$ and $A_3(3)$ models. As previously discussed, this is not due to the nature of these models, but rather to the parameterization in Table 1, which determines the extent of nonlinearity in the states for each model.

Figures 4 and 5 provides more perspective by scatter plotting the 500 individual $RMSE_k^{UKF}(h)$ on the y-axis against the corresponding $RMSE_k^{EKF}(h)$ on the x-axis for the one-week forecast horizon ($h = 1$) when caps are included in the filtering of states. The UKF outperforms the EKF when the plots fall below the 45-degree line. Figures 4 and 5 are quite striking. Note that there are few observations above the 45-degree line. These figures provide a more visual and intuitive assessment of the performance of the two filtering methods. The figures confirm that the UKF-implied forecasts substantially outperform the EKF-implied forecasts.

Our third main conclusion is that the UKF generally delivers forecast RMSEs that are lower

¹³To save space, Table 6 only reports on results when caps are included in filtering, and we report ratios of RMSE statistics only. The levels are available upon request. They are consistent with those observed in Tables 4 and 5 and tend to be increasing with the forecast horizon, as expected. For forecasts based on states filtered using LIBOR and swap rates only, the EKF to UKF ratios are either one or a bit greater than one, consistent with the evidence in Table 4.

than those obtained using the EKF when caps are included to filter states.

4.5 Implications for Long-Maturity Caps

So far we have run two versions of each Monte Carlo experiment for each filter: One where only LIBOR and swap rates are used in filtering, and another one where in addition 1-yr and 5-yr caps are used. All results in Tables 4-6 are based on pricing the same securities that are used in filtering.

We now instead consider an application of the term structure models in which 7-yr caps must be priced. These contracts have not been used in any of the filters when extracting states. We restrict attention to contemporaneous pricing just as in Tables 4-5. We report results for the EKF and the UKF, as well as for the PF.

Table 7 contains the results. We again compute pricing RMSE and Bias computed from the true rates and prices as well as the extracted rates and prices obtained from each filter. In Panel A we report the results for the EKF and UKF when states are filtered on only LIBOR and swap rates. Not surprisingly in light of the previous results, the performance of the EKF and the UKF is similar, and not very different from the PF.

In Panel B of Table 7 we report on 7-yr cap pricing when states are filtered using 1-yr and 5-yr caps as well as the LIBOR and swap rates. Once again, the UKF outperforms the EKF in this case. The improvement in performance is again model-specific, and largest for the $A_0(3)$ and $A_1(3)$ models. Contrary to the results in Panel A, the PF outperforms the UKF for pricing the 7-yr caps in Panel B.

Our fourth main conclusion is that when caps are used in filtering, the UKF outperforms the EKF for the pricing of nonlinear securities, even when these securities have not been used in filtering the underlying state vectors.

5 Parameter Estimation

Our Monte Carlo experiments show that the extended Kalman filter is ill-suited to optimally exploit the rich information in securities that are nonlinear in the state variables. We propose the unscented Kalman filter as an alternative to address the nonlinearity in the measurement equation. In our Monte Carlo analysis, we have deliberately kept the structural parameters fixed at their true values. However, our results suggest that the choice of filter will also affect parameter estimation. The inclusion of caps in the sample makes it difficult to investigate this issue in a full-fledged Monte Carlo analysis. We therefore reduce the computational burden of the Monte-Carlo analysis using asymptotic arguments.¹⁴

The literature contains a large number of empirical methods that can be used to estimate multifactor affine models, including indirect inference, simulated method of moments (SML), and the efficient method of moments (EMM). Most papers use either quasi maximum likelihood (QML) or the Kalman filter with a likelihood based criterion.¹⁵ These techniques are popular because they are relatively easy to implement and because Duffee and Stanton (2004) demonstrate in an extensive Monte Carlo experiment that QML and Kalman filtering outperform more complex estimation techniques (EMM and SML) in finite samples.

In our Monte Carlo analysis we consider the quasi log-likelihood function assuming that $y_{k,t}$ is

¹⁴We are very grateful to an anonymous referee for suggesting this approach.

¹⁵See for example Babbs and Nowman (1999), Chen and Scott (1995), Dai and Singleton (2000), Duan and Simonato (1999), Duffee (1999), Duffie and Singleton (1997), and Pearson and Sun (1994). See Thompson (2008) for an alternative approach using Bayesian filtering.

normally distributed with mean $y_{k,t|t-1}$ and covariance matrix $P_{k,yy(t|t-1)}$, that is:¹⁶

$$\begin{aligned} \ell_k(\Theta) = & -\frac{1}{2} \sum_{t=1}^T N \log 2\pi + \log |P_{yy(t|t-1)}| \\ & + (y_{k,t} - y_{k,t|t-1})' P_{yy(t|t-1)}^{-1} (y_{k,t} - y_{k,t|t-1}) . \end{aligned} \quad (36)$$

The analysis of the quasi log-likelihood at the optimal parameter values sheds light on the potential impact of the filter on estimation via maximum likelihood, assuming that all rates are observed with error. Table 8 summarizes the results. To keep the analysis manageable, we focus on just one model, the A_1 (3) model.

On each of the 500 simulated trajectories, we approximate the score vector $s(\Theta)$ at the optimum parameter value Θ_0 , computing the numerical gradient of the quasi log likelihood ℓ with respect to the optimal parameters. That is, for each parameter Θ_i , we keep the optimal value for all other parameters, use the filters to extract the states at $\Theta_{0,i} \pm \epsilon$, compute the log-likelihood as in (36), and approximate the scores by

$$s_{k,KF}(\Theta_{0,i}) = (\nabla \ell_k)_i \simeq \frac{\ell_k(\Theta_{0,i} + \epsilon) - \ell_k(\Theta_{0,i} - \epsilon)}{2\epsilon}, \quad KF \in \{\text{EKF}, \text{UKF}\}.$$

We use $\epsilon = 10^{-6}$ in implementation. Using the optimal filter, the likelihood should be maximized at the optimal parameters; hence, consistency requires pseudo-scores with an expected value of zero. The results (not reported) indicate that for all parameters, the distribution of the scores is centered around zero, but the median score for the UKF is always closer to zero than for the EKF. In fact, the entire distribution of the scores obtained using the UKF is more concentrated around zero, while the scores for the EKF are much more dispersed.

We now use these simulation results to comment on parameter estimation errors, using an asymptotic argument. The parameter estimation error can be expressed as

$$\hat{\Theta}_{k,KF} - \Theta_0 \simeq H^{-1}(\Theta_0) s_{k,KF}(\Theta_0), \quad (37)$$

¹⁶The predictive covariance matrix for the EKF is given by equation (27); for the UKF, the measurement update yields $S_{yy(t|t-1)}$ and $P_{yy(t|t-1)} = S'_{yy(t|t-1)} S_{yy(t|t-1)}$.

where $H(\Theta_0)$ is the Hessian matrix at Θ_0 . Table 8 reports the distribution of the estimation errors based on (37). Note that one could take into account the dependence of the Hessian on the filter. However, this would require second-differencing in Θ , and therefore repeatedly re-filtering the states, which is extremely costly computationally. Instead, we perform the second-differencing assuming perfect filtering (i.e. using the simulated states x_t) and use the same (optimal) Hessian matrix for both Kalman filters.

Table 8 indicates that the median estimation error is virtually zero for all parameters for both filters. However, the distribution of the estimation errors for the EKF is much more dispersed than the distribution of the UKF-based errors. These results strongly suggest that the UKF will be superior to the EKF for the purpose of parameter estimation. Given that our analysis in Section 4 indicates that the UKF performs well compared with the PF for most models, both in terms of state extraction and of fit to the observed yields, we conjecture that the UKF will yield estimation results reasonably close to the PF, obtained at much lower computational cost.

6 Conclusion

The extended Kalman filter has become the standard tool to analyze a number of important problems in financial economics, and in term structure modeling in particular. While there is no need to look beyond the extended Kalman filter for some term structure applications, it is not clear how well the method performs for many situations of interest, when the measurement equation is nonlinear in the state variables. Examples include the pricing of fixed income derivatives such as caps, floors and swaptions, as well as modeling the cross section of swap yields. The unscented Kalman filter is moderately more costly from a computational perspective, but better suited to handling these nonlinear securities.

