

# Short rate forecasting based on the inference from the CIR model for multiple yield curve dynamics \*

LIN-YEE HIN<sup>†</sup> AND NIKOLAI DOKUCHAEV<sup>‡</sup>

Department of Mathematics & Statistics, Curtin University,  
GPO Box U1987, Perth, 6845 Western Australia

October 6, 2015

1

## Abstract

We propose a strategy to extract the information on the market participants' expectation of the future short rate from the cross-sectional zero coupon bond prices. In line with the current market practice of building different yield curves for different tenors, we construct multiple one-factor short rate processes to pin down the salient features of the yield curve at different tenors. We represent this information in the form of the Cox-Ingersoll-Ross model implied parameters, and show that this information can be used to forecast the future short rate. This approach of representing the information on the market participants' consensus in the form of implied model parameters and using these implied parameters for forecasting purposes resembles the approach of representing the market expectation of the underlying asset volatility reflected by stock option prices in the form of implied volatility, and using it to forecast the realized volatility. We illustrate the implementation of this method using historical US STRIPS prices and effective Federal Funds Rate.

---

\*The paper has particularly benefited from helpful comments and valuable suggestions made by an anonymous referee. We acknowledge provision of ICT support and computing resources by Curtin IT Services <http://cits.curtin.edu.au>. Curtin Information Technology Services (CITS) provides Information and Communication Technology systems and services in support of Curtin's teaching, learning, research and administrative activities. We acknowledge use of computing resources from the NeCTAR Research Cloud <http://www.nectar.org.au>. NeCTAR is an Australian Government project conducted as part of the Super Science initiative and financed by the Education Investment Fund. This work was supported by ARC grant of Australia DP120100928.

<sup>†</sup>Corresponding author. Department of Mathematics and Statistics, Curtin University, GPO Box U1987, Perth, 6845 Western Australia; email: [linyeehin@gmail.com](mailto:linyeehin@gmail.com)

<sup>‡</sup>Department of Mathematics and Statistics, Curtin University, GPO Box U1987, Perth, 6845 Western Australia; email: [N.Dokuchaev@curtin.edu.au](mailto:N.Dokuchaev@curtin.edu.au)

<sup>1</sup>This is a pre-copy-editing, author-produced PDF of an article accepted following peer review for publication in the Annals of Financial Economics 2015 World Scientific Publishing Company. The definitive publisher-authenticated version will be available at Journal URL <http://www.worldscientific.com/worldscinet/afe>

Keywords: interest rates, CIR spot rate model, optimization.

JEL codes: C51, C58, C61, G12, G17

## 1 Introduction

The classical asset pricing paradigm assumes the existence of a unique yield curve to forecast interest rates, and to discount cash flows. However, the subprime mortgage crisis of 2007 led to segmentation of interest rates in the market where the quoted interest rates are no longer consistent with one another. For example, the forward rates implied by two consecutive zero coupon bonds are no longer comparable to the forward rate agreement rate quotes or the forward rates implied by the overnight indexed swap rate quotes, thereby violating the usual no-arbitrage relationships (see, e.g., Mercurio, 2009; 2010; Moreni and Pallavicini, 2014; Baviera and Cassaro, 2015). Practitioners respond to this change by moving away from the classical single-curve framework to adopt the multi-curve framework of constructing as many yield curves as possible rate tenors to forecast future interest rates, and constructing a different set of curves to discount future cash flows. In the multi-curve framework, interest rates of different tenors are assumed to be associated with different short rate processes (see, e.g., Mercurio, 2010; Moreni and Pallavicini, 2014). In fact, the recent empirical evidence of interest rate market segmentation may be viewed as a verification of the so-called ‘Market Segmentation Hypothesis’ which suggests that different categories of market participants may prefer to invest in interest rate related instruments with different maturities for various reasons, including investment horizon preferences, institutional factors, and regulations (see, e.g., Culberston, 1957; Modigliani and Sutch, 1966; Van Horne, 1980). Extant literature in this area largely address the issue of coherent contingent claim valuation in the multi-curve framework. This is an emerging area of mathematical finance that is of practical importance. Historically, the recent adoption of the multi-curve framework interest rate modelling resembles the adoption of models to capture the volatility smile in that both shifts of paradigm were ushered in by empirical observations in the financial market.

In this paper, we propose a strategy to model the interest rates implied from the cross-sectional zero coupon bond prices within the multi-curve framework. We organize the cross section of yields in multiple triplets of zero coupon bonds in ascending order of tenor. Assuming that each triplet that corresponds to a different tenor interval is associated with a different short rate process, we model each of these short rate processes using a separate single-factor Cox-Ingersoll-Ross (CIR) model (Cox et al., 1985) specific for each of the tenor intervals considered. We extract the information on the market participants' aggregate view of the future short rate at various forecast horizons from the triplets with comparable tenors, and represent this information as different sets of implied CIR model parameters corresponding to these tenor specific short rate processes. This approach is similar to representing the market expectation of the underlying asset volatility reflected by stock option prices in the form of implied volatility. Our approach departs from the multi-curve HJM model (Moreni and Pallavicini, 2014), and the multi-curve LIBOR market model (Mercurio, 2010) in that we consider modelling the tenor specific short rate processes instead of modelling tenor specific basis spreads. Additionally, our strategy also differs from that proposed by Madan and Schoutens (2012) wherein the tenor specific single-factor CIR model is used to construct tenor specific discount curve to pin down different segments of the yield curve in a two-price economy, a pricing framework different from the one we consider.

Using these model parameters, we derive an algorithm to predict the future short rate. Additionally, we demonstrate that it is possible to make use of the convexity adjustment to estimate the variance of the integrated short rate process. We apply our proposed algorithm to 14 years of daily United States Separately Traded Registered Interest and Principal Securities (US STRIPS) prices to forecast the US effective Federal Funds rate. We report two empirical findings. Firstly, for the data we investigate, the forecast of future short rate under risk neutral measure, i.e., based on cross-sectional data, is reasonably accurate up to six months ahead. Deterioration in forecast performance beyond six months may be due to noticeable effect of the term premium. This implies that forecast of future short rate under physical measure, i.e., based on time series data, may be necessary in order to disentangle the effect of term premium, and achieve accurate forecast performance beyond six months. Secondly, the prices of zero coupon bonds with different tenors appear

to contain information on the market participants' expectation of future short rate at forecasting horizons comparable to their tenors. However, it appears that the prices of zero coupon bonds does not contain sufficient information to predict future short rate at forecast horizon beyond their tenors. We interpret this as an indication that the single-factor CIR model may have a role in interest rate forecast in the multi-curve framework. On the other hand, it verifies the well known fact that single-factor CIR model does not perform well in the single-curve framework, and is superseded by the multi-factors models (see, e.g., Litterman and Scheinkman, 1991; Duffie and Kan, 1996; Dai and Singleton, 2002; Duffee, 2002; Diebold and Li, 2006; Collin-Dufresne et al., 2008; Duffee, 2011; Huang et al., 2014).

The classical interest rate forecast literature within the single-curve framework is voluminous. Within the single-curve framework, Diebold and Li (2006), Diebold et al. (2008), Moench (2008), and Guidolin and Timmermann (2009), among many others, forecast interest rates based on the physical measure, while de Munnik and Schotman (1994), and Bams and Schotman (2003), among others, do so based on the risk neutral measure. With the recent adoption of the multi-curve framework for pricing and hedging purposes in the banking and financial industry, we attempt to contribute to the literature by proposing a strategy to forecast interest rate within the multi-curve framework. As far as we are aware, there has been no previous published literature along this line.

We choose to work with the single-factor CIR model because it is the simplest possible model that guarantees non-negativity of the short rate, and its zero coupon bond pricing formula is available in closed form. Madan and Schoutens (2012) adopted the single-factor CIR model in the multi-curve framework for a two-price economy for similar reasons. Additionally, Aït-Sahalia (1996) found empirical evidence that if the magnitude of the short rate is less than 0.09 the CIR model is a good approximation model for the seven-day Eurodollar deposit short rate dynamics in the single-curve framework. Similar to the role of the Black-Scholes option pricing formula in the equity sphere, the CIR model has been used by many researchers as a starting point in the single-curve framework from which more sophisticated models are constructed to better capture the salient features of the interest rate term structure (see, e.g., Hull and White, 1990; Jamshidian, 1995; Maghsoodi, 1996; Schlögl and Schlögl, 2000; Ahn and Thompson, 1988; Deng, 2014).

Our approach differs from the segmented CIR model (Schlögl and Schlögl, 2000) in several aspects. Firstly, the segmented CIR model is constructed within the single-curve framework, while we work within the multi-curve framework. Secondly, we assume that all three CIR model parameters may vary across different tenors, and estimate them from the corresponding triplets of zero coupon prices corresponding to different tenors, while the segmented CIR model keep the same speed of mean reversion fixed across all tenors, specify the volatilities as exogenous input based on the volatility parameters of some contingent claims of interest, and estimate the long-range mean short rate as a piece-wise constant that span the time intervals between the maturity dates of two adjacent zero coupon bonds arranged in increasing tenor. Thirdly, while Schlögl and Schlögl (2000) note that strongly downward sloping initial forward rate curves can lead to negative values of long-range mean short rate in their framework, our estimation procedure ensures that all the estimated CIR model parameters are positive.

We analyse the interest rate forecast performance under the risk neutral measure as we intend to investigate the impact of the term premium on the forecasting performance of our predictor. Fama (1976) discovered that the forward rates calculated from T-Bill rates do not perform well in predicting short rate at lead time periods beyond two years unless adjustment is made to account for the term premium, i.e., the compensation for enduring interest rate risk during the holding period of the bond. Complementary to this finding, Longstaff (2000) discovered that term premium for short-term repurchase rates is negligible in the short end of the term structure for tenors up to three months. Since Fama (1976) and Longstaff (2000) considered two different sets of interest rates, it is unclear at what point along the tenor horizon will the term premium start to show noticeable effect on the forecasting performance of the future short rate predictor constructed from the same set interest rates. In this paper, we seek to clarify this point for our predictor and the forward rate calculated from the zero coupon bond prices.

While it is known that the forecasting performance of short rate predictors can be improved if the expected future short rate can be disentangled from the expected term premium (see, e.g., Fama, 1976; Fama and Bliss, 1987; Huang and Lin, 1996; Dai and Singleton, 2002; Tzavalis, 2004; Cochrane and Piazzesi, 2005), there exists different definitions of term premium, and var-

ious methods have been proposed to estimate them (see, e.g., Fama and Bliss, 1987; Cochrane and Piazzesi, 2005; Collin-Dufresne et al., 2009; Gil-Alana and Moreno, 2012; Rudebusch, 2012; Dewachter et al., 2014). The estimation of the term premium is sensitive to the choice of estimator as well as the sample of data considered, rendering its practical implementation a challenging task. Kim (2007) noted that this lack of robustness in existing methodologies makes it less appealing to practitioners to adjust for the term premium in their analysis of interest rates.

The rest of the paper is as follows. Section 2 outlines some characteristics of the short rate process and the CIR model. Section 3 constructs the estimator of CIR model parameters in the multi-curve framework, describes the parameter estimation algorithm, and derives the predictor for future short rate and the estimator for the forward variance of the integrated CIR process. Section 4 reports the application of the proposed strategy to the historical US STRIPS prices and effective Federal Funds Rate. Section 5 concludes.

## 2 The model setting

Let the short rate process  $r(u), t \leq u \leq T$ , be a stochastic process. The price of a zero-coupon bond at the current time  $t$  and with maturity time  $T$  is defined by the conditional expectation

$$P(t, T) = \mathbf{E} \left[ e^{-\int_t^T r(u) du} \middle| \mathcal{F}_t \right], \quad (1)$$

where  $\mathbf{E}$  is the expectation taken over  $r(u)$  under some probability measure  $\mathbf{P}$ , and  $\{\mathcal{F}_t\}$  is some filtration generated by the flow of the currently observed market data. For our purposes, it is not necessary to specify the dynamics of  $r(u)$  and the choice of  $\mathbf{P}$  and  $\{\mathcal{F}_t\}$ . We assume, however, that the process  $r(u)$  is adapted to the filtration  $\{\mathcal{F}_t\}$ , that  $r(u)$  is non-negative, or at least that the process  $\min(r(u), 0)$  is bounded, and that  $\mathbf{E} \int_t^T r(u)^2 du < +\infty, \forall t \leq T$ . The case when the filtration  $\{\mathcal{F}_t\}$  is generated by the process  $r(u)$  is not excluded.

