Thammasat Review 2014, 17(2): 171-193

A Simplified Linear Projection Approach to Estimate Daily Real Yields in Thailand's Bond Market

Anya Khanthavit
Faculty of Commerce and Accountancy, Thammasat University
akhantha@tu.ac.th

Abstract

Certain approaches can be applied to estimate real yields on a daily basis for Thailand's bond market. The estimation is complicated, data-intensive and time-consuming; hence it is not very useful to practitioners. This study proposes a simple and practical approach which practitioners can actually use. Simplicity and practicality result from the use of readily available, lagged nominal yields for projection variables and from the choice of less computationally-intensive, qualified but less efficient diagonal matrix for minimum chi-square estimation. Using daily nominal yield data of up to 15-year maturity from July 30, 2013 to August 8, 2014 - 250 observations as are commonly chosen by practitioners - together with scaled average headline inflation, the study finds a normal shape for the average daily real curve. The estimation is successful and fast.

Keywords: Daily Real Yields, Affine Mutifactor Interest Rate Model, Daily Real-Yield Estimation, Minimum-Chi-Square Estimation

Introduction

Certain approaches can be applied to estimate daily real yields for Thailand's bond market. For example, in a conventional way one may assume a multifactor interest model such as Joyce *et al.* (2010) that describes both nominal and real yields and then estimate them by Kalman filtering using daily nominal yield data and scaled average daily inflation data. Recently, Khanthavit (2014a, b, c) proposed a linear projection approach that can estimate real yields on a daily basis from daily nominal yields and monthly inflation data. Although the estimation of daily real yields by these approaches is possible, it is not very useful to practitioners. The Kalman filtering approach is complicated and numerically challenging, while the less complicated and less numerically challenging approaches of Khanthavit (2014a, b, c) need projection variables which are not readily available to practitioners.

In this study, I propose a simple approach to estimate daily real yields for Thailand's bond market. It extends the Khanthavit (2014a, b, c) studies by considering lagged nominal yields as being projection variables. This simplifies the model estimation substantially because lagged nominal yields are readily available to practitioners. The estimation is less data demanding and intensive. Moreover, with respect to Hamilton and Wu's (2012) analysis, because in a latent multifactor interest model nominal yields are determined by latent factors. The latent factors can be inferred from the projection nominal yields. Based on this functional relationship, I am able to relate the model parameters with the regression coefficients of nominal yields on projection lagged nominal yields. It turns out that the number of parameters to be estimated is reduced from that in Khanthavit (2014a, b, c) by the number of latent factors times the number of projection variables. The resulting empirical model is much less complicated.

I estimate the model parameters by minimum chi-square estimation. Rothenberg (1973) suggests an efficient weighting matrix be used so that the resulting estimates are efficient. In this study the efficient weighting matrix is the inverse of autocorrelation consistent covariance matrix of the regression coefficients. But using the efficient weighting matrix introduces two practical problems. Firstly, the study considers a large number of nominal yields in the estimation. So, the system of regression equations is large and it is difficult to compute the autocorrelation consistent covariance matrix. Secondly, the number of regression coefficients is large. Even if the autocorrelation consistent covariance matrix is available, the minimization of the chi-square objective function is numerically challenging. The minimization problem is highly non-linear and it involves the inversion of a large covariance matrix.

Rothenberg (1973) explains that in minimum chi-square estimation any positive semi-definite weighting matrix can give unbiased and consistent estimators. Matrix efficiency enhances the estimators by making them efficient.

Being aware that practitioners weigh more for practicality, I propose to use a diagonal weighting matrix whose diagonal elements are the inverses of consistent variances of the regression coefficients. Although it is less efficient, the matrix is positive semi-definite and is therefore qualified. The variances can be estimated from single-regression equations of their corresponding yields, rather than from a system of regression equations. The chi-square minimization problem can avoid large matrix inversion and reduce to a sum-of-squared-scaled-error problem.

The estimation is fast, in which the solution can be obtained within less than 5 minutes.[1] Using daily nominal yield data of up to 15-year maturity from July 30, 2013 to August 8, 2014 –250 observations as are commonly chosen by practitioners - together with scaled average headline inflation, the study finds a normal shape for the average daily real term structure.

The Model

I adopt Joyce *et al.* (2010) to describe nominal and real yields in Thailand. The model is an essentially affine term structure model which relates the nominal and real yields with a set of latent factors linearly under a no-arbitrage condition in the real world. It is flexible for it allows time-varying risk premiums and real short rate. The number of latent factors can be raised to capture complex behavior of the yields. Moreover, a latent factor model is found in previous studies to fit yields better than a macro factor model.