We use an extensive Monte Carlo experiment to investigate the performance of the extended and unscented Kalman filter and compare them to the PF benchmark. We study three-factor affine

term structure models for LIBOR and swap rates, which are mildly nonlinear in the underlying state variables, and cap prices, which are highly nonlinear. We find that when caps are used, the filtering performance of the unscented Kalman filter is much superior to that of the extended Kalman filter. It filters the states more accurately, which leads to improved fitting of security prices and more accurate forecasts. We conduct a small-scale experiment that suggests that these advantages result in improved properties for parameter estimates obtained using the UKF. We also compare the performance of the UKF to that of the particle filter, and find that the UKF's performance is often very close to that of the PF, at a fraction of the computational cost.

These results therefore suggest that the UKF may prove to be a good approach for implementing term structure models in a wide variety of applications, including the estimation of term structure models using interest rate derivatives, the estimation of nonlinear term structure models such as quadratic models, and the estimation of the term structure of default risk (Benzoni, Collin-Dufresne, Goldstein, and Helwege (2011)). The unscented Kalman filter may also prove useful for estimating other types of term structure models, such as the unspanned stochastic volatility models of Collin-Dufresne and Goldstein (2002), and Collin-Dufresne, Goldstein, and Jones (2009), and more broadly for other financial applications that involve nonlinear state-space systems.

References

- [1] Ait-Sahalia, Y., and R. Kimmel, 2010, "Estimating Affine Multifactor Term Structure Models Using Closed-Form Likelihood Expansions," *Journal of Financial Economics*, 98, 113–144.
- [2] Almeida, C., J. Graveline and S. Joslin, 2011, "Do Options Contain Information About Excess Bond Returns?" *Journal of Econometrics*, 164, 35–44.
- [3] Andersen, T. G., and L. Benzoni, 2010, "Do Bonds Span Volatility Risk in the US Treasury Market? A Specification Test for Affine Term Structure Models," *Journal of Finance*, 65,

603–653.

- [4] Babbs, S., and B. Nowman, 1999, “Kalman Filtering of Generalized Vasicek Term Structure Models,” *Journal of Financial and Quantitative Analysis*, 34, 115–130.
- [5] Backus, D., S. Foresi, A. Mozumdar and L. Wu, 2001, “Predictable Changes in Yields and Forward Rates,” *Journal of Financial Economics*, 59, 281–311.
- [6] Bakshi, G., P. Carr, and L. Wu, 2008, “Stochastic Risk Premiums, Stochastic Skewness in Currency Options, and Stochastic Discount Factors in International Economies,” *Journal of Financial Economics*, 87, 132–156.
- [7] Balduzzi, P., S. Das, S. Foresi, and R. Sundaram, 1996, “A Simple Approach to Three Factor Affine Term Structure Models,” *Journal of Fixed Income*, 6, 43–53.
- [8] Benzoni, L., P. Collin-Dufresne, R. Goldstein, and J. Helwege, 2011, “Credit Contagion via the Updating of Fragile Beliefs,” Working Paper, University of Minnesota.
- [9] Bikbov, R., and M. Chernov, 2009, “Unspanned Stochastic Volatility in Affine Models: Evidence from Eurodollar Futures and Options,” *Management Science*, 55, 1292–1305.
- [10] Carr, P., and L. Wu, 2007, “Stochastic Skew in Currency Options,” *Journal of Financial Economics*, 86, 213–247
- [11] Chen, L., 1996, “Stochastic Mean and Stochastic Volatility—A Three-Factor Model of the Term Structure of Interest Rates and Its Application to the Pricing of Interest Rate Derivatives,” *Financial Markets, Institutions, and Instruments*, 5, 1–88.
- [12] Chen, R. R., and L. Scott, 1995, “Multi-Factor Cox-Ingersoll-Ross Models of the Term Structure: Estimates and Tests from a Kalman Filter Model,” Working Paper, University of Georgia.

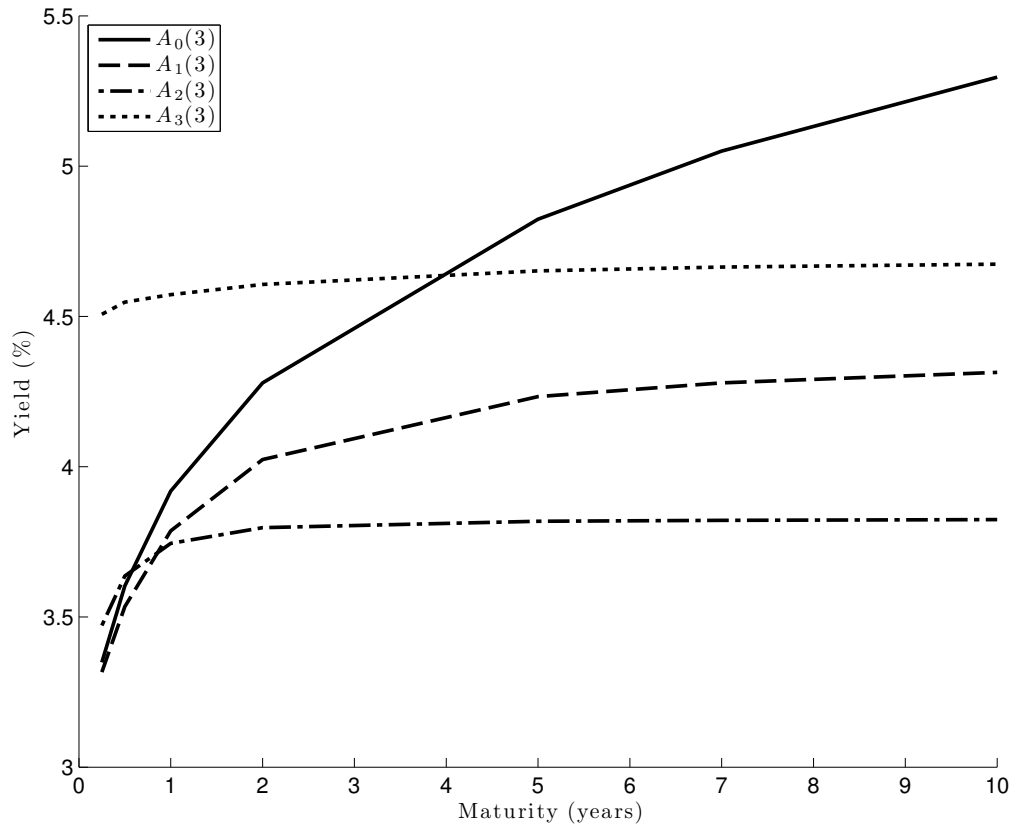
- [13] Chen, R. R., X. Cheng, F. Fabozzi, and B. Liu, 2008, “An Explicit, Multi-Factor Credit Default Swap Pricing Model with Correlated Factors,” *Journal of Financial and Quantitative Analysis*, 43, 123–160.
- [14] Cheridito, P., D. Filipović, and R. Kimmel, 2007, “Market Price of Risk Specifications for Affine Models: Theory and Evidence,” *Journal of Financial Economics*, 83, 123–170.
- [15] Collin-Dufresne, P., and B. Solnik, 2001, “On the Term Structure of Default Risk Premia in the Swap and Libor Markets,” *Journal of Finance*, 56, 1095–1115.
- [16] Collin-Dufresne, P., and R. Goldstein, 2002, “Do Bonds Span the Fixed Income Markets? Theory and Evidence for Unspanned Stochastic Volatility,” *Journal of Finance*, 57, 1685–1730.
- [17] Collin-Dufresne, P., R. Goldstein and C. Jones, 2009, “Can Interest Rate Volatility be Extracted from the Cross Section of Bond Yields? An Investigation of Unspanned Stochastic Volatility,” *Journal of Financial Economics*, 94, 47–66.
- [18] Dai, Q., and K. Singleton, 2000, “Specification Analysis of Affine Term Structure Models,” *Journal of Finance*, 55, 1943–1978.
- [19] Dai, Q., and K. Singleton, 2002, “Expectation Puzzles, Time Varying Risk Premia and Affine Models of the Term Structure,” *Journal of Financial Economics*, 63, 415–441.
- [20] de Jong, F., 2000, “Time Series and Cross-Section Information in Affine Term Structure Models,” *Journal of Business and Economic Statistics*, 18, 300–314.
- [21] Duan, J. C., and J. G. Simonato, 1999, “Estimating and Testing Exponential-Affine Term Structure Models by Kalman Filter,” *Review of Quantitative Finance and Accounting*, 13, 111–135.
- [22] Duarte, J., 2004, “Evaluating an Alternative Risk Preference in Affine Term Structure Models,” *Review of Financial Studies*, 17, 379–404.