## 2.1 Some characteristics of the short rate process

We will consider the ‘expected average integrated short rate’ observed at time  $t$  for the time interval  $[t, T]$  defined as

$$\rho(t; s, T) \triangleq \mathbf{E} \left[ \frac{1}{T-s} \int_s^T r(u) du \mid \mathcal{F}_t \right], \quad t \leq s \leq u \leq T. \quad (2)$$

Let

$$\eta(t, T) = -\frac{1}{T-t} \log P(t, T)$$

be the yield-to-maturity of the zero coupon bond maturing at the tenor  $T - t$ , and let

$$J(t, T) = \rho(t; t, T) - \eta(t, T) \quad (3)$$

be the corresponding ‘convexity adjustment’, (see, e.g. Fisher and Gilles, 1998). By the Jensen’s inequality,

$$\mathbf{E} \left[ e^{-\int_t^T r(u) du} \mid \mathcal{F}_t \right] \geq e^{-\mathbf{E} \left[ \int_t^T r(u) du \mid \mathcal{F}_t \right]},$$

and  $J(t, T) \geq 0$ . Clearly,  $J(t, T) = 0$  if the short rate process  $r$  is non-random. Let  $\eta(t, T_1)$  and  $\eta(t, T_2)$  be the yield-to-maturities of two zero coupon bonds  $P(t, T_1)$  and  $P(t, T_2)$  respectively where  $t < T_1 < T_2$ . The forward rate in the interval  $[T_1, T_2]$  can be inferred from these zero coupon bond prices by

$$F(t; T_1, T_2) = \frac{-\log P(t, T_2) + \log P(t, T_1)}{T_2 - T_1} = \frac{(T_2 - t) \eta(t, T_2) - (T_1 - t) \eta(t, T_1)}{T_2 - T_1}. \quad (4)$$

We may use (3) and (4) to express the relation between  $\rho(t; T_1, T_2)$  and  $F(t; T_1, T_2)$  as

$$\rho(t; T_1, T_2) - F(t; T_1, T_2) = \frac{(T_2 - t) J(t, T_2) - (T_1 - t) J(t, T_1)}{T_2 - T_1}. \quad (5)$$

The difference between  $\rho(t; T_1, T_2)$  and  $F(t; T_1, T_2)$  is the weighted difference between the convexity premiums for the intervals  $[t, T_2]$  and  $[t, T_1]$  respectively. Let

$$\gamma(t; s, T) = \mathbf{E} \left[ \int_s^T r(u) du \mid \mathcal{F}_t \right], \quad \Gamma(t; s, T) = \mathbf{E} \left[ \left( \int_s^T r(u) du \right)^2 \mid \mathcal{F}_t \right], \quad t \leq s \leq u \leq T.$$

We expand  $e^{-\int_t^{T_1} r(u)du}$  in (1) up to the second order to obtain

$$P(t, T_k) \approx 1 - \gamma(t; t, T_k) + \frac{1}{2}\Gamma(t; t, T_k) ,$$

and expand the log term up to the second order to obtain

$$\begin{aligned} & \log\left(1 - \gamma(t; t, T_k) + \frac{1}{2}\Gamma(t; t, T_k)\right) \\ & \approx -\gamma(t; t, T_k) + \frac{1}{2}\Gamma(t; t, T_k) - \frac{1}{2}\left(-\gamma(t; t, T_k) + \frac{1}{2}\Gamma(t; t, T_k)\right)^2 \approx \log P(t, T_k) . \end{aligned}$$

Discarding the terms higher than the second order, we obtain

$$\begin{aligned} \log P(t, T_k) & \approx -\gamma(t; t, T_k) + \frac{1}{2}\Gamma(t; t, T_k) - \frac{1}{2}\gamma(t; t, T_k)^2 \\ & = -\gamma(t; t, T_k) + \frac{1}{2} \text{Var}\left[\int_t^{T_k} r(u)du \mid \mathcal{F}_t\right] . \end{aligned} \quad (6)$$

We substitute (6) into (4) for  $k = 1, 2$  to obtain, to the second order of approximation,

$$F(t; T_1, T_2) \approx \rho(t; T_1, T_2) - \frac{\delta(t; T_1, T_2)}{2(T_2 - T_1)} ,$$

where

$$\delta(t; T_1, T_2) = \text{Var}\left[\int_t^{T_2} r(u)du \mid \mathcal{F}_t\right] - \text{Var}\left[\int_t^{T_1} r(u)du \mid \mathcal{F}_t\right] .$$

Let

$$\Delta(t; T_1, T_2) = 2(\rho(t; T_1, T_2) - F(t; T_1, T_2)) , \quad (7)$$

we obtain

$$\rho(t; T_1, T_2) - F(t; T_1, T_2) \approx \frac{\delta(t; T_1, T_2)}{2(T_2 - T_1)} \approx \frac{1}{2}\Delta(t; T_1, T_2) . \quad (8)$$

We can regard  $\Delta(t; T_1, T_2)$  as the ‘approximate average forward variance’, a gauge of the volatility of the integrated short rate in the interval  $[T_1, T_2]$  that can be approximated using  $\delta(t; T_1, T_2)$ .

## 2.2 The Cox-Ingersoll-Ross model for the short rate process

The definitions in Section 2.1 do not rely on a particular model for the short rate process. Moreover, we do not assume that a particular short rate model can capture all the salient features of the interest

rate term structure. However, for the reasons outlined in Section 1, we choose to use the CIR model as a working model to describe the dynamics of the tenor specific short rate processes in a multi-curve framework. We will be using this model as a mapping tool to extract the information on the market participants' aggregate view of the future short rate at various forecast horizons from the triplets of zero coupon bond prices with comparable tenors, and represent this information as different sets of implied CIR model parameters corresponding to the short rate processes for these tenors.

The CIR model for the short rate process  $r_{CIR}(u), t \leq u \leq T$ , is assumed to satisfy the following stochastic Itô equation

$$r_{CIR}(T) = r_{CIR}(t) + \int_t^T \kappa(\theta - r_{CIR}(u))du + \int_t^T \sigma \sqrt{r_{CIR}(u)}dW(u), \quad t < T, \quad (9)$$

where  $\kappa$  is the mean-reversion speed of the short rate,  $\theta$  is the long-range short rate,  $\sigma$  is the volatility of the process,  $r_{CIR}(t) = r(t)$  where  $r(t)$  is the short rate at time  $t$ , and  $W(u)$  is the standard Wiener process. The solution has the following properties: (1) if  $\kappa = 0$  or  $\theta = 0$ ,  $r_{CIR}(T)$  reaches zero almost surely and the point zero is absorbing, (2) if  $2\kappa\theta \geq \sigma^2$ ,  $r_{CIR}(T)$  is a transient process that stays positive and never reaches zero, and (3) if  $0 < 2\kappa\theta < \sigma^2$ ,  $r_{CIR}(T)$  is instantaneously reflective at point zero (Feller, 1951). It is known that

$$\mathbf{E}[r_{CIR}(T)|\mathcal{F}_t] = \theta + (r(t) - \theta) e^{-\kappa(T-t)} \quad (10)$$

and

$$\gamma_{CIR}(t; t, T) = \theta(T - t) + \left( \frac{r(t) - \theta}{\kappa} \right) (1 - e^{-\kappa(T-t)}) . \quad (11)$$

In particular,

$$\rho_{CIR}(t; t, T) = \theta + \frac{r(t) - \theta}{\kappa(T - t)} (1 - e^{-\kappa(T-t)}) \quad (12)$$

(see, e.g., Dufresne, 2001). Using (2) and (12), we obtain the expected forward average integrated short rate as

$$\rho_{CIR}(t; T_1, T_2) = \theta + \left( \frac{r(t) - \theta}{\kappa(T_2 - T_1)} \right) (e^{-\kappa(T_1-t)} - e^{-\kappa(T_2-t)}) , \quad (13)$$

where, as  $T_1 \rightarrow \infty$  and  $T_2 \rightarrow \infty$ ,  $\rho_{CIR}(t; T_1, T_2) \rightarrow \theta$ .

Additionally, it follows from (3) that  $\rho_{CIR}(t; t, T) = \eta(t, T) + J_{CIR}(t, T)$ , and the relation between  $\rho_{CIR}(t; T_1, T_2)$  and  $F(t; T_1, T_2)$  can be expressed as

$$\rho_{CIR}(t; T_1, T_2) - F(t; T_1, T_2) = \frac{(T_2 - t) J_{CIR}(t, T_2) - (T_1 - t) J_{CIR}(t, T_1)}{T_2 - T_1}. \quad (14)$$

From (8), it is clear that (14) may be used as an estimator of the approximate average forward variance when the short rate dynamics is modelled using the CIR process, i.e.,

$$\Delta_{CIR}(t; T_1, T_2) = 2(\rho_{CIR}(t; T_1, T_2) - F(t; T_1, T_2)), \quad (15)$$

and the weighted difference between the convexity adjustment terms  $J_{CIR}(t, T_1)$  and  $J_{CIR}(t, T_2)$  is approximately half the average difference between variance of the integrated CIR short rate process spanning  $[t, T_2]$  and  $[t, T_1]$

$$\frac{(T_2 - t) J_{CIR}(t, T_2) - (T_1 - t) J_{CIR}(t, T_1)}{T_2 - T_1} \approx \frac{\delta_{CIR}(t; T_1, T_2)}{2(T_2 - T_1)},$$

where

$$\delta_{CIR}(t; T_1, T_2) = \text{Var}_{CIR} \left[ \int_t^{T_2} r_{CIR}(u) du \middle| \mathcal{F}_t \right] - \text{Var}_{CIR} \left[ \int_t^{T_1} r_{CIR}(u) du \middle| \mathcal{F}_t \right],$$

and  $\text{Var}_{CIR}(\cdot)$  denote the variance expressions for the integrated CIR process.

### 3 Inferring CIR model parameters from cross section zero coupon bond prices

We organize the set of cross section zero coupon bonds data in increasing tenors, and group them into successive triplets. Assuming that each triplet that corresponds to a different tenor interval is associated with a different short rate process, we model each of these short rate processes using a separate single-factor CIR model specific for each of the tenor intervals considered. Since the zero coupon bond pricing formula for the CIR model under the risk-neutral measure is given by

$$P_{CIR}(t, T; \kappa, \theta, \sigma, r(t)) = e^{A_1 - A_2 r(t)}, \quad (16)$$

where

$$\begin{aligned}
A_1 &= \frac{\kappa\theta}{\sigma^2} \left( (\kappa + A_3)(T - t) - 2 \log \left( 1 + \frac{(\kappa + A_3)(e^{A_3(T-t)} - 1)}{2A_3} \right) \right), \\
A_2 &= \frac{2(e^{A_3(T-t)} - 1)}{(\kappa + A_3)(e^{A_3(T-t)} + 2A_3)}, \\
A_3 &= \sqrt{\kappa^2 + 2\sigma^2},
\end{aligned}$$

and  $r(t)$  is the initial short rate, we construct a system of nonlinear equations for each tenor interval considered by mapping the corresponding triplet of cross section zero coupon bond prices to their respective set of implied CIR model parameters via (16). Using this approach, we extract the information on the market participants' aggregate view of the future short rate at various forecast horizons from the triplets with comparable tenors, and represent this information as different sets of implied CIR parameters corresponding to these short rate curves. This approach is similar to representing the market expectation of the underlying asset volatility reflected by stock option prices in the form of implied volatility.

### 3.1 CIR model parameters as solution to a system of nonlinear equations

Let  $P(t, T_{j,k})$ ,  $j = 1, \dots, n$ ,  $k = 1, 2, 3$ , be the prices of a set of zero coupon bonds observed at time  $t$  maturing at different times  $T_{j,k}$  that are strictly increasing in  $j$  and  $k$  where  $T_{j,k+1} > T_{j,k} > t$  and  $T_{j+1,1} > T_{j,3} > t$ ,  $j = 1, \dots, n$ . We adopt the multi-curve framework (see, e.g., Mercurio, 2010; Madan and Schoutens, 2012; Moreni and Pallavicini, 2014) in the modelling of the interest rate term structure. We assume that each triplet of prices  $\{P(t, T_{j,k})\}_{k=1}^3$  that corresponds to a different tenor interval is associated with a different short rate process, and they can be mapped to a triplet of CIR model parameters  $(\kappa_j(t), \theta_j(t), \sigma_j(t))$ . We express the dependence of CIR model parameters on time  $t$  to indicate that these parameters are to be estimated from the cross section set of zero coupon prices observed at time  $t$ . If all three model parameters are regarded as unknown variables, they cannot be uniquely determined from the price of one zero coupon bond because there is only one equation but there are three unknown variables. Based on this assumption, we formulate a system of nonlinear equations mapping the observed prices of the set of zero coupon bonds to the

model parameters  $\kappa_j(t)$ ,  $\theta_j(t)$ , and  $\sigma_j(t)$  via the CIR zero coupon bond pricing formula (16) as

$$P_{CIR}(t, T_{j,k}; \kappa_j(t), \theta_j(t), \sigma_j(t), r(t)) = P(t, T_{j,k}), \quad j = 1, \dots, n, \quad k = 1, 2, 3, \quad (17)$$

where  $(\kappa_j(t), \theta_j(t), \sigma_j(t))$  is the set of CIR model parameters inferred from the corresponding set of zero coupon prices  $\{P(t, T_{j,k})\}_{k=1}^3$ . For each set of three nonlinear equations, there are three unknown parameters.

Some extant literature (see, e.g., Brown and Dybvig, 1986; de Munnik and Schotman, 1994) regard the initial short rate  $r(t)$  as a latent state variable, and estimate it alongside  $\kappa_j(t)$ ,  $\theta_j(t)$ , and  $\sigma_j(t)$ , while others (see, e.g., Aït-Sahalia, 1996; Chapman et al., 1999; Chapman and Pearson, 2001) choose interest rates observed in the market as proxies for  $r(t)$ . While Schlögl and Schlögl (2000) raised the concern that using a market rate, such as an overnight rate, as a proxy for  $r(t)$  may not match the term structure implied by the CIR model parameters well, the empirical results reported by Chapman et al. (1999) demonstrated that in an affine bond price model such as the single-factor CIR model, accurate estimates of the drift and diffusion of the short-rate process can be obtained even using proxies for  $r(t)$  with maturities as long as three months for US data. In this paper, we represent  $r(t)$  by an observable overnight interest rate in the same economy as  $P(t, T_{j,k})$  that reflect the credit and liquidity risk levels that are comparable to those of  $P(t, T_{j,k})$ . We work in a multi-curve framework where we assume that the proxy short rate process  $\tilde{r}_j(t')$ ,  $t' \in [t, T_{j,3}]$ ,  $\tilde{r}_j(t) = r(t)$ , can be used to model the short rate process  $r_j(t')$ ,  $t' \in [t, T_{j,3}]$ ,  $r_j(t) = r(t)$ , associated with the triplet  $\{P(t, T_{j,k})\}_{k=1}^3$ .