- [23] Duffee, G., 1999, Estimating the Price of Default Risk, *Review of Financial Studies*, 12, 197–226.
- [24] Duffee, G., 2002, “Term Premia and Interest Rate Forecasts in Affine Models,” *Journal of Finance*, 57, 405–443.
- [25] Duffee, G., and R. Stanton, 2004, “Estimation of Dynamic Term Structure Models,” Working Paper, Haas School of Business, University of California at Berkeley.
- [26] Duffie, D., and R. Kan, 1996, “A Yield-Factor Model of Interest Rates,” *Mathematical Finance*, 6, 379–406.
- [27] Duffie, D., J. Pan and K. Singleton, 2000, “Transform Analysis and Asset Pricing for Affine Jump-Diffusions,” *Econometrica*, 68, 1343–1376.
- [28] Duffie, D., and K. Singleton, 1997, “An Econometric Model of the Term Structure of Interest Rate Swap Yields,” *Journal of Finance*, 52, 1287–1323.
- [29] Dunik, J., M. Simandel, and O. Straka, 2012, “Unscented Kalman Filter: Aspects and Adaptive Setting of Scaling Parameter,” *IEEE Transactions on Automatic Control* 57, 2411–2416.
- [30] Egorov, A., Y. Hong and H. Li, 2006, “Validating Forecasts of the Joint Probability Density of Bond Yields: Can Affine Models Beat Random Walk?” *Journal of Econometrics*, 135, 255–284.
- [31] Feldhutter, P., and D. Lando, 2008, “Decomposing Swap Spreads,” *Journal of Financial Economics*, 88, 375–405.
- [32] Fontaine, J.-S., and R. Garcia, 2012, “Bond Liquidity Premia,” *Review of Financial Studies*, 25, 1207–1254.
- [33] Jarrow, R., H. Li and F. Zhao, 2007, “Interest Rate Caps ‘Smile’ Too! But Can the LIBOR Market Models Capture Smile?” *Journal of Finance*, 62, 345–382.

- [34] Julier, S. J., 2000, “The Spherical Simplex Unscented Transformation,” Proceedings of the IEEE American Control Conference.
- [35] Julier, S. J., and J. K. Uhlmann, 2004, “Unscented Filtering and Nonlinear Estimation,” *IEEE Review*, 92, March.
- [36] Li, H. and F. Zhao, 2006, “Unspanned Stochastic Volatility: Evidence from Hedging Interest Rate Derivatives,” *Journal of Finance*, 61, 341–378.
- [37] Li, J., 2011, “Sequential Bayesian Analysis of Time-Changed Infinite Activity Derivatives Pricing Models,” *Journal of Business & Economic Statistics*, 29, 468–480.
- [38] Li, J., 2013, “An Unscented Kalman Smoother for Volatility Extraction: Evidence from Stock Prices and Options ” *Computational Statistics & Data Analysis*, 58, 15–26.
- [39] Liu, J., and R. Chen (1998), “Sequential Monte Carlo Methods for Dynamic Systems,” *Journal of the American Statistical Association*, 93, 1032–1044.
- [40] Lund, J., 1997, “Non-Linear Kalman Filtering Techniques for Term Structure Models,” Working Paper, Aarhus School of Business.
- [41] Pearson, N. D., and T.-S. Sun, 1994, “Exploiting the Conditional Density in Estimating the Term Structure: an Application to the Cox, Ingersoll, and Ross Model, *Journal of Finance*, 49, 1279–1304.
- [42] Thompson, S., 2008, “Identifying Term Structure Volatility from the LIBOR-Swap Curve,” *Review of Financial Studies*, 21, 819–854.
- [43] van Binsbergen, J., and R. Koijen, 2012, “Predictive Regressions: A Present-Value Approach,” *Journal of Finance*, 65, 1439–1471.

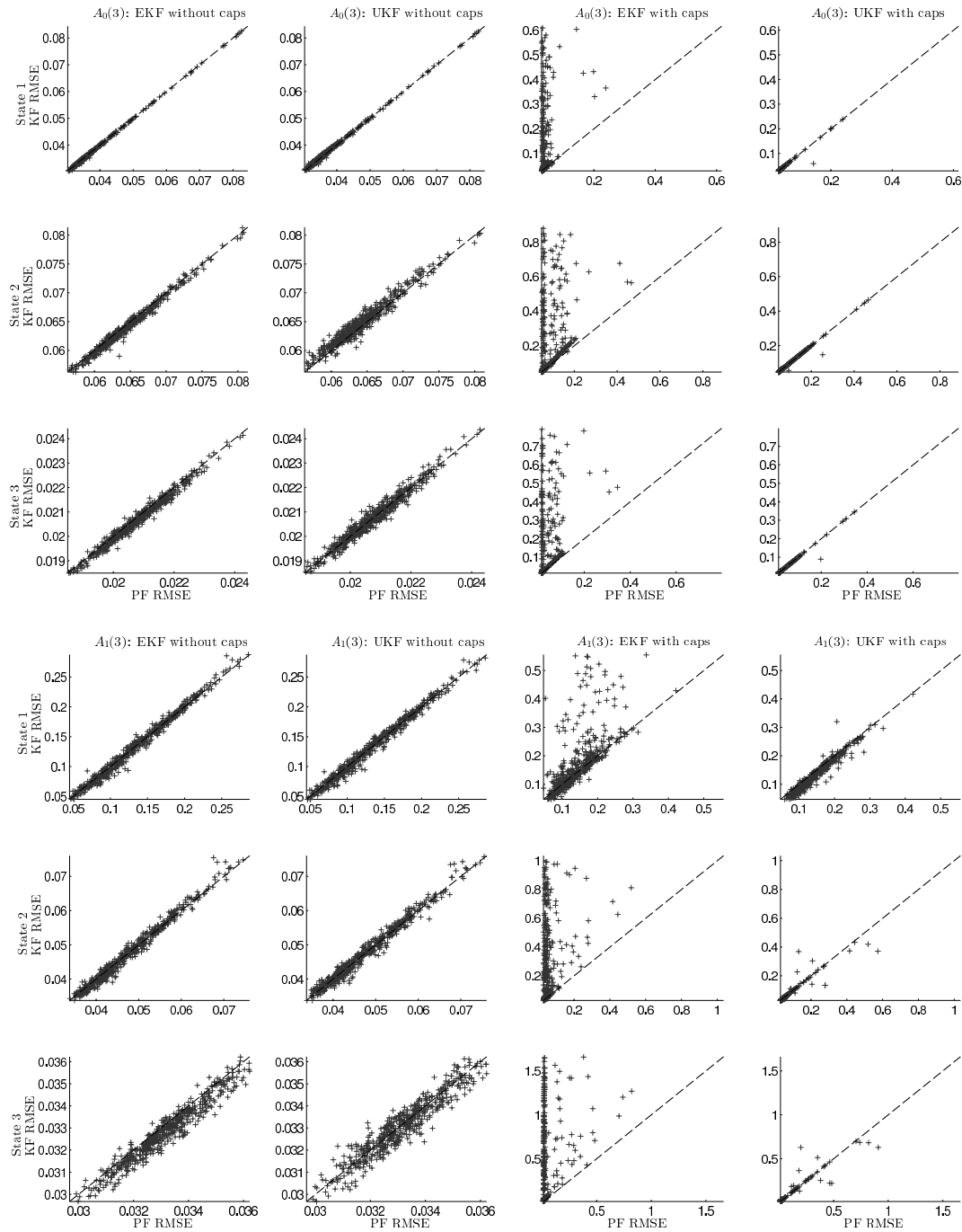
- [44] van der Merwe, R., and E. A. Wan, 2002, “The Square-Root Unscented Kalman Filter for State and Parameter-Estimation,” Proceedings of the 2001 IEEE International Conference On Acoustics, Speech, and Signal Processing, 3461–3464.

Figure 1: Unconditional Term Structures of Interest Rates. $A_M(3)$ Models



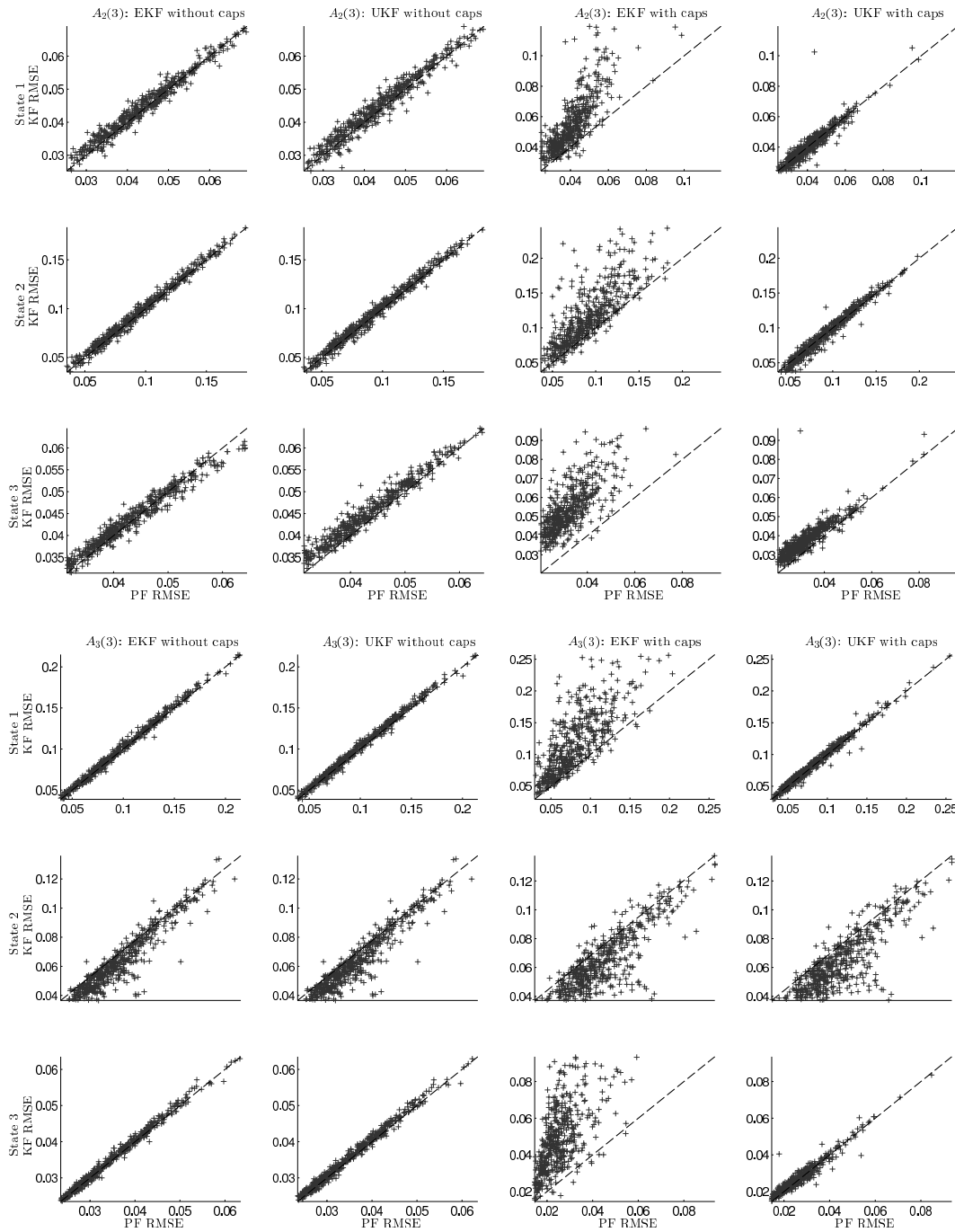
Notes: We display the unconditional term structure of interest rates implied by the four $A_M(3)$ models we consider, using the parameter values in Table 1.

Figure 2: Filtered State RMSEs. EKF & UKF versus Particle Filter. $A_0(3)$ & $A_1(3)$ Models



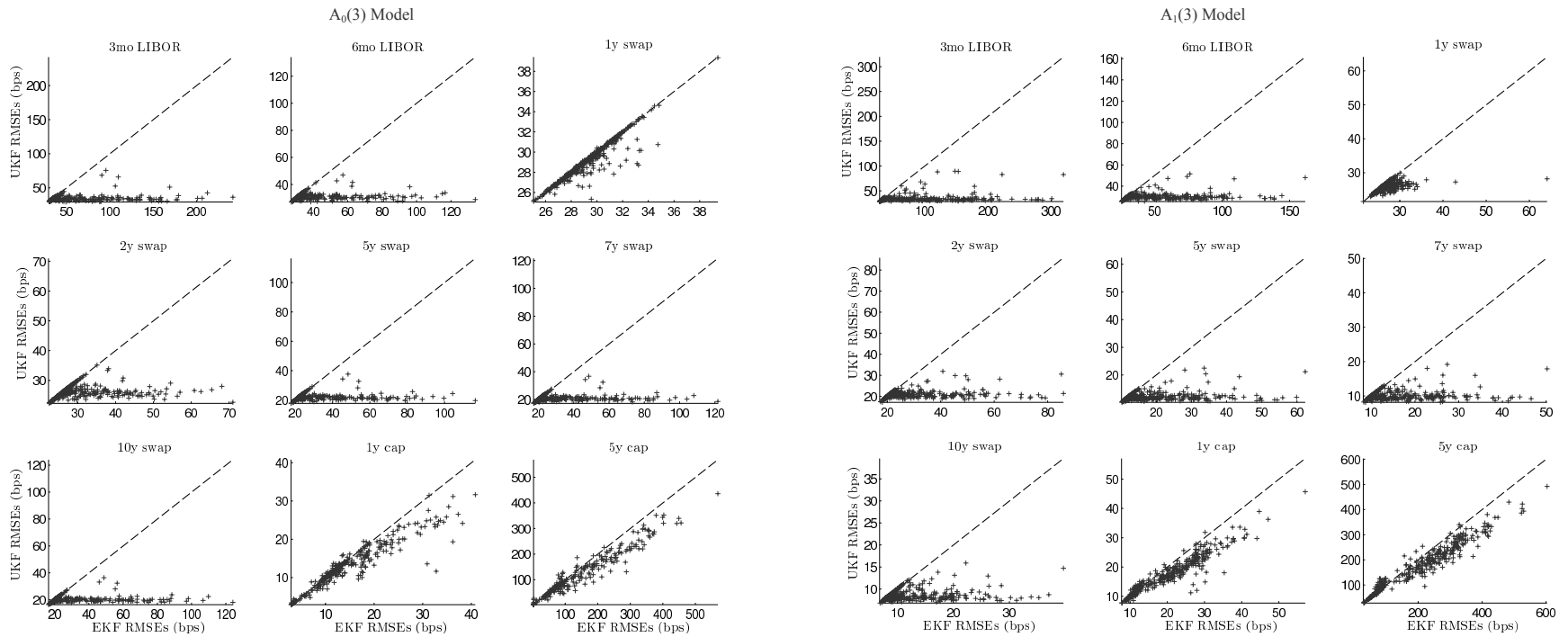
Notes: We scatter plot the filtered state RMSE of the Kalman filters against that of the particle filter for the $A_0(3)$ model (top three rows) and $A_1(3)$ model (bottom three rows). Errors are in basis points. Each row of panels depicts a different state variable. The two left-side columns show states filtered using LIBOR and swap rates only; the two right-side columns show filtered states obtained using the rates as well as the cap prices. The PF outperforms the Kalman filter when marks are above the dashed 45-degree line.

Figure 3: Filtered State RMSEs. EKF & UKF versus Particle Filter. $A_2(3)$ & $A_3(3)$ Models