### 3.2 Numerical approximate solution of nonlinear equations

For a current time  $t$ , we suggest to seek numerical approximation to the solution of the unknown parameters in (17). Let  $\boldsymbol{\kappa}(t) = (\kappa_1(t), \dots, \kappa_n(t))$ ,  $\boldsymbol{\theta}(t) = (\theta_1(t), \dots, \theta_n(t))$ ,  $\boldsymbol{\sigma}(t) = (\sigma_1(t), \dots, \sigma_n(t))$ , where  $\kappa_j(t) > 0$ ,  $\theta_j(t) > 0$ , and  $\sigma_j(t) > 0$  for  $j = 1, \dots, n$ . Let  $\boldsymbol{\Xi}(t) = (\Xi_1(t), \dots, \Xi_{3n}(t)) = (\boldsymbol{\kappa}(t), \boldsymbol{\theta}(t), \boldsymbol{\sigma}(t))$ . We seek approximate solution to (17) by

$$\text{Minimize } \Psi(\boldsymbol{\Xi}(t)) \quad \text{over } \boldsymbol{\Xi}(t) \quad (18)$$

where

$$\Psi(\Xi(t)) = \sum_{j=1}^n \sum_{k=1}^3 \left( P(t, T_{j,k}) - P_{CIR}(t, T_{j,k}; \kappa_j(t), \theta_j(t), \sigma_j(t), r(t)) \right)^2 \quad (19)$$

is the  $L_2$ -norm objective function. Since (19) is not convex with respect to  $\kappa_j(t)$ ,  $\theta_j(t)$ , and  $\sigma_j(t)$ , we seek numerical approximation to  $\Xi(t)$  via the Zhang-Sanderson's differential evolution algorithm (Zhang and Sanderson, 2009) that is designed to locate the global minimum in the presence of multiple co-existing local minima.

For the rest of this subsection, we describe the implementation of the Zhang-Sanderson's algorithm to solve (17). We follow the exposition style of Zhang and Sanderson (2009). We will consider iteration steps  $g = 1, \dots, G$ , with  $g = 0$  as the initialization stage.

*Initialization stage ( $g = 0$ ):* Generate a set  $\Xi_{I,0}(t) = \{(\Xi_{J,I,0}(t))_{J=1,\dots,3n}, I = 1, \dots, N_p\} \subset \mathbb{R}^{3n}$ , where  $\Xi_{J,I,0}(t)$  is a random number drawn from the uniform distribution  $U[\Xi_J^{Lower}(t), \Xi_J^{Upper}(t)]$ , and where  $[\Xi_J^{Lower}(t), \Xi_J^{Upper}(t)]$  is a user specified interval within which  $\Xi_J(t)$ ,  $J = 1, \dots, 3n$ , is assumed to be located.

*Iteration stage ( $g = 1, \dots, G$ ):* After initialization, the algorithm enters an iterative loop indexed by  $g = 1, \dots, G$ . Each iteration consists of three steps known as *mutation*, *crossover* and *selection*.

*Mutation step for iteration stage  $g$ :* We construct vectors  $\mathbf{v}_{I,g}(t) = \{v_{J,I,g}(t)\}_{J=1}^{3n}$ ,  $I = 1, \dots, N_p$ , by

$$\mathbf{v}_{I,g}(t) = \Xi_{I,g}(t) + F_{I,g}(t) \left( \Xi_{best,g}^p(t) - \Xi_{I,g}(t) \right), \quad I = 1, \dots, N_p,$$

where  $\Xi_{best,g}^p(t)$  is randomly chosen from the subset of  $\{\Xi_{I,g}(t), I = 1, \dots, N_p\}$  such that  $\Psi(\Xi_{best,g}^p(t))$  is located within the lowest  $p$  percentiles among the values  $\{\Psi(\Xi_{I,g}(t)), I = 1, \dots, N_p\}$ , while  $F_{I,g}(t)$  is drawn randomly from the Cauchy distribution with location parameter  $\mu_{F,g}(t)$  and scale parameter 0.1, and then to be either truncated to 1 if  $F_{I,g}(t) \geq 1$  or regenerated if  $F_{I,g}(t) \leq 0$ . Let  $\mu_{F,0}(t) = 0.5$ , let  $\mu_{F,g+1}(t) = (1 - c)\mu_{F,g}(t) + c\mu_{L,S_{F,g,g}}(t)$ , where  $c$  is a user specified parameter,  $S_{F,g}(t) = \{F_{I,g}(t) : \Xi_{I,g}(t) \neq \Xi_{I,g+1}(t)\}$ , and  $\mu_{L,S_{F,g,g}}(t) = \sum_{F_{I,g} \in S_{F,g}} F_{I,g}^2(t) / \sum_{F_{I,g} \in S_{F,g}} F_{I,g}(t)$ . Any element  $v_{J,I,g}(t)$  that violates the user specified parameter constraint interval  $[\Xi_J^{Lower}(t), \Xi_J^{Upper}(t)]$  will be reset to a

value within the interval by the rule

$$v_{J,I,g}(t) = \begin{cases} \Xi_J^{Lower}(t) + \zeta_{J,I,g}(t) \left( \Xi_J^{Upper}(t) - \Xi_J^{Lower}(t) \right), & \text{if } v_{J,I,g}(t) < \Xi_J^{Lower}(t), \\ \Xi_J^{Upper}(t) - \zeta_{J,I,g}(t) \left( \Xi_J^{Upper}(t) - \Xi_J^{Lower}(t) \right), & \text{if } v_{J,I,g}(t) > \Xi_J^{Upper}(t), \end{cases}$$

where  $\zeta_{J,I,g}(t)$  is randomly drawn from uniform distribution  $U[0, 1]$ .

*Crossover step for iteration stage g:* We construct vectors  $\mathbf{u}_{I,g}(t) = \{u_{J,I,g}(t)\}_{J=1}^{3n}$ ,  $I = 1, \dots, N_p$ ,

$$u_{J,I,g}(t) = \begin{cases} v_{J,I,g}(t), & \text{if } \xi_{J,I,g}(t) \leq CR_{I,g}, \\ \Xi_{J,I,g}(t), & \text{otherwise,} \end{cases}$$

where  $\xi_{J,I,g}(t)$  is randomly drawn from uniform distribution  $U[0, 1]$ , while  $CR_{I,g}(t)$  is generated by truncating a random draw from a Gaussian distribution  $\mathcal{N}(\mu_{CR,g}(t), 0.01)$  with respect to a floor of 0 and a ceiling of 1. Let  $\mu_{CR,0}(t) = 0.5$ , and let  $\mu_{CR,g+1}(t) = (1 - c)\mu_{CR,g}(t) + c\mu_{A,CR,g}(t)$ , where  $c$  is a user specified parameter.

*Selection step for iteration stage g:* We update  $\Xi_{I,g+1}(t)$  by the rule

$$\Xi_{I,g+1}(t) = \begin{cases} \mathbf{u}_{I,g}(t), & \text{if } \Psi(\mathbf{u}_{I,g}(t)) < \Psi(\Xi_{I,g}(t)), \\ \Xi_{I,g}(t), & \text{otherwise.} \end{cases}$$

The selection is successful if  $\Xi_{I,g+1}(t) \neq \Xi_{I,g}(t)$ .

*Algorithm termination:* The parameter vector  $\Xi_{best,G}(t)$  corresponding to the minimum of  $\Psi(\Xi_{I,G}(t))$  over  $I = 1, \dots, N_p$  is the best estimate of the parameters in the set  $\{\Xi_{I,G}(t), I = 1, \dots, N_p\}$ , and  $\Xi_{best,G}(t)$  is regarded as the approximate solution of (17) by minimization of objective function (19) over a given number of iteration steps  $G$ . In our experiments, we used  $G = 2,000$ ,  $c = 0.15$ ,  $p = 0.05$ , and  $N_p = 25 \times 3n$ . This was sufficient to achieves an acceptable level of precision.

Let  $\Xi_{best,G}(t) = \{\Xi_{best,G,j}(t)\}_{j=1}^{3n} = (\boldsymbol{\kappa}_{best,G}(t), \boldsymbol{\theta}_{best,G}(t), \boldsymbol{\sigma}_{best,G}(t))$ , where  $\boldsymbol{\kappa}_{best,G}(t) = \{\kappa_{best,G,j}(t)\}_{j=1}^n$ ,  $\boldsymbol{\theta}_{best,G}(t) = \{\theta_{best,G,j}(t)\}_{j=1}^n$ ,  $\boldsymbol{\sigma}_{best,G}(t) = \{\sigma_{best,G,j}(t)\}_{j=1}^n$ , are the term structures of CIR model parameters estimated from the corresponding zero coupon prices along the respective tenors  $\{T_{j,3} - t\}_{j=1}^n$ . The triplets  $(\kappa_{best,G,j}(t), \theta_{best,G,j}(t), \sigma_{best,G,j}(t))$  are regarded as the sets of estimated CIR model parameters for the short rate process that span the tenors  $T_{j,3} - t$ . The estimated parameters  $\Xi_{best,G}(t)$  are approximate solutions to (18). Uncertainty bound for estimated parameters may be constructed based on

$\{\kappa_{I,G,j}(t)\}_{I=1}^{N_p}$ ,  $\{\theta_{I,G,j}(t)\}_{I=1}^{N_p}$ , and  $\{\sigma_{I,G,j}(t)\}_{I=1}^{N_p}$  based on the strategy proposed in Hin and Dokuchaev (2015).

*Computation acceleration by parallel implementation:* From a typical collection of cross section US STRIPS zero coupon day-close prices, we can formulate approximately 40 triplets of nonlinear equations. Numerical solution of such a large system of equations is computationally intensive. To accelerate the numerical computation involved, we deployed 48 parallel cores on the National eResearch Collaboration Tools and Resources (NeCTAR) eResearch cloud infrastructure to carry out all the calculations reported in this paper. We used the parallel implementation of the Zhang-Sanderson algorithm in the form of DEoptim (Ardia et al., 2012) package for the R computing environment (R Core Team, 2012). To give an idea of the computation acceleration, we estimated the CIR model parameters based on day-close US STRIPS prices on January 2nd 2001. There were 168 parameters to be estimated. Using one processor, it took 3 hours and 10 minutes to carry out 2000 iterations. Using 48 parallel processors, it took only 46 minutes, leading to a 76% reduction in computation time.

### **3.3 Forecast of the short rate and its forward variance using implied CIR model parameters**

We construct two sets of predictors to forecast the future short rate. The first set of predictors, (20) to (23) of Section 3.3.1, is designed for the multi-curve framework, while the second set of predictors, (24) to (26) of Section 3.3.2, is designed for the single-curve framework.

#### **3.3.1 Forecast within the multi-curve framework**

We formulate the forecast strategy in the multi-curve framework by carrying out future short rate prediction based on multiple one-factor CIR models, each mapping the market participants' expectation of the future short rate at different forecast horizons from the triplets with comparable tenors to different sets of implied model parameters, assuming that each triplet is associated with a

different short rate process. The estimated CIR model parameters  $\{\kappa_{best,G,j}(t), \theta_{best,G,j}(t), \sigma_{best,G,j}(t)\}$  are used to model the short rate process associated with the triplet  $\{P(t, T_{j,k})\}_{k=1}^3$ . We then use the projection of the future short rate based on the corresponding tenor-specific short rate curve in order to forecast  $r(T_{j,3})$ , the realized short rate at time  $T_{j,3}$ .

We use (11) to estimate  $\gamma_{CIR}(t; t, T_{j,3})$  at the forecast horizon  $T_{j,3} - t$  as

$$\hat{\gamma}_{CIR}(t; t, T_{j,3}) = \theta_{best,G,j}(t) (T_{j,3} - t) + \frac{r(t) - \theta_{best,G,j}(t)}{\kappa_{best,G,j}(t)} (1 - e^{-\kappa_{best,G,j}(t) (T_{j,3} - t)}), \quad j = 1, \dots, n, \quad (20)$$

and use (13) to estimate  $\rho_{CIR}(t; T_{j,3}, T_{j+1,3})$  for the time intervals  $[T_{j+1,3}, T_{j,3}]$  as

$$r_j(T_{j,3}) = \hat{\rho}_{CIR}(t; T_{j,3}, T_{j+1,3}) = \frac{\hat{\gamma}_{CIR}(t; t, T_{j+1,3}) - \hat{\gamma}_{CIR}(t; t, T_{j,3})}{T_{j+1,3} - T_{j,3}}, \quad j = 1, \dots, n-1. \quad (21)$$

We use (4) to calculate the forward rates in the intervals  $[T_{j,3}, T_{j+1,3}]$

$$F(t; T_{j,3}, T_{j+1,3}) = \frac{\log P(t, T_{j,3}) - \log P(t, T_{j+1,3})}{T_{j+1,3} - T_{j,3}}, \quad j = 1, \dots, n-1. \quad (22)$$

as another predictor of the future short rate  $r(T_{j,3})$  at the same forecast horizons. Additionally, we suggest to use (21) and  $F(t; T_{j,3}, T_{j+1,3})$  to estimate the average variance of the integrated CIR process within the time interval  $[T_{j,3}, T_{j+1,3}]$  as

$$\begin{aligned} \hat{\Delta}_{CIR}(t; T_{j,3}, T_{j+1,3}) &= 2 \left( \hat{\rho}_{CIR}(t; T_{j,3}, T_{j+1,3}) - F(t; T_{j,3}, T_{j+1,3}) \right) \\ &\approx \frac{\hat{\delta}_{CIR}(t; T_{j,3}, T_{j+1,3})}{T_{j+1,3} - T_{j,3}}, \end{aligned} \quad (23)$$

where

$$\hat{\delta}_{CIR}(t; T_{j,3}, T_{j+1,3}) = \widehat{\text{Var}}_{CIR} \left[ \int_t^{T_{j+1,3}} r_{CIR}(u) du \middle| \mathcal{F}_t \right] - \widehat{\text{Var}}_{CIR} \left[ \int_t^{T_{j,3}} r_{CIR}(u) du \middle| \mathcal{F}_t \right],$$

and where  $\widehat{\text{Var}}_{CIR}(\cdot)$  denote the variance expressions calculated using all  $n$  triplets of the estimated CIR model parameters  $\{\theta_{best,G,j}(t), \kappa_{best,G,j}(t), \sigma_{best,G,j}(t)\}_{j=1}^n$ .

### 3.3.2 Forecast within the single-curve framework

We formulate an alternative forecast strategy based on the usual approach of using a single one-factor CIR models to model the dynamics of the entire interest rate term structure in the classical

single-curve framework. This strategy is different from that described in Section 3.3.1. Specifically, the estimated CIR model parameters  $\{\kappa_{best,G,1}(t), \theta_{best,G,1}(t), \sigma_{best,G,1}(t)\}$  implied from the triplet  $\{P(t, T_{1,k})\}_{k=1}^3$  with the shortest tenors in the set of cross section zero coupon bond prices are used to model the short rate process and to forecast the realized short rate at different forecast horizons that may span beyond  $T_{1,3} - t$  ahead.

In this context, we use (11) to estimate  $\gamma_{CIR}(t; t, T_{j,3})$  at the forecast horizon  $T_{j,3} - t$  as

$$\hat{\gamma}_{CIR,1}(t; t, T_{j,3}) = \theta_{best,G,1}(t) (T_{j,3} - t) + \frac{r(t) - \theta_{best,G,1}(t)}{\kappa_{best,G,1}(t)} (1 - e^{-\kappa_{best,G,1}(t)(T_{j,3} - t)}), \quad j = 1, \dots, n, \quad (24)$$

and use (13) to estimate  $\rho_{CIR}(t; T_{j,3}, T_{j+1,3})$  for the time intervals  $[T_{j+1,3}, T_{j,3}]$  as

$$\hat{\rho}_{CIR,1}(t; T_{j,3}, T_{j+1,3}) = \frac{\hat{\gamma}_{CIR,1}(t; t, T_{j+1,3}) - \hat{\gamma}_{CIR,1}(t; t, T_{j,3})}{T_{j+1,3} - T_{j,3}}, \quad j = 1, \dots, n - 1. \quad (25)$$

Here, we use (25) as a predictor of future short rate  $\{r(T_{j,3})\}_{j=1}^{n-1}$  at the forecast horizons  $\{T_{j,3} - t\}_{j=1}^{n-1}$  in the single-curve framework. Additionally, we use  $\hat{\rho}_{CIR,1}(t; T_{j,3}, T_{j+1,3})$  and  $F(t; T_{j,3}, T_{j+1,3})$  to estimate the average variance of the integrated CIR process within the time interval  $[T_{j,3}, T_{j+1,3}]$  as

$$\begin{aligned} \hat{\Delta}_{CIR,1}(t; T_{j,3}, T_{j+1,3}) &= 2 \left( \hat{\rho}_{CIR,1}(t; T_{j,3}, T_{j+1,3}) - F(t; T_{j,3}, T_{j+1,3}) \right) \\ &\approx \frac{\hat{\delta}_{CIR,1}(t; T_{j,3}, T_{j+1,3})}{T_{j+1,3} - T_{j,3}}, \end{aligned} \quad (26)$$

where

$$\hat{\delta}_{CIR,1}(t; T_{j,3}, T_{j+1,3}) = \widehat{\text{Var}}_{CIR,1} \left[ \int_t^{T_{j+1,3}} r_{CIR}(u) du \middle| \mathcal{F}_t \right] - \widehat{\text{Var}}_{CIR,1} \left[ \int_t^{T_{j,3}} r_{CIR}(u) du \middle| \mathcal{F}_t \right],$$

and where  $\widehat{\text{Var}}_{CIR,1}(\cdot)$  denote the variance expressions calculated using the first of the  $n$  triplets of the estimated CIR model parameters  $(\theta_{best,G,1}(t), \kappa_{best,G,1}(t), \sigma_{best,G,1}(t))$ . Note that the triplet of the CIR model parameters  $\{\theta_{best,G,1}(t), \kappa_{best,G,1}(t), \sigma_{best,G,1}(t)\}$  used in (24) to (26) are inferred from  $\{P(t, T_{1,k})\}_{k=1}^3$ , the prices of the three zero coupon bonds with the shortest tenor observed at time  $t$ .

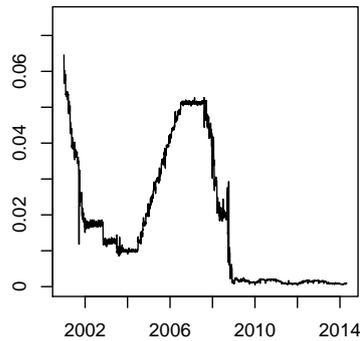


Figure 1: Time series of effective Federal Funds rate from January 2nd 2001 to April 28th 2014.

## 4 Empirical example

### 4.1 The data

We use the day close prices of the US STRIPS zero coupon bonds from January 2nd 2001 to April 28th 2014. They are retrieved from the Thomson Reuters Tick History (TRTH) supplied by the Securities Industry Research Centre of Asia-Pacific (SIRCA). The US STRIPS are zero coupon bonds created by stripping the Treasury notes and bonds. They can be reconstituted to form the original Treasury notes and bonds from the corresponding zero coupon bonds if so desired (see, e.g., Gregory and Livingston, 1992; Grinblatt and Longstaff, 2000). The market for the US STRIPS zero coupons is highly liquid and these zero coupons contain negligible credit risk. Following the rationale of Carr et al. (2011), we use the average of the day close bid and ask prices of each zero coupons as the day close price in our calculations.

The historical data of the US effective Federal Funds rate was downloaded from the Board of Governors of the Federal Reserve System. The quoted rates are converted to continuously compound convention for analysis. Figure 1 depict the time series of effective Federal Funds rate from January 2nd 2001 to April 28th 2014.

We use the US effective Federal Funds rate as the proxy for the short rate process to provide input

for  $r(t)$  in the CIR model, and we use the information extracted from the day close bid-ask average prices of the US STRIPS to forecast the effective Federal Funds rate at various forecast horizons. There are two reasons for our choice of this short rate proxy.

Firstly, the credit and liquidity risks of the effective Federal Funds rate are comparable to those of the US STRIPS, the former a weighted average of the uncollateralized overnight borrowing rate for a group of federal funds broker, the latter an entity derived from the uncollateralized US Treasuries.

Secondly, the effective Federal Funds rate and the US STRIPS are closely related entities in the same economy. Garfinkel and Thornton (1995) reported empirical evidence that the effective Federal Funds rate and the daily three-month T-bills rate are cointegrated, and they exhibit bidirectional Granger causality at various lags. Additionally, it has been long recognized that the effective Federal Funds rate and US Treasury bill (T-bill) rates tend to move together (see, e.g., Anderson, 1997; Hall et al., 1992; Rudebusch, 1995; Sarno and Thornton, 2003), and the T-bill rates have been used to predict the effective Federal Funds rates (see, e.g., Hardouvelis, 1988; Simon, 1990; Roberds et al., 1996; Guidolin and Timmermann, 2009). Since the US STRIPS are derived from US T-bills and T-bonds, we envisage that it may contain information for the prediction of the effective Federal Funds rate.

## 4.2 Term structure of implied CIR model parameters

Let  $t = t_i, i = 1, \dots, 3234$ , index the date of the cross section day close prices of the US STRIPS for each of the 3,234 trading days between January 2nd 2001 and April 28th 2014 denoted by  $\{P(t_i, T_{i,j,k})\}_{k=1,2,3}$ ,  $i = 1, \dots, 3234, j = 1, \dots, n_i$ , where  $T_{i,j,k}$  is the maturity date of the corresponding zero coupon bond. For each  $t_i$ , we construct a system of nonlinear equations using (17), and estimate the term structure of CIR model parameters using the numerical strategy described in Section 3.2. Let the term structure of CIR model parameters estimated from the zero coupon prices observed on the date indexed by  $t_i$  be  $\kappa_{best,G}(t_i) = (\kappa_{best,G,1}(t_i), \dots, \kappa_{best,G,n}(t_i))$ ,  $\theta_{best,G} = (\theta_{best,G,1}(t_i), \dots, \theta_{best,G,n}(t_i))$ ,  $\sigma_{best,G} = (\sigma_{best,G,1}(t_i), \dots, \sigma_{best,G,n}(t_i))$ , representing the term structure

of CIR model parameters estimated from the corresponding zero coupon prices for the respective tenors  $T_{i,j,3} - t_i, j = 1, \dots, n_i$ . The parameters  $(\kappa_{best,G,j}(t_i), \theta_{best,G,j}(t_i), \sigma_{best,G,j}(t_i))$  are regarded as the sets of estimated CIR model parameters for the single-factor CIR processes that span the tenors  $T_{i,j,3} - t_i$  where  $j = 1, \dots, n_i$  respectively.

Figure 2 depicts the term structure of average CIR model parameters implied from the US STRIPS day close prices from January 2nd 2001 to April 28th 2014 grouped into 100 intervals of tenors  $(\tau_\ell, \tau_{\ell+1}]$ , where  $\tau_\ell$  is the  $\ell$ -th percentile of tenors in the data for  $\ell = 0, \dots, 99$ . The non-constant relationship between the CIR model parameters and the tenor may be interpreted as an indication that if we choose to use one-factor CIR model to model the interest rate term structure, we cannot do so adequately in the single-curve framework because one set of CIR model parameters cannot capture the salient features of the entire term structure adequately. Instead, if we work in the multi-curve framework where each segment of the term structure that corresponds to a different tenor is assumed to be associated with a different short rate process, we then can model these short rate processes using multiple one-factor CIR models with different sets of CIR model parameters to capture the features of different segments of the yield curve.

The term structures of the estimated values of the mean reversion rate  $\hat{\kappa}$  and the long term interest rate  $\hat{\theta}$  are high at the short end of the tenor, rapidly decreases until maturities about one year, then increases until maturities up to 20 years, and finally decreases again. This “snake shape” profile of the  $\hat{\kappa}$  and  $\hat{\theta}$  term structures bears resemblance to the term structure of volatility of the London Interbank Offer Rates and swap rates depicted in Figure 6 of Piazzesi (2005) even though  $\hat{\kappa}$  and  $\hat{\theta}$  are inferred from the US STRIP prices and the effective Federal Funds rate, two economic entities that have credit and liquidity risk profiles that are different from the Libor rates. The shorter tenors correspond to higher levels of interest rate volatility (Piazzesi, 2005), implying that market participants’ tend to have more diverse views on the expectation of future short rate at shorter tenors. It is probably due to this reason that the sensitivities of  $\hat{\kappa}$  and  $\hat{\theta}$  with respect to the tenor are higher at shorter tenor. The bucket-wise standard deviation of  $\hat{\kappa}$  and  $\hat{\theta}$  are widest at around 1.5 to 2 years, corresponding to the “hump” of the volatility term structure depicted in Figure 6 of Piazzesi (2005), and in Figure 3 of Dai and Singleton (2000), reflect the need for a wider range of  $\hat{\kappa}$  and  $\hat{\theta}$  to

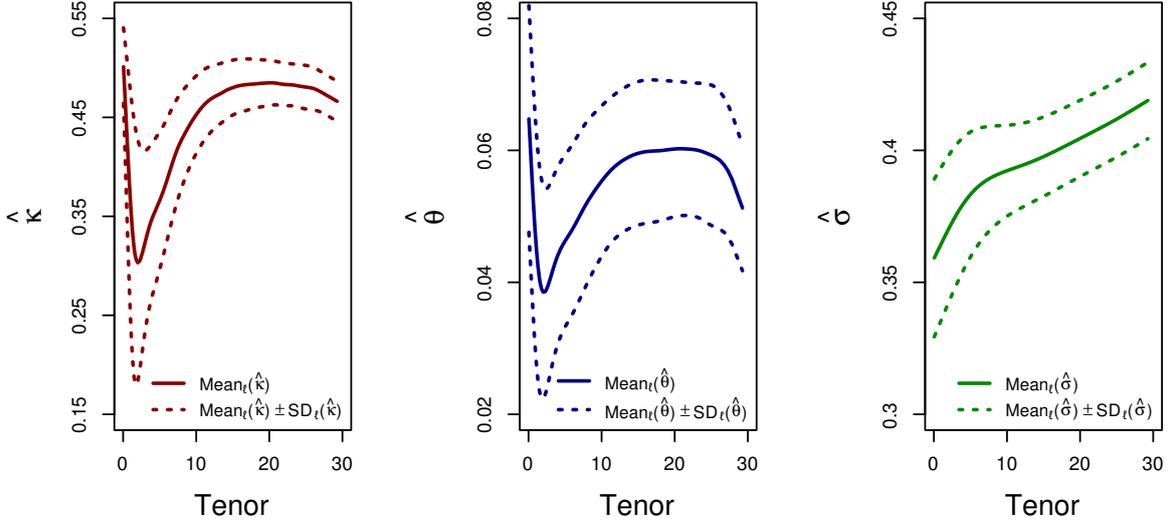


Figure 2: Average CIR model parameters implied from US STRIPS day close prices from January 2nd 2001 to April 28th 2014 grouped into 40 intervals of tenors  $(\tau_\ell, \tau_{\ell+1}]$ , where  $\tau_\ell$  is the  $q$ -th percentile of tenors in the data for  $q = 2.5\ell, \ell = 0, \dots, 39$ .

capture the more diverse views on the expectation of future short rate among market participants. Additionally, the wide bucket-wise standard deviation of  $\hat{\theta}$  may be interpreted as a reflection of the market participants' heterogeneous views of the long-term interest rate, while the increasing trend of  $\hat{\sigma}$  with respect to tenor imply the market participants' perception that interest rate risk increases as tenor increases.