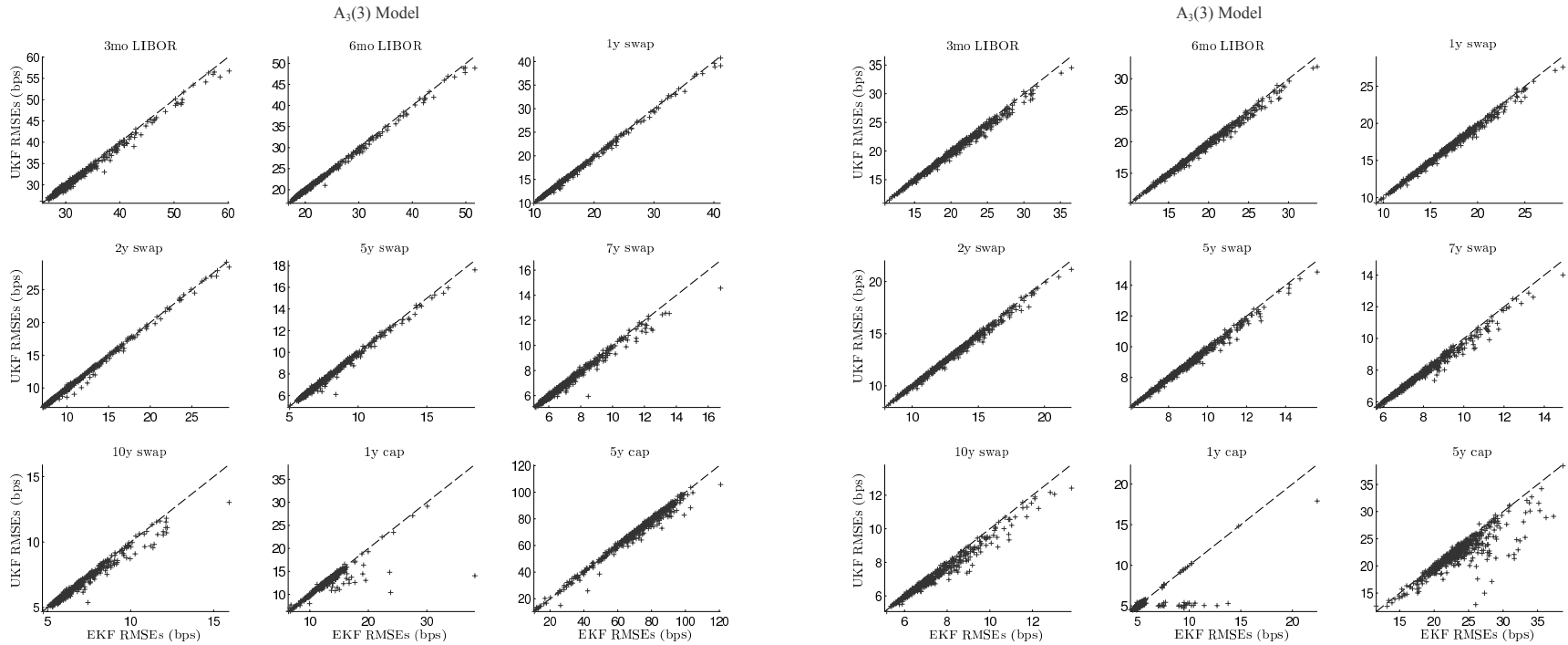
Notes: We scatter plot the filtered state RMSE of the Kalman filters against that of the particle filter for the $A_2(3)$ model (top three rows) and $A_3(3)$ model (bottom three rows). Errors are in basis points. Each row of panels depicts a different state variable. The two left-side columns show states filtered using LIBOR and swap rates only; the two right-side columns show filtered states obtained using the rates as well as the cap prices. The PF outperforms the Kalman filter when marks are above the dashed 45-degree line.

Figure 4: Rate and Price Forecast RMSEs. UKF versus EKF. States Filtered using LIBOR, Swap Rates and Caps. $A_0(3)$ & $A_1(3)$ Models



Notes: For each of the nine rates and prices, we scatter the 500 simulated one-week-ahead forecast RMSEs of the UKF model against the corresponding RMSEs for the EKF. The UKF outperforms the EKF when marks fall below the dashed 45-degree line. The state variables are filtered using LIBOR, swap rates and caps.

Figure 5: Rate and Price Forecast RMSEs. UKF versus EKF. States Filtered using LIBOR, Swap Rates and Caps. $A_2(3)$ & $A_3(3)$ Models



Notes: For each of the nine rates and prices, we scatter the 500 simulated one-week-ahead forecast RMSEs of the UKF model against the corresponding RMSEs for the EKF. The UKF outperforms the EKF when marks fall below the dashed 45-degree line. The state variables are filtered using LIBOR, swap rates and caps.

Table 1: Parameters for the $A_M(3)$ Models

Parameter	$A_0(3)$			$A_1(3)$			$A_2(3)$			$A_3(3)$		
	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3
δ_0	0.030			0.030			0.030			0.030		
δ_{ij}	0.0048	-0.0130	0.0241	0.0028	0.0052	0.0281	0.0194	0.0028	0.0391	0.0028	0.0002	0.0145
κ_{ij}	0.0168			0.0390			0.1600	0		0.0370	0	0
κ_{2j}	0.4000	2.9600		0	0.8800		0	0.0380		0	5.7100	0
κ_{3j}	-0.6400	-2.5600	0.8410	0	-2.3200	2.6900	0	0	5.6500	0	0	0.8100
θ_j				0.05			0.05	0.05		0.05	0.05	1.00
λ_{0j}	-0.190	0.610	-0.970	0.001	-0.120	-0.970	0.680	-0.035	-1.100	-0.034	-0.010	-0.100
α_j	1	1	1		1	1			1			
β_1				1	0	0	1		0	1		
β_2								1	0		1	
β_3												1

Notes: We report the parameter values used in the Monte Carlo simulations for the four $A_M(3)$ models. Empty entries indicate zero parameter values that are implicit to the normalized form of the models or imposed for identification. Grey-shaded zero entries indicate restrictions placed on the parameters in order to obtain closed-form solutions to the Riccati equations. With some exceptions, the parameters are from Table 8 in Ait-Sahalia and Kimmel (2010). The exceptions are motivated by numerical considerations in the simulations and filtering. See Section 4.1 for the details.

Table 2: State RMSEs from States Filtered using LIBOR and Swap Rates Only**Panel A: $A_0(3)$ Model**

	Factor 1			Factor 2			Factor 3		
	EKF	UKF	PF	EKF	UKF	PF	EKF	UKF	PF
Mean(RMSE)	0.0378	0.0382	0.0380	0.0640	0.0650	0.0645	0.0207	0.0208	0.0208
Ratio with PF	0.99	1.00	1.00	0.99	1.01	1.00	0.99	1.00	1.00
Median(RMSE)	0.0352	0.0356	0.0354	0.0631	0.0641	0.0635	0.0206	0.0207	0.0207
Ratio with PF	0.99	1.01	1.00	0.99	1.01	1.00	0.99	1.00	1.00
Stdev(RMSE)	0.0099	0.0098	0.0099	0.0050	0.0049	0.0050	0.0012	0.0012	0.0012
Ratio with PF	1.00	0.99	1.00	1.01	0.99	1.00	0.99	1.00	1.00

Panel B: $A_1(3)$ Model

	Factor 1			Factor 2			Factor 3		
	EKF	UKF	PF	EKF	UKF	PF	EKF	UKF	PF
Mean(RMSE)	0.1352	0.1348	0.1356	0.0476	0.0478	0.0478	0.0328	0.0331	0.0331
Ratio with PF	1.00	0.99	1.00	1.00	1.00	1.00	0.99	1.00	1.00
Median(RMSE)	0.1305	0.1280	0.1293	0.0457	0.0461	0.0459	0.0327	0.0331	0.0331
Ratio with PF	1.01	0.99	1.00	1.00	1.01	1.00	0.99	1.00	1.00
Stdev(RMSE)	0.0557	0.0550	0.0539	0.0097	0.0094	0.0092	0.0014	0.0014	0.0014
Ratio with PF	1.03	1.02	1.00	1.06	1.02	1.00	1.00	1.00	1.00

Panel C: $A_2(3)$ Model

	Factor 1			Factor 2			Factor 3		
	EKF	UKF	PF	EKF	UKF	PF	EKF	UKF	PF
Mean(RMSE)	0.0445	0.0448	0.0438	0.0954	0.0956	0.0940	0.0438	0.0454	0.0435
Ratio with PF	1.02	1.02	1.00	1.02	1.02	1.00	1.01	1.04	1.00
Median(RMSE)	0.0440	0.0440	0.0431	0.0915	0.0931	0.0903	0.0430	0.0443	0.0425
Ratio with PF	1.02	1.02	1.00	1.01	1.03	1.00	1.01	1.04	1.00
Stdev(RMSE)	0.0098	0.0102	0.0102	0.0334	0.0337	0.0333	0.0070	0.0074	0.0079
Ratio with PF	0.95	0.99	1.00	1.00	1.01	1.00	0.89	0.94	1.00

Panel D: $A_3(3)$ Model

	Factor 1			Factor 2			Factor 3		
	EKF	UKF	PF	EKF	UKF	PF	EKF	UKF	PF
Mean(RMSE)	0.1016	0.1023	0.1001	0.0652	0.0652	0.0710	0.0357	0.0360	0.0354
Ratio with PF	1.02	1.02	1.00	0.92	0.92	1.00	1.01	1.02	1.00
Median(RMSE)	0.0952	0.0959	0.0939	0.0598	0.0598	0.0668	0.0339	0.0344	0.0336
Ratio with PF	1.01	1.02	1.00	0.90	0.90	1.00	1.01	1.02	1.00
Stdev(RMSE)	0.0393	0.0395	0.0388	0.0227	0.0227	0.0207	0.0088	0.0088	0.0087
Ratio with PF	1.01	1.02	1.00	1.10	1.10	1.00	1.01	1.01	1.00

Notes: For each model, we report the mean, median, and standard deviation of the state RMSEs from the extended and the unscented Kalman filters, and for the particle filter (3000 particles) using 500 simulated paths. For each statistic, the ratio of the KF to PF RMSE is also reported. In each of the 500 simulations, 260 weekly LIBOR and swap rates are generated using the parameters from Table 1. States are filtered using LIBOR and swap rates only. The total number of rates evaluated is (in millions) 25.5 using EKF, 47.3 using UKF, and 10,967 using PF.