### 4.3 Prediction of short rate and estimation of forward variance of integrated CIR process

#### 4.3.1 Forecast performance of short rate prediction in the multi-curve framework

Following the strategy detailed in Section 3.3.2, we use  $\hat{\rho}_{CIR}(t; T_{j,3}, T_{j+1,3})$  to predict  $r(T_{i,j,3}), j = 1, \dots, n_i - 1$ . We define the differences between the realized short rates,  $r(T_{i,j,3})$ , and the predicted short rates,  $\hat{\rho}_{CIR}(t_i; T_{i,j,3}, T_{i,j+1,3})$ , at forecast horizons  $T_{i,j,3} - t_i$  to be

$$\text{Error}(t_i; r(T_{i,j,3}), \hat{\rho}_{CIR}(t_i; T_{i,j,3}, T_{i,j+1,3})) = r(T_{j,3}) - \hat{\rho}_{CIR}(t_i; T_{i,j,3}, T_{i,j+1,3}), \quad j = 1, \dots, n_i - 1. \quad (27)$$

We define the differences between the realized short rates,  $r(T_{i,j,3})$ , and forward rates,  $F(t_i; T_{i,j,3}, T_{i,j+1,3})$ , at forecasting horizons of  $T_{i,j,3} - t_i$  to be

$$\text{Error}(t_i; r(T_{i,j,3}), F(t_i; T_{i,j,3}, T_{i,j+1,3})) = r(T_{i,j,3}) - F(t_i; T_{i,j,3}, T_{i,j+1,3}), \quad j = 1, \dots, n_i - 1. \quad (28)$$

We investigate the forecast performance of  $\hat{\rho}_{CIR}(t_i; T_{i,j,3}, T_{i,j+1,3})$  and  $F(t_i; T_{i,j,3}, T_{i,j+1,3})$  in predicting  $r(T_{i,j,3})$  at different forecast horizons. We choose 100 intervals of forecast horizons  $(\tau_\ell, \tau_{\ell+1}]$ , where  $\tau_\ell$  is the  $\ell$ -th percentiles of  $\{T_{i,j,3} - t_i, i = 1, \dots, 3234, j = 1, \dots, n_i - 1\}$  and  $\ell = 0, \dots, 99$ . For each  $(\tau_\ell, \tau_{\ell+1}]$ , we use (27) to compute the forecast horizons interval specific mean squared error (MSE)

$$\text{MSE}_\ell(\hat{\rho}_{CIR}) = \frac{1}{n_\ell} \sum_{i=1}^{3234} \sum_j^{n_i} \mathbf{1}_{\tau_\ell < T_{i,j,3} - t_i \leq \tau_{\ell+1}} \left( \text{Error}(t_i; r(T_{i,j,3}), \hat{\rho}_{CIR}(t_i; T_{i,j,3}, T_{i,j+1,3})) \right)^2 \quad (29)$$

and the mean absolute error (MAE)

$$\text{MAE}_\ell(\hat{\rho}_{CIR}) = \frac{1}{n_\ell} \sum_{i=1}^{3234} \sum_j^{n_i} \mathbf{1}_{\tau_\ell < T_{i,j,3} - t_i \leq \tau_{\ell+1}} \left| \text{Error}(t_i; r(T_{i,j,3}), \hat{\rho}_{CIR}(t_i; T_{i,j,3}, T_{i,j+1,3})) \right|, \quad (30)$$

where  $n_\ell = \sum_{i=1}^{3234} \sum_j^{n_i} \mathbf{1}_{\tau_\ell < T_{i,j,3} - t_i \leq \tau_{\ell+1}}$ , and  $\mathbf{1}$  is the indicator function. These error measures allow us to assess the prediction performance of  $\hat{\rho}_{CIR}(t; T_{j,3}, T_{j+1,3})$  across different forecast horizons. Additionally, we use (28) to calculate the forecast horizons interval specific MSE

$$\text{MSE}_\ell(F) = \frac{1}{n_\ell} \sum_{i=1}^{3234} \sum_j^{n_i} \mathbf{1}_{\tau_\ell < T_{i,j,3} - t_i \leq \tau_{\ell+1}} \left( \text{Error}(t_i; r(T_{i,j,3}), F(t_i; T_{i,j,3}, T_{i,j+1,3})) \right)^2, \quad (31)$$

and MAE

$$\text{MAE}_\ell(F) = \frac{1}{n_\ell} \sum_{i=1}^{3234} \sum_j^{n_i} \mathbf{1}_{\tau_\ell < T_{i,j,3} - t_i \leq \tau_{\ell+1}} \left| \text{Error}(t_i; r(T_{i,j,3}), F(t_i; T_{i,j,3}, T_{i,j+1,3})) \right|, \quad (32)$$

to assess to assess the prediction performance of  $F(t; T_{j,3}, T_{j+1,3})$  across different forecast horizons. Here,  $\min(n_\ell) = 449$ , and  $\max(n_\ell) = 584$ , and  $\sum_{\ell=0}^{99} n_\ell = 20,963$ .

Table 1 depicts the the forecast horizon specific performance of  $\hat{\rho}_{CIR}(t_i; T_{i,j,3}, T_{i,j+1,3})$ , the predictor for the multi-curve framework, and  $F(t_i; T_{i,j,3}, T_{i,j+1,3})$ , the forward rate, and  $\hat{\rho}_{CIR,1}(t_i; T_{i,j,3}, T_{i,j+1,3})$ , the predictor for the single-curve framework. We use the random walk as the benchmark to evaluate the short rate forecast accuracy (Duffee, 2002; Diebold and Li, 2006) where the short rate  $r(t)$  is

used as the forecast of the future short rate at the forecast horizon. The predictor for the multi-curve framework and the forward rate performs marginally better than the random walk benchmark for some forecast horizons especially for the shorter forecast horizons. However, the predictor for the single-curve framework perform less favourably against the random walk.

Table 1: Forecasting performance of effective Federal Funds Rate at different forecast horizons.  $MSE(\hat{\rho}_{CIR})$ ,  $MSE(\hat{\rho}_{CIR,1})$ ,  $MSE(F)$ , and  $MSE(RW)$  are the forecast horizon-specific MSE for the multi-curve and single-curve CIR predictors, forward rate, and the random walk benchmark respectively, while  $MAE_\ell(\hat{\rho}_{CIR})$ ,  $MAE_\ell(\hat{\rho}_{CIR,1})$ ,  $MAE_\ell(F)$ , and  $MAE_\ell(RW)$  are the forecast horizon-specific MAE for these predictors.

Forecast horizon (years)	$MSE_\ell(\hat{\rho}_{CIR})$ $\times 10^{-4}$	$MSE_\ell(\hat{\rho}_{CIR,1})$ $\times 10^{-4}$	$MSE_\ell(F)$ $\times 10^{-4}$	$MSE_\ell(RW)$ $\times 10^{-4}$	$MAE_\ell(\hat{\rho}_{CIR})$ $\times 10^{-3}$	$MAE_\ell(\hat{\rho}_{CIR,1})$ $\times 10^{-3}$	$MAE_\ell(F)$ $\times 10^{-3}$	$MAE_\ell(RW)$ $\times 10^{-3}$
0.4	0.42	1.72	0.16	0.19	4.91	11.83	2.97	2.79
0.8	1.95	5.61	3.05	1.98	10.58	20.05	10.52	10.64
1.2	0.60	5.28	1.18	2.28	6.35	18.68	9.51	11.70
1.6	5.96	12.03	8.08	6.45	18.51	29.10	18.80	20.49
1.8	8.14	15.91	10.19	4.36	23.00	36.31	22.05	14.18
2.2	8.29	20.11	6.44	3.76	24.55	41.36	22.25	11.81
2.4	13.06	16.72	9.93	9.44	30.99	34.45	28.03	25.56
2.8	9.04	14.57	6.82	9.05	23.15	29.80	21.91	25.20
3.2	14.52	23.11	11.31	12.37	32.54	41.84	29.18	30.99
3.4	11.30	17.77	7.92	7.89	28.04	35.94	24.83	23.04
3.8	16.27	26.80	9.85	11.15	36.21	45.76	26.73	28.95
4.2	17.75	27.76	10.51	10.10	38.47	48.00	28.82	27.44
4.6	26.51	33.09	18.00	12.24	48.08	52.87	37.90	30.63
5.0	20.86	30.20	12.93	6.23	42.33	49.34	31.50	20.75
5.4	32.09	37.86	21.59	12.97	55.14	57.15	44.63	31.20
5.8	26.16	33.57	16.77	6.99	48.13	53.58	37.66	22.36
6.2	37.07	41.46	24.26	14.41	60.30	60.75	48.78	33.70
6.6	35.04	39.80	22.52	6.74	57.77	61.17	45.49	19.48
7.1	39.65	47.44	26.76	10.70	62.38	68.12	50.98	28.16
7.6	44.43	43.51	29.17	8.03	65.81	62.55	53.32	23.99
8.1	46.80	48.70	32.86	8.63	67.70	68.32	56.41	25.32
8.5	48.79	47.90	33.51	5.28	68.61	66.61	56.88	19.59
9.0	56.31	41.24	37.57	7.77	74.21	60.77	60.54	23.19
9.6	56.50	47.90	39.94	4.93	74.30	67.26	62.59	18.89

We depict, in Panels A and B of Figure 3,  $MSE_\ell(\hat{\rho}_{CIR})$ ,  $MSE_\ell(F)$ ,  $MAE_\ell(\hat{\rho}_{CIR})$  and  $MAE_\ell(F)$  for the prediction of  $r(T_{i,j,3})$ . Their profiles are similar, being considerably smaller at shorter forecast horizons. Both  $\hat{\rho}_{CIR}(t_i; T_{i,j,3}, T_{i,j+1,3})$  and  $F(t_i; T_{i,j,3}, T_{i,j+1,3})$  appear to be reasonably good predictors of future short rate up to a forecast horizon of 6 months with an average absolute prediction error of less than 0.005 (Panel B, Figure 3). The effect of term premium on forecasting ability of both  $\hat{\rho}_{CIR}(t_i; T_{i,j,3}, T_{i,j+1,3})$  and  $F(t_i; T_{i,j,3}, T_{i,j+1,3})$  becomes noticeable beyond a forecast horizon of 6 months, leading to progressive deterioration of the forecasting performance. This is consistent with the findings of Fama (1976) and Longstaff (2000). Additionally, the corresponding  $\ell$ -interval

specific MSE and MAE for the random walk are depicted in Figure 3 for comparison.