Table 3: State RMSEs from States Filtered using LIBOR, Swap Rates and Cap Prices**Panel A: $A_0(3)$ Model**

	Factor 1			Factor 2			Factor 3		
	EKF	UKF	PF	EKF	UKF	PF	EKF	UKF	PF
Mean(RMSE)	0.1227	0.0385	0.0396	0.2172	0.0854	0.0863	0.1541	0.0376	0.0379
Ratio with PF	3.10	0.97	1.00	2.52	0.99	1.00	4.07	0.99	1.00
Median(RMSE)	0.0396	0.0342	0.0352	0.1002	0.0596	0.0599	0.0392	0.0199	0.0199
Ratio with PF	1.12	0.97	1.00	1.67	1.00	1.00	1.97	1.00	1.00
Stdev(RMSE)	0.1707	0.0173	0.0178	0.2479	0.0504	0.0505	0.2273	0.0360	0.0367
Ratio with PF	9.59	0.97	1.00	4.91	1.00	1.00	6.19	0.98	1.00

Panel B: $A_1(3)$ Model

	Factor 1			Factor 2			Factor 3		
	EKF	UKF	PF	EKF	UKF	PF	EKF	UKF	PF
Mean(RMSE)	0.1859	0.1304	0.1437	0.2496	0.0546	0.0567	0.3875	0.0460	0.0426
Ratio with PF	1.29	0.91	1.00	4.40	0.96	1.00	9.10	1.08	1.00
Median(RMSE)	0.1448	0.1225	0.1370	0.0730	0.0431	0.0454	0.0481	0.0232	0.0184
Ratio with PF	1.06	0.89	1.00	1.61	0.95	1.00	2.61	1.26	1.00
Stdev(RMSE)	0.1450	0.0557	0.0527	0.3124	0.0467	0.0514	0.5276	0.0872	0.0940
Ratio with PF	2.75	1.06	1.00	6.08	0.91	1.00	5.61	0.93	1.00

Panel C: $A_2(3)$ Model

	Factor 1			Factor 2			Factor 3		
	EKF	UKF	PF	EKF	UKF	PF	EKF	UKF	PF
Mean(RMSE)	0.0570	0.0415	0.0415	0.1213	0.0917	0.0924	0.0556	0.0381	0.0328
Ratio with PF	1.37	1.00	1.00	1.31	0.99	1.00	1.69	1.16	1.00
Median(RMSE)	0.0511	0.0400	0.0402	0.1087	0.0882	0.0889	0.0514	0.0363	0.0312
Ratio with PF	1.27	0.99	1.00	1.22	0.99	1.00	1.65	1.16	1.00
Stdev(RMSE)	0.0221	0.0113	0.0107	0.0547	0.0329	0.0310	0.0173	0.0093	0.0097
Ratio with PF	2.07	1.05	1.00	1.77	1.06	1.00	1.77	0.95	1.00

Panel D: $A_3(3)$ Model

	Factor 1			Factor 2			Factor 3		
	EKF	UKF	PF	EKF	UKF	PF	EKF	UKF	PF
Mean(RMSE)	0.1204	0.0832	0.0818	0.0661	0.0655	0.0767	0.0496	0.0266	0.0259
Ratio with PF	1.47	1.02	1.00	0.86	0.85	1.00	1.91	1.03	1.00
Median(RMSE)	0.1076	0.0764	0.0754	0.0603	0.0599	0.0731	0.0467	0.0249	0.0241
Ratio with PF	1.43	1.01	1.00	0.83	0.82	1.00	1.93	1.03	1.00
Stdev(RMSE)	0.0575	0.0354	0.0349	0.0225	0.0227	0.0204	0.0200	0.0087	0.0090
Ratio with PF	1.65	1.01	1.00	1.10	1.11	1.00	2.21	0.96	1.00

Notes: For each model, we report the mean, median, and standard deviation of the state RMSEs from the extended and the unscented Kalman filters, and for the particle filter (300 particles) using 500 simulated paths. For each statistic, the ratio of the KF to PF is also reported. In each of the 500 simulations, 260 weekly LIBOR and swap rates are generated using the parameters from Table 1. States are filtered using LIBOR, swap rates, and caps. The total number of rates and caplets priced under each method is (in millions) 25.5 and 80.1 using EKF, 47.3 and 148.7 using UKF, and 1139 and 3581 using PF.

Table 4: Rate and Price Fit of $A_M(3)$ Models. States Filtered using LIBOR and Swap Rates Only.

	$A_0(3)$			$A_1(3)$			$A_2(3)$			$A_3(3)$		
	EKF	UKF	PF	EKF	UKF	PF	EKF	UKF	PF	EKF	UKF	PF
3-mo LIBOR												
Bias	0.0589	0.0298	0.0456	-0.0090	-0.3150	-0.0461	0.2669	-1.0762	0.0539	0.0979	-0.2204	0.0577
RMSE	2.8807	2.6780	2.9228	2.7410	2.6076	2.7985	2.6107	3.1009	2.8845	4.0919	4.0478	4.1311
RMSE/PF	0.99	0.92	1.00	0.98	0.93	1.00	0.91	1.08	1.00	0.99	0.98	1.00
6-mo LIBOR												
Bias	0.0713	0.0344	0.0582	0.0866	-0.2896	0.0302	0.1807	-0.9708	0.0092	0.0801	-0.2223	0.0389
RMSE	4.3794	4.3662	4.3876	4.1631	4.1719	4.1887	4.3363	4.6642	4.2718	4.2374	4.2179	4.2610
RMSE/PF	1.00	1.00	1.00	0.99	1.00	1.00	1.02	1.09	1.00	0.99	0.99	1.00
1-yr Swap												
Bias	0.1020	0.0596	0.0896	0.1616	-0.2532	0.0930	0.1144	-0.8136	-0.0173	0.0619	-0.2079	0.0234
RMSE	4.2251	4.1191	4.2376	4.0565	3.9915	4.1088	4.5292	4.6827	4.4915	4.4865	4.4852	4.4937
RMSE/PF	1.00	0.97	1.00	0.99	0.97	1.00	1.01	1.04	1.00	1.00	1.00	1.00
2-yr Swap												
Bias	0.0866	0.0450	0.0752	0.0880	-0.2868	0.0267	0.0163	-0.6854	-0.0833	-0.0193	-0.2420	-0.0512
RMSE	4.4751	4.3969	4.4864	4.0292	3.9868	4.0604	4.6604	4.7472	4.6378	4.6911	4.6912	4.6914
RMSE/PF	1.00	0.98	1.00	0.99	0.98	1.00	1.00	1.02	1.00	1.00	1.00	1.00
5-yr Swap												
Bias	0.1115	0.0764	0.1009	-0.0484	-0.2962	-0.0817	-0.0403	-0.4890	-0.1091	-0.0823	-0.2388	-0.1038
RMSE	4.8206	4.8017	4.8263	4.6495	4.6077	4.6639	4.8262	4.8356	4.8397	4.8364	4.7803	4.8465
RMSE/PF	1.00	0.99	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.00	0.99	1.00
7-yr Swap												
Bias	0.1045	0.0714	0.0941	-0.0958	-0.3041	-0.1207	-0.0653	-0.4500	-0.1264	-0.1112	-0.2488	-0.1297
RMSE	4.6647	4.6237	4.6735	4.7391	4.6627	4.7625	4.8280	4.8063	4.8641	4.8310	4.7520	4.8465
RMSE/PF	1.00	0.99	1.00	1.00	0.98	1.00	0.99	0.99	1.00	1.00	0.98	1.00
10-yr Swap												
Bias	0.0954	0.0640	0.0853	-0.1282	-0.3029	-0.1466	-0.0852	-0.4180	-0.1400	-0.1338	-0.2555	-0.1498
RMSE	4.4746	4.4012	4.4870	4.7899	4.6866	4.8236	4.8104	4.7632	4.8650	4.8350	4.7373	4.8560
RMSE/PF	1.00	0.98	1.00	0.99	0.97	1.00	0.99	0.98	1.00	1.00	0.98	1.00
1-yr Cap												
Bias	-0.6545	-0.6483	-0.6542	-0.0276	-0.0673	-0.0399	-0.1438	0.1145	-0.1071	-0.0064	0.0201	-0.0011
RMSE	6.6074	6.6645	6.7035	7.6802	7.8455	7.7070	6.5112	6.4334	6.2677	5.2711	5.4185	5.3739
RMSE/PF	0.99	0.99	1.00	1.00	1.02	1.00	1.04	1.03	1.00	0.98	1.01	1.00
5-yr Cap												
Bias	-0.4960	-0.4493	-0.4880	-0.1368	-0.1290	-0.1128	-0.9026	1.6071	-0.5310	-0.4481	-0.0352	-0.4221
RMSE	27.7210	28.7350	27.7130	46.5200	46.6780	46.3120	17.8450	18.8800	17.4990	11.7840	11.9470	11.7340
RMSE/PF	1.00	1.04	1.00	1.00	1.01	1.00	1.02	1.08	1.00	1.00	1.02	1.00

Notes: RMSE and Bias estimates are obtained from 300,000 simulated rates and prices (500 trajectories, 260 weeks), and the corresponding fitted values using the EKF, the UKF or the PF. The RMSE ratio of the filters to the PF is also reported. Caps are not used when filtering the states in this table, only LIBOR and swap rates are used for filtering.

Table 5: Rate and Price Fit of $A_M(3)$ Models. States Filtered using LIBOR, Swap Rates and Cap Prices.

	$A_0(3)$			$A_1(3)$			$A_2(3)$			$A_3(3)$		
	EKF	UKF	PF	EKF	UKF	PF	EKF	UKF	PF	EKF	UKF	PF
3-mo LIBOR												
Bias	10.1520	0.7175	0.5191	16.6930	0.5515	0.5683	-1.5280	-0.5461	-0.0163	-0.2374	-0.0167	-0.0072
RMSE	33.3070	8.9777	9.3658	54.0460	6.9102	7.6562	7.6306	3.5063	4.4995	6.0769	4.4107	4.5636
RMSE/PF	3.56	0.96	1.00	7.06	0.90	1.00	1.70	0.78	1.00	1.33	0.97	1.00
6-mo LIBOR												
Bias	5.5900	0.3637	0.2949	8.7181	0.1562	0.2845	-0.3165	-0.6816	-0.0108	-0.1942	-0.0437	-0.0151
RMSE	19.3050	6.5922	6.7378	29.1880	5.7043	5.8876	5.9207	4.6463	4.6207	5.8629	4.4898	4.6013
RMSE/PF	2.87	0.98	1.00	4.96	0.97	1.00	1.28	1.01	1.00	1.27	0.98	1.00
1-yr Swap												
Bias	0.8067	0.0374	0.0892	0.8806	-0.2271	0.0219	0.4927	-0.7029	-0.0023	-0.1032	-0.0749	-0.0094
RMSE	5.7578	4.2562	4.4439	7.0156	4.1431	4.3708	5.2823	4.6991	4.6926	5.5401	4.6478	4.7158
RMSE/PF	1.30	0.96	1.00	1.61	0.95	1.00	1.13	1.00	1.00	1.17	0.99	1.00
2-yr Swap												
Bias	-2.5148	-0.1705	-0.0695	-3.0517	-0.4767	-0.1448	0.6853	-0.6598	-0.0547	-0.0285	-0.1742	-0.0525
RMSE	10.6490	4.9183	5.0696	13.2450	4.6535	4.8939	5.0824	4.7661	4.7992	5.1960	4.7604	4.7992
RMSE/PF	2.10	0.97	1.00	2.71	0.95	1.00	1.06	0.99	1.00	1.08	0.99	1.00
5-yr Swap												
Bias	-4.1305	-0.1653	-0.0864	-2.4759	-0.4708	-0.1118	0.5728	-0.4975	-0.0794	0.1218	-0.2606	-0.0614
RMSE	14.9080	5.3275	5.4338	11.2470	5.0117	5.1677	5.1082	4.8391	4.9404	5.1675	4.8630	4.9290
RMSE/PF	2.74	0.98	1.00	2.18	0.97	1.00	1.03	0.98	1.00	1.05	0.99	1.00
7-yr Swap												
Bias	-4.3090	-0.1594	-0.0915	-1.9633	-0.4530	-0.1017	0.5052	-0.4626	-0.0976	0.1478	-0.2941	-0.0760
RMSE	15.2360	5.1308	5.2791	9.6844	4.9644	5.1554	5.1557	4.8175	4.9655	5.2173	4.8556	4.9449
RMSE/PF	2.89	0.97	1.00	1.88	0.96	1.00	1.04	0.97	1.00	1.06	0.98	1.00
10-yr Swap												
Bias	-4.3584	-0.1542	-0.0951	-1.5432	-0.4288	-0.0944	0.4446	-0.4332	-0.1122	0.1629	-0.3178	-0.0885
RMSE	15.1570	4.8770	5.0730	8.3952	4.9139	5.1353	5.1714	4.7820	4.9676	5.2866	4.8664	4.9719
RMSE/PF	2.99	0.96	1.00	1.63	0.96	1.00	1.04	0.96	1.00	1.06	0.98	1.00
1-yr Cap												
Bias	-0.6059	-0.7926	-0.6474	-0.2408	-0.3000	-0.1103	1.0803	-0.1672	-0.0027	0.0512	0.0005	0.0083
RMSE	6.9819	5.6874	4.9379	10.3270	6.7769	5.2854	6.0638	5.3421	5.6532	5.4416	5.3687	5.3305
RMSE/PF	1.41	1.15	1.00	1.95	1.28	1.00	1.07	0.94	1.00	1.02	1.01	1.00
5-yr Cap												
Bias	6.8152	-0.8712	-0.1509	6.2130	-0.6298	0.7970	4.7040	-0.0021	-0.1763	0.5735	-0.6274	-0.1753
RMSE	51.4360	25.3060	9.6486	93.0390	44.6070	14.7130	16.3620	12.7500	8.5306	9.0795	7.6691	6.8271
RMSE/PF	5.33	2.62	1.00	6.32	3.03	1.00	1.92	1.49	1.00	1.33	1.12	1.00

Notes: RMSE and Bias estimates are obtained from 300,000 simulated rates and prices (500 trajectories, 260 weeks), and the corresponding fitted values using the EKF, the UKF or the PF. The RMSE ratio of the filters to the PF is also reported. Caps as well as LIBOR and swap rates are used when filtering the states in this table.

Table 6: Rate and Price Forecasting Performance. States Filtered using LIBOR, Swap Rates and Cap Prices

Panel A: Average Forecast RMSE (ratios)

Forecast horizon	A ₀ (3) Model			A ₁ (3) Model			A ₂ (3) Model			A ₃ (3) Model		
	1 week	4 weeks	12 weeks	1 week	4 weeks	12 weeks	1 week	4 weeks	12 weeks	1 week	4 weeks	12 weeks
3-mo LIBOR	1.47	1.12	1.00	1.96	1.24	1.00	1.01	1.00	1.00	1.02	1.00	1.00
6-mo LIBOR	1.22	1.03	1.00	1.45	1.06	1.01	1.01	1.00	1.00	1.02	1.00	1.00
1-yr Swap	1.01	1.00	1.01	1.03	1.03	1.05	1.01	1.00	1.00	1.01	1.00	1.00
2-yr Swap	1.13	1.05	1.03	1.29	1.14	1.09	1.01	1.00	1.00	1.01	1.00	1.00
5-yr Swap	1.31	1.12	1.06	1.41	1.20	1.11	1.01	1.01	1.00	1.02	1.01	1.00
7-yr Swap	1.35	1.14	1.06	1.39	1.20	1.11	1.02	1.01	1.00	1.02	1.01	1.00
10-yr Swap	1.37	1.15	1.06	1.35	1.19	1.11	1.03	1.02	1.01	1.03	1.02	1.01
1-yr Cap	1.10	1.07	1.05	1.13	1.10	1.07	1.03	1.01	1.01	1.05	1.05	1.04
5-yr Cap	1.16	1.12	1.07	1.15	1.14	1.09	1.02	1.01	1.02	1.05	1.02	1.02