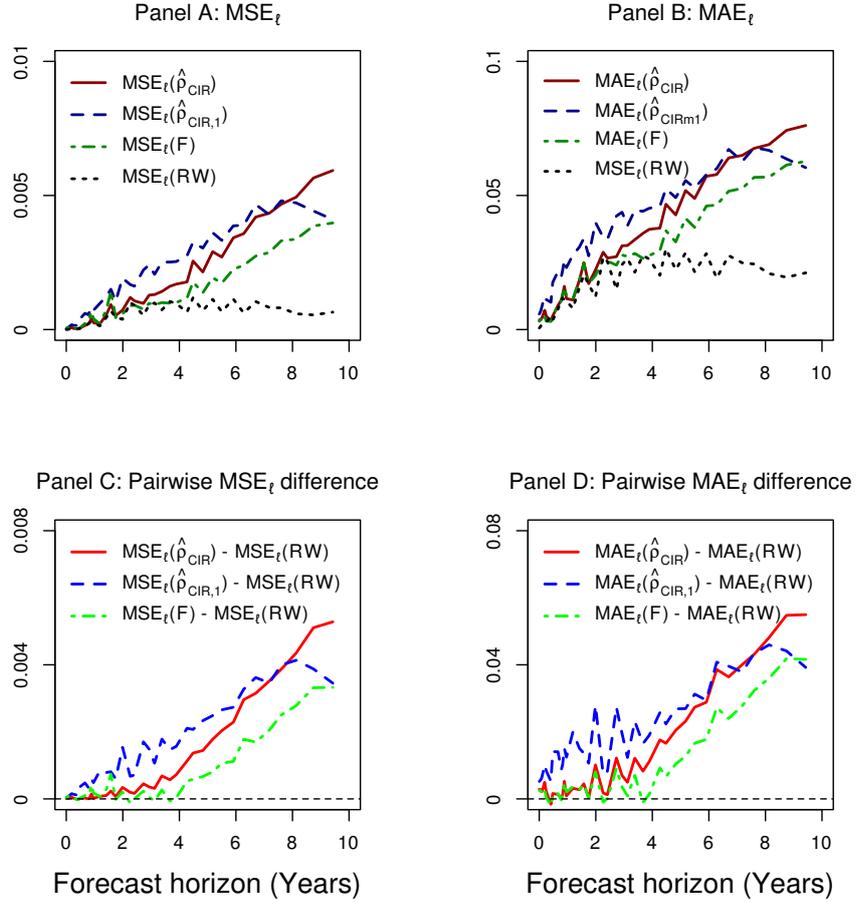


Figure 3: Prediction performance of  $\hat{\rho}_{CIR}(t_i; T_{i,j,3}, T_{i,j+1,3})$ ,  $\hat{\rho}_{CIR,1}(t_i; T_{i,j,3}, T_{i,j+1,3})$ ,  $F(t_i; T_{i,j,3}, T_{i,j+1,3})$ , and random walk (RW) at different forecast horizons.

The forecast horizon interval specific MSE and MAE for  $\hat{\rho}_{CIR}(t; T_{j,3}, T_{j+1,3})$  and  $F(t; T_{j,3}, T_{j+1,3})$  are similar for forecast horizons of up to 2 years, beyond which those of  $\hat{\rho}_{CIR}(t_i; T_{i,j,3}, T_{i,j+1,3})$  become noticeably larger than those of  $F(t_i; T_{i,j,3}, T_{i,j+1,3})$  (Panels C and D, Figure 3). This phenomenon can be partly explained by the variance of the integrated CIR process in the interval  $[T_{i,j,3}, T_{i,j+1,3}]$ .

Recall from (23) that

$$\begin{aligned}
\frac{1}{2}\hat{\Delta}_{CIR}(t_i; T_{i,j,3}, T_{i,j+1,3}) &= \hat{\rho}_{CIR}(t_i; T_{i,j,3}, T_{i,j+1,3}) - F(t_i; T_{i,j,3}, T_{i,j+1,3}) \\
&= \left(\hat{\rho}_{CIR}(t_i; T_{i,j,3}, T_{i,j+1,3}) - r(T_{i,j,3})\right) - \left(F(t_i; T_{i,j,3}, T_{i,j+1,3}) - r(T_{i,j,3})\right) \\
&= \text{Error}\left(t_i; r(T_{i,j,3}), \hat{\rho}_{CIR}(t_i; T_{i,j,3}, T_{i,j+1,3})\right) - \text{Error}\left(t_i; r(T_{i,j,3}), F(t_i; T_{i,j,3}, T_{i,j+1,3})\right) \\
&\approx \frac{\hat{\delta}_{CIR}(t; T_{i,j,3}, T_{i,j+1,3})}{2(T_{i,j+1,3} - T_{i,j,3})}, \tag{33}
\end{aligned}$$

it follows that

$$\begin{aligned}
&\text{MSE}_\ell(\hat{\rho}_{CIR}) - \text{MSE}_\ell(F) \\
&\approx \frac{1}{n_\ell} \sum_{i=1}^{3234} \sum_j^{n_i} \mathbf{1}_{\tau_\ell < T_{i,j,3} - t_i \leq \tau_{\ell+1}} \text{Error}\left(t_i; r(T_{i,j,3}), F(t_i; T_{i,j,3}, T_{i,j+1,3})\right) \frac{\hat{\delta}(t; T_{i,j,3}, T_{i,j+1,3})}{T_{i,j+1,3} - T_{i,j,3}} \\
&+ \frac{1}{n_\ell} \sum_{i=1}^{3234} \sum_j^{n_i} \mathbf{1}_{\tau_\ell < T_{i,j,3} - t_i \leq \tau_{\ell+1}} \left(\frac{\hat{\delta}(t; T_{i,j,3}, T_{i,j+1,3})}{2(T_{i,j+1,3} - T_{i,j,3})}\right)^2.
\end{aligned}$$

The progressively increasing trend of  $\text{MSE}_\ell(\hat{\rho}_{CIR}) - \text{MSE}_\ell(F)$  along the forecast horizons implies that  $\hat{\delta}(t; T_{i,j,3}, T_{i,j+1,3})$  is larger at longer forecast horizons.

For each  $(\tau_\ell, \tau_{\ell+1}]$ ,  $\ell = 0, \dots, 99$ , we calculate the interval specific Theil's inequality coefficient  $U$  to assess the forecast performance of the two predictors of future short rate considered. The Theil's inequality coefficient  $U$  is bounded between zero in the case of perfect forecasting and one in the case of very poor forecasting performance. The nearer the Theil's  $U$  statistic is to zero, the better the forecasting performance (Theil, 1966; Bliemel, 1973). We calculate the interval specific Theil's inequality coefficient  $U$  for  $\hat{\rho}_{CIR}(t_i; T_{i,j,3}, T_{i,j+1,3})$  as

$$U_\ell(\hat{\rho}_{CIR}) = \frac{[\text{MSE}_\ell(\hat{\rho}_{CIR})]^{1/2}}{[\text{MSS}_\ell(\hat{\rho}_{CIR})]^{1/2} + [\text{MSS}_\ell(r(T_{i,j,3}))]^{1/2}}, \tag{34}$$

where  $\text{MSE}_\ell(\hat{\rho}_{CIR})$  is defined in (29),

$$\text{MSS}_\ell(\hat{\rho}_{CIR}) = \frac{1}{n_\ell} \sum_{i=1}^{3234} \sum_j^{n_i} \mathbf{1}_{\tau_\ell < T_{i,j,3} - t_i \leq \tau_{\ell+1}} \left(\hat{\rho}_{CIR}(t_i; T_{i,j,3}, T_{i,j+1,3})\right)^2,$$

and

$$\text{MSS}_\ell(r(T_{i,j,3})) = \frac{1}{n_\ell} \sum_{i=1}^{3234} \sum_j^{n_i} \mathbf{1}_{\tau_\ell < T_{i,j,3} - t_i \leq \tau_{\ell+1}} \left(r(T_{i,j,3})\right)^2.$$

We also calculate the interval specific Theil's inequality coefficient  $U$  for  $F(t_i; T_{i,j,3}, T_{i,j+1,3})$  as

$$U_\ell(F) = \frac{[\text{MSE}_\ell(F)]^{1/2}}{[\text{MSS}_\ell(F)]^{1/2} + [\text{MSS}_\ell(r(T_{i,j,3}))]^{1/2}}, \quad (35)$$

where  $\text{MSE}_\ell(F)$  is defined in (31), and

$$\text{MSS}_\ell(F) = \frac{1}{n_\ell} \sum_{i=1}^{3234} \sum_j^{n_i} \mathbf{1}_{\tau_\ell < T_{i,j,3} - t_i \leq \tau_{\ell+1}} \left( F(t_i; T_{i,j,3}, T_{i,j+1,3}) \right)^2.$$

Figure 4 depict  $U_\ell(\hat{\rho}_{CIR})$  and  $U_\ell(F)$  with respect to different lead time period intervals  $(\tau_\ell, \tau_{\ell+1}]$ ,  $\ell = 0, \dots, 99$ . Both of them show an increasing trend with respect to the forecast horizon, indicating that the forecasting performance is better at shorter forecast horizon. In particular, the values of  $U_\ell(\hat{\rho}_{CIR})$  and  $U_\ell(F)$ , are approximately 0.2 or less when the forecast horizons are less than 6 months, implying that both  $\hat{\rho}_{CIR}(t_i; T_{i,j,3}, T_{i,j+1,3})$  and  $F(t_i; T_{i,j,3}, T_{i,j+1,3})$  are good predictors of the future effective Federal Funds rate up to 6 months ahead. This is consistent with the trend depicted in Figure 3.

### 4.3.2 Forecast performance of short rate prediction in the single-curve framework

In order to investigate whether the prices of zero coupon bonds can be used to predict future short rate at forecasting horizons beyond their tenor, we use  $\hat{\rho}_{CIR,1}(t; T_{j,3}, T_{j+1,3})$  to predict  $r(T_{i,j,3})$ ,  $j = 1, \dots, n_i - 1$ , the future short rate at different forecasting horizons, following the strategy detailed in Section 3.3.2. The evaluation of  $\hat{\rho}_{CIR,1}(t; T_{j,3}, T_{j+1,3})$  is based on  $(\kappa_{best,G,1}(t_i), \theta_{best,G,1}(t_i), \sigma_{best,G,1}(t_i))$  inferred from  $\{P(t_i, T_{i,1,k})\}_{k=1,2,3}$ .

We define the differences between the realized short rates,  $r(T_{i,j,3})$ , and the predicted short rates  $\hat{\rho}_{CIR,1}(t_i; T_{i,j,3}, T_{i,j+1,3})$  at forecast horizons  $T_{i,j,3} - t_i$  to be

$$\text{Error}(t_i; r(T_{i,j,3}), \hat{\rho}_{CIR,1}(t_i; T_{i,j,3}, T_{i,j+1,3})) = r(T_{i,j,3}) - \hat{\rho}_{CIR,1}(t_i; T_{i,j,3}, T_{i,j+1,3}), \quad j = 1, \dots, n_i - 1. \quad (36)$$

We investigate the forecast performance of  $\hat{\rho}_{CIR,1}(t_i; T_{i,j,3}, T_{i,j+1,3})$  in predicting  $r(T_{i,j,3})$  at different forecast horizons. We choose 100 intervals of forecast horizons  $(\tau_\ell, \tau_{\ell+1}]$ , where  $\tau_\ell$  is the  $\ell$ -th percentiles of  $\{T_{i,j,3} - t_i, i = 1, \dots, 3234, j = 1, \dots, n_i - 1\}$  where  $\ell = 0, \dots, 99$ . For each  $(\tau_\ell, \tau_{\ell+1}]$ ,

we use (27) to compute the forecast horizons interval specific MSE

$$\text{MSE}_\ell(\hat{\rho}_{CIR,1}) = \frac{1}{n_\ell} \sum_{i=1}^{3234} \sum_j^{n_i} \mathbf{1}_{\tau_\ell < T_{i,j,3} - t_i \leq \tau_{\ell+1}} \left( \text{Error} \left( t_i; r(T_{i,j,3}), \hat{\rho}_{CIR,1}(t_i; T_{i,j,3}, T_{i,j+1,3}) \right) \right)^2 \quad (37)$$

and MAE

$$\text{MAE}_\ell(\hat{\rho}_{CIR,1}) = \frac{1}{n_\ell} \sum_{i=1}^{3234} \sum_j^{n_i} \mathbf{1}_{\tau_\ell < T_{i,j,3} - t_i \leq \tau_{\ell+1}} \left| \text{Error} \left( t_i; r(T_{i,j,3}), \hat{\rho}_{CIR,1}(t_i; T_{i,j,3}, T_{i,j+1,3}) \right) \right|, \quad (38)$$

where  $n_\ell = \sum_{i=1}^{3234} \sum_j^{n_i} \mathbf{1}_{\tau_\ell < T_{i,j,3} - t_i \leq \tau_{\ell+1}}$ , and  $\mathbf{1}$  is the indicator function. These error measures allow us to assess the prediction performance of  $\hat{\rho}_{CIR,1}(t; T_{j,3}, T_{j+1,3})$  across different forecast horizons. We depict, in Panels A and B of Figure 3,  $\text{MSE}_\ell(\hat{\rho}_{CIR,1})$ , and  $\text{MAE}_\ell(\hat{\rho}_{CIR,1})$  for the prediction of  $r(T_{i,j,3})$ . It appears that the forecasting performance of  $\hat{\rho}_{CIR,1}(t_i; T_{i,j,3}, T_{i,j+1,3})$  is inferior to  $\hat{\rho}_{CIR}(t_i; T_{i,j,3}, T_{i,j+1,3})$ ,  $F(t_i; T_{i,j,3}, T_{i,j+1,3})$ , and the random walk at all forecast horizons.

We calculate the interval specific Theil's inequality coefficient U for  $\hat{\rho}_{CIR,1}(t_i; T_{i,j,3}, T_{i,j+1,3})$  as

$$U_\ell(\hat{\rho}_{CIR,1}) = \frac{[\text{MSE}_\ell(\hat{\rho}_{CIR,1})]^{1/2}}{[\text{MSS}_\ell(\hat{\rho}_{CIR,1})]^{1/2} + [\text{MSS}_\ell(r(T_{i,j,3}))]^{1/2}}, \quad (39)$$

where  $\text{MSE}_\ell(\hat{\rho}_{CIR,1})$  is defined in (37), and

$$\text{MSS}_\ell(\hat{\rho}_{CIR,1}) = \frac{1}{n_\ell} \sum_{i=1}^{3234} \sum_j^{n_i} \mathbf{1}_{\tau_\ell < T_{i,j,3} - t_i \leq \tau_{\ell+1}} \left( \hat{\rho}_{CIR,1}(t_i; T_{i,j,3}, T_{i,j+1,3}) \right)^2.$$

Figure 4 depicts  $U_\ell(\hat{\rho}_{CIR,1})$  for different lead time period intervals  $(\tau_\ell, \tau_{\ell+1}]$ ,  $\ell = 0, \dots, 99$ . The values of  $U_\ell(\hat{\rho}_{CIR,1})$  are higher than  $U_\ell(\hat{\rho}_{CIR})$  and  $U_\ell(F)$ , implying that the forecasting performance of  $U_\ell(\hat{\rho}_{CIR,1})$  is inferior to  $\hat{\rho}_{CIR,1}(t_i; T_{i,j,3}, T_{i,j+1,3})$  and  $F(t_i; T_{i,j,3}, T_{i,j+1,3})$  for different forecasting horizons.