Panel B: Median Forecast RMSE (ratios)

Forecast horizon	A ₀ (3) Model			A ₁ (3) Model			A ₂ (3) Model			A ₃ (3) Model		
	1 week	4 weeks	12 weeks	1 week	4 weeks	12 weeks	1 week	4 weeks	12 weeks	1 week	4 weeks	12 weeks
3-mo LIBOR	1.04	1.03	1.01	1.06	1.05	1.00	1.01	1.00	1.00	1.01	1.01	1.00
6-mo LIBOR	1.04	1.02	1.00	1.05	1.03	1.01	1.00	1.00	1.00	1.01	1.00	1.00
1-yr Swap	1.01	1.00	1.01	1.02	1.02	1.04	1.01	1.01	1.00	1.01	1.00	1.00
2-yr Swap	1.04	1.03	1.02	1.06	1.05	1.04	1.01	1.00	1.00	1.01	1.00	1.00
5-yr Swap	1.05	1.05	1.04	1.08	1.08	1.07	1.01	1.01	1.00	1.02	1.00	1.01
7-yr Swap	1.05	1.06	1.05	1.11	1.09	1.07	1.02	1.01	1.01	1.01	1.01	1.00
10-yr Swap	1.06	1.06	1.05	1.11	1.09	1.07	1.02	1.02	1.01	1.03	1.02	1.01
1-yr Cap	1.01	1.02	1.02	0.97	1.00	1.00	1.02	1.01	1.02	1.00	1.00	1.00
5-yr Cap	1.04	1.01	1.02	0.87	0.98	1.00	1.01	1.01	1.02	1.04	1.02	1.03

Panel C: Forecast RMSE Standard Deviation (ratios)

Forecast horizon	A ₀ (3) Model			A ₁ (3) Model			A ₂ (3) Model			A ₃ (3) Model		
	1 week	4 weeks	12 weeks	1 week	4 weeks	12 weeks	1 week	4 weeks	12 weeks	1 week	4 weeks	12 weeks
3-mo LIBOR	8.83	3.74	1.01	8.37	5.29	1.03	1.06	1.01	1.01	1.05	1.02	1.02
6-mo LIBOR	7.76	1.48	1.00	9.01	2.03	1.05	1.04	1.00	1.01	1.05	1.02	1.02
1-yr Swap	1.03	1.00	1.02	2.16	1.37	1.37	1.02	1.00	1.01	1.04	1.01	1.02
2-yr Swap	4.18	1.69	1.13	7.58	3.62	1.79	1.01	1.00	1.00	1.03	1.01	1.02
5-yr Swap	7.92	3.19	1.38	6.73	3.80	1.80	1.02	1.00	1.01	1.06	1.02	1.02
7-yr Swap	8.54	3.48	1.43	5.79	3.27	1.67	1.06	1.01	1.01	1.09	1.03	1.01
10-yr Swap	8.61	3.58	1.44	4.94	2.71	1.51	1.11	1.02	1.01	1.12	1.03	1.01
1-yr Cap	1.30	1.16	1.09	1.40	1.41	1.35	1.20	1.04	1.09	1.51	1.61	1.44
5-yr Cap	1.29	1.23	1.15	1.31	1.36	1.33	1.03	1.04	1.09	1.13	1.10	1.17

Notes: We forecast rates and cap prices using the EKF and UKF filters. For each of 500 simulations, we compute the forecast RMSE. The EKF to UKF ratio of the mean, median, and standard deviation of these RMSEs is reported. States are using LIBOR, swap rates and cap prices.

Table 7: Fit of 7-Year Cap Prices. $A_M(3)$ Models

Panel A: States Filtered using LIBOR and Swap Rates Only

7-yr Cap	$A_0(3)$			$A_1(3)$			$A_2(3)$			$A_3(3)$		
	EKF	UKF	PF	EKF	UKF	PF	EKF	UKF	PF	EKF	UKF	PF
Bias	-0.454	-0.348	-0.400	-0.210	-0.086	-0.151	-1.259	2.325	-0.714	-0.669	-0.007	-0.594
RMSE	40.193	41.460	39.922	64.756	64.868	64.411	22.547	24.141	22.257	15.580	15.782	15.543
RMSE/PF	1.01	1.04	1.00	1.01	1.01	1.00	1.01	1.08	1.00	1.00	1.02	1.00

Panel B: States Filtered using LIBOR, Swap Rates and 1-yr and 5-yr Caps

7-yr Cap	$A_0(3)$			$A_1(3)$			$A_2(3)$			$A_3(3)$		
	EKF	UKF	PF	EKF	UKF	PF	EKF	UKF	PF	EKF	UKF	PF
Bias	15.447	-0.662	0.274	13.208	-0.416	1.484	6.026	0.206	-0.257	0.910	-0.943	-0.263
RMSE	88.811	40.422	17.666	141.490	64.675	24.153	22.875	17.017	13.484	13.979	11.517	10.986
RMSE/PF	5.03	2.29	1.00	5.86	2.68	1.00	1.70	1.26	1.00	1.27	1.05	1.00

Notes: Estimates of RMSE and Bias are obtained from 130,000 simulated 7-year cap prices (500 trajectories, 260 weeks), and the corresponding fitted prices using the EKF, the UKF or the PF. In Panel A only LIBOR and swap rates are used when filtering the underlying states; the PF uses 3000 particles. In Panel B, 1-year and 5-year caps are used in addition when filtering the state; the PF uses 300 particles. 7-year caps are not used in the filtering step.

Table 8: Estimation Errors (Hessian-Scaled Scores) for the $A_1(3)$ Model

	Percentiles						
	5 th	10 th	25 th	50 th	75 th	90 th	95 th
δ_0							
EKF	-7.56	-1.07	-0.05	0.00	0.12	4.26	24.37
UKF	-0.02	0.00	0.00	0.00	0.00	0.00	0.01
δ_{11}							
EKF	-11.38	-1.12	-0.01	0.00	0.00	0.51	5.42
UKF	0.00	0.00	0.00	0.00	0.00	0.00	0.00
δ_{12}							
EKF	-1.35	-0.09	0.00	0.00	0.00	0.09	1.35
UKF	0.00	0.00	0.00	0.00	0.00	0.00	0.00
δ_{13}							
EKF	-1.56	-0.14	0.00	0.00	0.00	0.08	0.95
UKF	0.00	0.00	0.00	0.00	0.00	0.00	0.00
κ_{11}							
EKF	-7773.80	-1106.00	-30.85	0.00	51.55	1554.90	9238.50
UKF	-8.65	-2.63	-0.38	-0.01	0.49	3.07	7.21
κ_{22}							
EKF	-961.18	-150.88	-2.94	0.01	4.17	280.49	1213.90
UKF	-2.16	-0.13	-0.02	0.00	0.01	0.07	0.73
κ_{32}							
EKF	-2447.10	-544.96	-10.29	0.00	7.60	414.53	2178.60
UKF	-1.91	-0.26	-0.04	0.00	0.04	0.22	2.78
κ_{33}							
EKF	-3100.40	-970.22	-16.64	0.00	9.61	598.31	3205.90
UKF	-2.23	-0.26	-0.06	0.00	0.09	0.54	6.70
θ_1							
EKF	-13058.00	-2790.90	-47.67	-0.01	54.79	1585.40	7708.60
UKF	-14.67	-3.47	-0.36	0.00	0.21	1.81	7.90
λ_{01}							
EKF	-12046.00	-2175.60	-48.04	0.00	28.47	989.17	6933.60
UKF	-7.25	-2.95	-0.49	0.00	0.41	3.17	10.80
λ_{02}							
EKF	-541.94	-63.69	-0.74	0.00	0.35	35.99	193.60
UKF	-0.25	-0.06	0.00	0.00	0.00	0.07	0.23
λ_{03}							
EKF	-1475.40	-254.16	-4.55	0.00	12.16	809.33	4262.30
UKF	-1.96	-0.45	-0.07	0.00	0.04	0.29	1.60

Notes: This table reports the distribution of asymptotic estimation errors for the $A_1(3)$ model at the optimal parameters. In each of 500 simulations, we use the EKF or the UKF to filter the states using caps for the optimal and perturbed parameters ($\varepsilon=10^{-6}$). Given the obtained rates and prices, we then compute the gradient of the log-likelihood associated with the observation equation. The Hessian matrix used to scale the scores is the same for both Kalman filters as it is obtained assuming perfect filtering (i.e. using the simulated states x_t).