### 4.3.3 The performance of the short rate predictors

For our dataset, forecast of the short rate based on the cross section data, i.e., under the risk-neutral measure, using our proposed algorithm in the multi-curve framework appears to be reasonably accurate up to a forecast horizon of about 6 months. We interpret the deterioration of the forecast

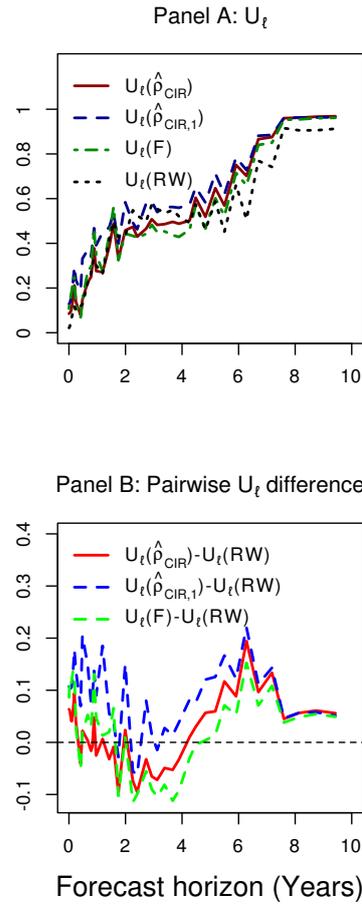


Figure 4: Theil's inequality coefficient  $U$  at different forecast horizons.

performance beyond 6 months as the result of, among other factors, the noticeable effect of the term premium associated with the short rate at forecast horizons beyond 6 months. It affirms the current practice of forecasting interest rate with respect to the physical measure as it is necessary to use the time series data of the yield curve to disentangle the term premium from the expected short rate at some forecast horizon. In this paper, we consider forecast under the risk-neutral measure instead because we intend to elucidate the forecast horizon beyond which the term premium effect may lead to noticeable deterioration of the short rate prediction performance if it were not accounted for.

It is clear from Figure 3 that our proposed predictor for the multi-curve framework performs better at all forecast horizons than the predictor based on the classical one-factor CIR model in the single-

curve framework. It is well known that the one-factor CIR model is incapable of generating a single short rate curve that is sufficiently flexible to capture the salient features of the entire yield curve spanning all available tenors, and hence it performs poorly at short rate forecasting. Our empirical results show better forecasting performance for our proposed approach of using tenor-specific one-factor CIR processes in the multi-curve framework to capture the salient features of different segments of the yield curve, and use this information to forecast short rate.

Term premium aside, other factors may affect the forecast performance reported above. The assessment of the forecast performance of  $\hat{\rho}_{CIR}(t; T_{j,3}, T_{j+1,3})$  in predicting the future short rate  $r(T_{i,j,3})$  in the present context is a joint test of whether  $\rho_{CIR}(t; T_{j,3}, T_{j+1,3})$  defined in (2) is a good predictor of the future short rate, and whether  $\hat{\rho}_{CIR}(t; T_{j,3}, T_{j+1,3})$  defined in (21) is a good estimator of  $\rho_{CIR}(t; T_{j,3}, T_{j+1,3})$ . The former is affected by the model specification risk, the latter by the model estimation risk.

The model specification risk arise due to the fact that  $\rho_{CIR}(t; T_{j,3}, T_{j+1,3})$  is a model-based predictor. While the extant literature suggests that the drift and diffusion coefficients of the short rate process may not be linear (see, e.g., Aït-Sahalia, 1996; Stanton, 1997; Jones, 2003; Chapman and Pearson, 2000), the one-factor CIR process is nonetheless a relatively good approximation if the short rate process being modelled lies within the range between 0 and 0.09 (Aït-Sahalia, 1996). The time series of effective Federal Funds rate, our short rate proxy, for the period we consider lies within this boundary (Figure 1). While this lends support to our choice of using CIR process to model this short rate dynamics, the CIR model is only an approximation.

The model estimation risk arise due to the fact that  $\hat{\rho}_{CIR}(t; T_{j,3}, T_{j+1,3})$  is evaluated using the estimated CIR model parameters inferred from the data as approximate solutions to a system of non-linear equations (17), leading to some uncertainty in the point estimation of  $\rho_{CIR,1}(t; T_{j,3}, T_{j+1,3})$ . In general, the model estimation risk is an issue faced by all parametric models constructed to model the short rate dynamics due to the necessity for numerical estimation of the model parameters.

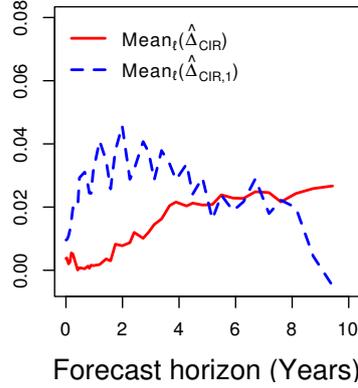


Figure 5: The average variance of the integrated CIR process at different forecast horizons.

#### 4.3.4 Estimation of forward variance of integrated short rate

For the multiple one-factor CIR models in the multi-curve framework described in Section 3.3.1, we use (33) to compute

$$\text{Mean}_\ell(\hat{\Delta}_{CIR}) = \frac{1}{n_\ell} \sum_{i=1}^{3234} \sum_j^{n_i} \mathbf{1}_{\tau_\ell < T_{i,j,3} - t_i \leq \tau_{\ell+1}} \hat{\Delta}_{CIR}(t_i; T_{i,j,3}, T_{i,j+1,3}), \quad (40)$$

to estimate the average forward variance of the integrated CIR process based on all available triplets of inferred CIR model parameters. For the single one-factor CIR model in the single-curve framework described in Section 3.3.2, we use

$$\frac{1}{2} \hat{\Delta}_{CIR,1}(t_i; T_{i,j,3}, T_{i,j+1,3}) \approx \frac{\hat{\delta}_{CIR,1}(t; T_{i,j,3}, T_{i,j+1,3})}{2(T_{i,j+1,3} - T_{i,j,3})} \quad (41)$$

instead to compute

$$\text{Mean}_\ell(\hat{\Delta}_{CIR,1}) = \frac{1}{n_\ell} \sum_{i=1}^{3234} \sum_j^{n_i} \mathbf{1}_{\tau_\ell < T_{i,j,3} - t_i \leq \tau_{\ell+1}} \hat{\Delta}_{CIR,1}(t_i; T_{i,j,3}, T_{i,j+1,3}), \quad (42)$$

instead to estimate the average forward variance of the integrated CIR process (Figure 5). It appears that  $\text{Mean}_\ell(\hat{\Delta}_{CIR})$  produces a more reasonable profile of average forward variance plot than  $\text{Mean}_\ell(\hat{\Delta}_{CIR,1})$  as one would expect that the further out the forecast horizon, the larger the average forward variance.

It is well known that the interest rate dynamics consists of both diffusion and jump components (see, e.g., Das, 2002; Piazzesi, 2005; Wright and Zhou, 2009). In this paper, we have constructed the forward variance of integrated CIR short rate process  $\hat{\Delta}_{CIR}(t_i; T_{i,j,3}, T_{i,j+1,3})$  as a model-based estimate of the expected future realized variance of the short rate based on the CIR model in the multi-curve framework. Admittedly, it is not capable of capturing potential jumps in the effective Federal Funds rate (Piazzesi, 2005). That said, our intention is to outline the strategy to extract information on the market participants' expectation of the variance of the integrated future short rate volatility in a multi-curve framework. We leave the issue of constructing a prediction model based on a more flexible stochastic process to more accurately capture the realized volatility for future research.

While the sum of squared log return of asset price is a widely used nonparametric estimator of the realized variance, i.e., the square of the realized volatility, in the context of equity price dynamics (see, e.g., Merton, 1980; Ait-Sahalia et al., 2003; Andersen et al., 2007; McAleer and Medeiros, 2008b), Andersen and Benzoni (2010) proposed an estimator for the annualized realized volatility of the bond yield over the time interval  $[t, t + h]$ , given a sampling frequency  $h/n$ , by

$$RV_{\eta(t,T)}(n, h) = \frac{1}{n} \sum_{i=1}^n \left( -\frac{1}{T - t - \frac{ih}{n}} \log P\left(t + \frac{ih}{n}, T\right) + \frac{1}{T - t - \frac{(i-1)h}{n}} \log P\left(t + \frac{(i-1)h}{n}, T\right) \right)^2,$$

where  $P(t, T)$  and  $P(t + h, T)$  are the prices of a zero coupon bond observed at times  $t$  and  $t + h$  respectively where  $t < t + h$ . Andersen and Benzoni (2010) also derived estimators of the realized volatility of the short rate for the affine term structure model, of which the CIR model is a special case, and for the extended affine term structure model that take into consideration random jumps of the state variables.

The implied volatility may be viewed as the information on the market participants' aggregate view of future realized volatility that is usually represented in the form of model parameter estimated from some observable financial state variables based on some asset pricing model. The definition of implied volatility depends on the modelling framework. If a one factor log-normal stochastic process is used to model the dynamics of the log return of an asset price, the coefficient of the diffusion term of this process can be regarded as the implied volatility; it projects the quadratic variation

of the process as the forecast of the realized volatility. If a one factor affine jump diffusion process is used instead, the forecast of the realized volatility consists of two components; the quadratic variation which is related to the coefficient of the diffusion term, and the sum of squares of the magnitude of jumps which is related to the jump intensity and the random jump size distribution. In this case, the coefficient of the diffusion component of the affine jump diffusion process can no longer be regarded as the implied volatility. Instead, the notion of implied volatility is embedded in the affine jump diffusion model parameters, i.e., the coefficient of the diffusion component, the jump intensity, and the distribution characteristics of the random jump size. In the context of interest rate modelling, the coefficient of the diffusion term of a Gaussian short rate process may be regarded as the implied volatility as it is the only model parameter that contributes towards the quadratic variation (see, e.g., Steeley, 1997). In contrast, the notion of implied volatility for the CIR process is embedded within the three CIR model parameters as all of them contribute towards the volatility of the CIR process; the coefficient of the diffusion component, i.e.,  $\theta$  in (9), cannot be regarded as the implied volatility.

Emerging evidence from recent empirical studies (see, e.g., Duffee, 2011; Collin-Dufresne et al., 2009; Andersen and Benzoni, 2010) reveal that the cross-sectional interest rate term structure does not contain information on the short rate volatility. In this light, one should extract this information from the time series of the interest rates in order to forecast the realized volatility of the short rate. One may adopt time series modelling strategies such as the extended EGARCH-type model to estimate the implied volatility from this class of highly parametric time series model in order to forecast the realized volatility of interest rates (see, e.g., Andersen and Benzoni, 2010). Alternatively, one may use the Heterogeneous Autoregressive model of Realized Volatility (HAR-RV) proposed by Corsi (2009) to forecast the realized volatility of yield by linear projection based on the historical volatilities estimated at different lags. Empirical studies revealed that the out-of-sample forecast performance of the HAR-RV model is superior to the lag-1 and the lag-3 Autoregressive models, and the Autoregressive Fractional Integrated Moving Average (ARFIMA) model (Corsi, 2009), while its forecast performance is comparable to that of the extended EGARCH-type model (Andersen and Benzoni, 2010). The HAR-RV model has been extended in various directions in the

context of realized volatility forecast; for example, McAleer and Medeiros (2008a) extended the HAR-RV model to capture nonlinearities and long-range dependence in the time series dynamics via a flexible multiple regime smooth transition model, while Andersen et al. (2007) modified it to forecast the realized volatility of affine jump-diffusion stochastic processes.

## **5 Conclusion**

In this paper, we have proposed a strategy to extract the information on the market participants' expectation of the future short rate from the zero coupon bond prices. Instead of working within the classical single-curve framework and attempt to model the entire yield curve with one short rate process, we have approached this problem from a multi-curve perspective where multiple short rate processes have been constructed to pin down the salient features of the yield curve at different tenors. For our data, the proposed approach appears to deliver reasonably accurate forecast within a six-month forecast horizon.

The practice of multi-curve framework for yield curve modelling (see, e.g., Mercurio, 2009; 2010; Madan and Schoutens, 2012) is a important recent departure from the classical yield curve modelling framework, triggered by the mismatch between the classical theory of yield curve modelling and the empirical phenomenon observed in the market. Given this additional flexibility in the interest rate term structure modelling, we envisage that it will be a fertile area of research in the development of more sophisticated theory and strategies to capture the information from the market data. We leave this for future research.

## **Acknowledgements**

We acknowledge provision of ICT support and computing resources by Curtin IT Services. Curtin Information Technology Services (CITS) provides Information and Communication Technology systems and services in support of Curtin's teaching, learning, research and administrative activi-

ties. We acknowledge use of computing resources from the NeCTAR Research Cloud. NeCTAR is an Australian Government project conducted as part of the Super Science initiative and financed by the Education Investment Fund.

This work was supported by ARC grant of Australia DP120100928.

## References

- Ahn, C. and Thompson, H. (1988). Jump-Diffusion Processes and the Term Structure of Interest Rates, *The Journal of Finance* **43**(1): 155–174.
- Aït-Sahalia, Y. (1996). Nonparametric Pricing of Interest Rate Derivative Securities, *Econometrica* **64**(3): 527–560.
- Aït-Sahalia, Y., Mykland, P. and Zhang, L. (2003). How Often to Sample a Continuous-Time Process in the Presence of Market Microstructure Noise, *Review of Financial Studies* **18**: 351–416.
- Andersen, T. and Benzoni, L. (2010). Do Bonds Span Volatility Risk in the U.S. Treasury Market? A Specification Test for Affine Term Structure Models, *The Journal of Finance* **65**(2): 603–653.
- Andersen, T., Bollerslev, T. and Diebold, F. (2007). Roughing it up: Including Jump Components in the Measurement, Modeling, and Forecasting of Return Volatility, *The Review of Economics and Statistics* **89**(4): 701–720.
- Anderson, H. (1997). Transaction costs and no-linear adjustment towards equilibrium in the US Treasury Bill market, *Oxford Bulletin of Economics and Statistics* **59**(4): 465–484.
- Ardia, D., Mullen, K., Peterson, B. and Ulrich, J. (2012). *DEoptim: Differential Evolution in 'R'*. version 2.2-2.
- Bams, D. and Schotman, P. (2003). Direct Estimation of the Risk Neutral Factor Dynamics of Gaussian Term Structure Models, *Journal of Econometrics* **117**(1): 179–206.

- Baviera, R. and Cassaro, A. (2015). A Note on Dual-Curve Construction: Mr. Crab's Bootstrap, *Applied Mathematical Finance* **22**(2): 105–132.
- Bliemel, F. (1973). Theil's Forecast Accuracy Coefficient: A Clarification, *Journal of Marketing Research* **10**(4): 444–446.
- Brown, S. and Dybvig, P. (1986). The Empirical Implications of the Cox, Ingersoll, Ross Theory of the Term Structure of Interest Rates, *The Journal of Finance* **51**(3): 617–630.
- Carr, P., Madan, D. and Vicente Alvarez, J. (2011). Markets, Profits, Capital, Leverage and Return, *The Journal of Risk* **14**(1): 95–122.
- Chapman, D., J.B.Jr., L. and Pearson, N. (1999). Using Proxies for the Short Rate: When are Three Months Like an Instant, *The Review of Financial Studies* **12**(4): 763–806.
- Chapman, D. and Pearson, N. (2000). Is the Short Rate Drift Actually Nonlinear?, *The Journal of Finance* **55**(1): 355–388.
- Chapman, D. and Pearson, N. (2001). Recent Advances in Estimating Term-Structure Models, *Financial Analyst Journal* **57**(4): 77–95.
- Cochrane, J. and Piazzesi, M. (2005). Bond Risk Premia, *The American Economic Review* **95**(1): 138–160.
- Collin-Dufresne, P., Goldstein, R. and Jones, C. (2008). Identification of Maximal Affine Term Structure Models, *The Journal of Finance* **63**(2): 743–795.
- Collin-Dufresne, P., Goldstein, R. and Jones, C. (2009). Can Interest Rate Volatility be extracted from the Cross Section of Bond Yields?, *Journal of Financial Economics* **94**(1): 47–66.
- Corsi, F. (2009). A Simple Approximate Long-Memory Model of Realized Volatility, *Journal of Financial Econometrics* **7**(2): 174–196.
- Cox, J., Ingersoll, J. and Ross, S. (1985). A theory of the term structure of interest rates, *Econometrica* **53**(2): 385–407.

- Culberston, J. (1957). The Term Structure of Interest Rates, *Quarterly Journal of Economics* **71**(4): 485–517.
- Dai, Q. and Singleton, K. (2000). Specification Analysis of Affine Term Structure Models, *The Journal of Finance* **55**(5): 1943–1978.
- Dai, Q. and Singleton, K. (2002). Expectation puzzles, time-varying risk premia, and affine models of the term structure, *Journal of Financial Economics* **63**(3): 415–441.
- Das, S. (2002). The surprise element: jumps in interest rates, *Journal of Econometrics* **106**(1): 27–65.
- de Munnik, F. and Schotman, P. (1994). Cross-sectional versus time series estimation of term structure models: Empirical results for the Dutch bond market, *Journal of Banking & Finance* **18**(5): 997–1025.
- Deng, G. (2014). Pricing American put option on zero-coupon bond in a jump-extended CIR model, *Commun Nonlinear Sci Numer Simulat* . <http://dx.doi.org/10.1016/j.cnsns.2014.10.003>.
- Dewachter, H., Iania, L. and Lyrio, M. (2014). Information in the Yield Curve: A macro-finance approach, *Journal of Applied Econometrics* **29**(1): 42–64.
- Diebold, F. and Li, C. (2006). Forecasting the term structure of government bond yields, *Journal of Econometrics* **130**(2): 337–364.
- Diebold, F., Li, C. and Yue, V. (2008). Global Yield Curve Dynamics and Interactions: A Dynamic Nelson-Siegel Approach, *Journal of Econometrics* **146**(2): 351–363.
- Duffee, G. (2002). Term Premia and Interest Rate Forecasts in Affine Models, *The Journal of Finance* **57**(1): 405–443.
- Duffee, G. (2011). Information in (and not in) the Term Structure, *The Review of Financial Studies* **24**(9): 2895–2934.

- Duffie, D. and Kan, R. (1996). A Yield-Factor Model of Interest Rates, *Mathematical Finance* **6**(4): 379–406.
- Dufresne, D. (2001). The integrated square root process. Research paper number 90, University of Montreal.
- Fama, E. (1976). Forward rates as predictors of future spot rates, *Journal of Financial Economics* **3**(4): 361–377.
- Fama, E. and Bliss, R. (1987). The Information in Long-Maturity Forward Rates, *The American Economic Review* **77**(4): 680–692.
- Feller, W. (1951). Two Singular Diffusion Problems, *Annals of Mathematics, Second Series* **54**(1): 173–182.
- Fisher, M. and Gilles, C. (1998). Around and around: The Expectations Hypothesis, *The Journal of Finance* **53**(1): 365–383.
- Garfinkel, M. and Thornton, D. (1995). The Information Content of the Federal Funds Rate: Is it Unique?, *Journal of Money, Credit and Banking* **27**(3): 838–847.
- Gil-Alana, L. and Moreno, A. (2012). Uncovering the US term premium: An alternative route, *Journal of Banking & Finance* **36**(4): 1181–1193.
- Gregory, D. and Livingston, M. (1992). Development of the Market for U.S. Treasury STRIPS, *Financial Analyst Journal* **48**(2): 68–74.
- Grinblatt, M. and Longstaff, F. (2000). Financial Innovation and the Role of Derivative Securities: An Empirical Analysis of the Treasury STRIPS Program, *The Journal of Finance* **55**(3): 1415–1436.
- Guidolin, M. and Timmermann, A. (2009). Forecasts of US short-term interest rates: A flexible forecast combination approach, *Journal of Econometrics* **150**(2): 297–311.

- Hall, A., Anderson, H. and Granger, C. (1992). A Cointegration Analysis of Treasury Bill Yields, *The Review of Economics and Statistics* **74**(1): 116–126.
- Hardouvelis, G. (1988). The Predictive Power of the Term Structure During Recent Monetary Regimes, *The Journal of Finance* **43**(2): 339–356.
- Hin, L.-Y. and Dokuchaev, N. (2015). Computation of the implied discount rate and volatility for an overdefined system using stochastic optimization, *IMA Journal of Management Mathematics* . DOI:10.1093/imaman/dpv007  
<http://imaman.oxfordjournals.org/cgi/content/abstract/dpv007?ijkey=qMTk0IzgOY6vikX&keytype=ref> .
- Huang, R. and Lin, C. (1996). An analysis of nonlinearities in term premiums and forward rates, *Journal of Empirical Finance* **3**(4): 347–368.
- Huang, T., Sun, B. and Chen, X. (2014). Re-specification of Affine Term Structure Models: The Linkage to Empirical Investigations, *Applied Mathematical Finance* **21**(6): 523–554.
- Hull, J. and White, A. (1990). Pricing Interest-Rate-Derivative Securities, *The Review of Financial Studies* **3**(4): 573–592.
- Jamshidian, F. (1995). A simple class of square-root interest-rate models, *Applied Mathematical Finance* **2**(1): 61–72.
- Jones, C. (2003). Nonlinear Mean Reversion in the Short-Term Interest Rate, *The Review of Financial Studies* **16**(3): 793–843.
- Kim, D. (2007). The bond market term premium: what is it, and how can we measure it? BIS Quarterly Review. Available at [http://www.bis.org/publ/qtrpdf/r\\_qt0706e.pdf](http://www.bis.org/publ/qtrpdf/r_qt0706e.pdf).
- Litterman, R. and Scheinkman, J. (1991). Common factors affecting bond returns, *Journal of Fixed Income* **1**(June): 54–61.
- Longstaff, F. (2000). The term structure of very short-term rates: New evidence for the expectations hypothesis, *Journal of Financial Economics* **58**(3): 397–415.

- Madan, D. and Schoutens, W. (2012). Tenor specific pricing, *International Journal of Theoretical and Applied Finance* **15**(6): 1–21.
- Maghsoodi, Y. (1996). Solution of the extended CIR term structure and bond option valuation, *Mathematical Finance* **6**(1): 89–109.
- McAleer, M. and Medeiros, M. (2008a). A multiple regime smooth transition heterogeneous autoregressive model for long memory and asymmetries, *Journal of Econometrics* **147**(1): 104–119.
- McAleer, M. and Medeiros, M. (2008b). Realized Volatility: A Review, *Econometric Reviews* **27**(1-3): 10–45.
- Mercurio, F. (2009). Interest Rates and The Credit Crunch: New Formulas and Market Models. Bloomberg Portfolio Research Paper No. 2010-01-FRONTIERS. Available at SSRN: <http://ssrn.com/abstract=1332205>.
- Mercurio, F. (2010). Modern Libor Market Models: Using Different Curves for Projecting Rates and for Discounting, *International Journal of Theoretical and Applied Finance* **13**(1): 113–137.
- Merton, R. C. (1980). On Estimating the Expected Return on the Market: An exploratory investigation, *Journal of Financial Economics* **8**(4): 323–361.
- Modigliani, F. and Sutch, R. (1966). Innovations in Interest Rate Policy, *The American Economic Review* **56**(1): 178–197.
- Moench, E. (2008). Forecasting the Yield Curve in a data-rich environment: A no-arbitrage factor-augmented VAR approach, *Journal of Econometrics* **146**(1): 26–43.
- Moreni, N. and Pallavicini, A. (2014). Parsimonious HJM Modelling for Multiple Yield Curve Dynamics, *Quantitative Finance* **14**(2): 199–210.
- Piazzesi, M. (2005). Bond Yields and the Federal Reserve, *Journal of Political Economy* **113**(2): 311–344.

- R Core Team (2012). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.  
**URL:** <http://www.R-project.org/>
- Roberds, W., Runkle, D. and Whiteman, C. (1996). A Daily View of Yield Spreads and Short-Term Interest Rate Movements, *Journal of Money, Credit and Banking* **28**(1): 34–53.
- Rudebusch, G. (1995). Federal Reserve interest rate targeting, rational expectations, and the term structure, *Journal of Monetary Economics* **35**(2): 245–274.
- Rudebusch, G. (2012). The Bond Premium in a DSGE Model with Long-Run Real and Nominal Risks, *American Economic Journal: Macroeconomics* **4**(1): 105–143.
- Sarno, L. and Thornton, D. (2003). The dynamic relationship between the Federal Funds Rate and the Treasury Bill Rate: An empirical investigation, *Journal of Banking & Finance* **27**(6): 1079–1110.
- Schlögl, E. and Schlögl, L. (2000). A square root interest rate model fitting discrete initial term structure data, *Applied Mathematical Finance* **7**(3): 183–209.
- Simon, D. (1990). Expectations and the Treasury Bill-Federal Funds Rate Spread over Recent Monetary Policy Regimes, *The Journal of Finance* **45**(2): 567–577.
- Stanton, R. (1997). A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk, *The Journal of Finance* **52**(5): 1973–2002.
- Steeley, J. M. (1997). Implied volatility from the term structure: a simple analytical approximation, *Economics Letters* **57**(3): 345–352.
- Theil, H. (1966). *Applied Economic Forecasts*, North Holland, Amsterdam.
- Tzavalis, E. (2004). The term premium and the puzzles of the expectations hypothesis of the term structure, *Economic Modelling* **21**(1): 73–93.

- Van Horne, J. (1980). The Term Structure: A Test of the Segmented Market Hypothesis, *Southern Economic Journal* **46**(4): 1129–1140.
- Wright, J. and Zhou, H. (2009). Bond risk premia and realized jump risk, *Journal of Banking & Finance* **33**(12): 2333–2345.
- Zhang, J. and Sanderson, A. (2009). JADE: Adaptive Differential Evolution with Optional External Archive, *IEEE Transactions on Evolutionary Computation* **13**(5): 945–958.