Turn first to the pricing of real and nominal bonds. In a no-arbitrage environment, the time-t price $P_t^{n,R}$ of a zero-coupon real bond with an n-period maturity must be given by (Cochrane (2005) where M_{t+j} is the real pricing

$$P_t^{n,R} = E_t\{M_{t+1}M_{t+2} \dots M_{t+n}\}, \qquad (1)$$

kernel in j periods hence and $E_t\{.\}$ is the conditional expectation operator in the real world. The price $P_t^{n,N}$ of a zero-coupon nominal bond is given in a similar way but with the nominal pricing kernel $M_{t+j}^* = M_{t+j}\frac{I_{t+j-1}}{I_{t+j}}$ being substituted for M_{t+j}. I_{t+j} is the consumer price index at time t+j.

$$P_t^{n,N} = E_t\{M_{t+1}^* M_{t+2}^* \dots M_{t+n}^*\}. \qquad (2)$$

From eqs. (1) and (2), because the real yield $y_t^{n,R}$ and nominal yield $y_t^{n,N}$ are $-\frac{1}{n}Ln\{P_t^{n,R}\}$ and $-\frac{1}{n}Ln\{P_t^{n,N}\}$, up to a second order approximation the yields must equal

$$y_t^{n,R} = -\frac{1}{n}\left\{E_t\left(\sum_{j=1}^{n} m_{t+j}\right) + \frac{1}{2}V_t\left(\sum_{j=1}^{n} m_{t+j}\right)\right\} \qquad (3.1)$$

[1] I demonstrate simplicity and practicality of the proposed approach by estimating the model parameters and inferring real yields in Microsoft Excel spreadsheets. The computation time depends on starting values and computer speed. Readers may obtain the Excel spreadsheets from the author upon request.

$$y_t^{n,N} = -\frac{1}{n}\left\{E_t\left(\sum_{j=1}^{n}\left(m_{t+j} - \pi_{t+j}\right)\right) + \frac{1}{2}V_t\left(\sum_{j=1}^{n}\left(m_{t+j} - \pi_{t+j}\right)\right)\right\}, \qquad (3.2)$$

where $m_{t+j} = Ln\{M_{t+j}\}$. $\pi_{t+j} = Ln\left\{\frac{I_{t+j-1}}{I_{t+j}}\right\}$ is log inflation. $V_t(.)$ is the variance operator conditioned on the information at time t.

Turn next to the stochastic behavior of pricing kernels. The logged, real pricing kernel m_{t+1} takes on the form as in eq. (4).

$$m_{t+1} = -(\bar{r} + \boldsymbol{\gamma}^T \mathbf{z}_t) - \frac{\Lambda_t' \Omega \Lambda_t}{2} - \boldsymbol{\Lambda}_t' \Omega^{\frac{1}{2}} \boldsymbol{\varepsilon}_{t+1} \qquad (4)$$

The term $(\bar{r} + \boldsymbol{\gamma}' \mathbf{z}_t)$ is the real short rate. It can vary over time with a set of K latent factors $\mathbf{z}_t' = [z_{1,t}, \dots, z_{K,t}]$. The real short rate is constant if $\boldsymbol{\gamma}' = [\gamma_1, \dots, \gamma_K]$ is a zero vector. Vector $\boldsymbol{\Lambda}_t' \Omega^{\frac{1}{2}}$ is time-varying risk premiums.

$$\boldsymbol{\Lambda}_t = \boldsymbol{\lambda} + \beta \mathbf{z}_t. \qquad (5)$$

Vector $\boldsymbol{\lambda}' = [\lambda_1, \dots, \lambda_K]$ and matrix $\beta = \begin{bmatrix} \beta_{11} & \dots & \beta_{1K} \\ \vdots & \ddots & \vdots \\ \beta_{K1} & \cdots & \beta_{KK} \end{bmatrix}$. The risk premium for factor k is constant if vector $[\beta_{k1}, \dots, \beta_{kK}]$ is zero. $\boldsymbol{\varepsilon}_{t+1}' = [\varepsilon_{1,t+1}, \dots, \varepsilon_{K,t+1}]$ are Gaussian shocks of factors $\mathbf{z}_{t+1}$. Their mean vector is zero and their covariance matrix is $\Omega = \begin{bmatrix} \sigma_1^2 & 0 \dots & 0 \\ 0 & \ddots & \vdots \\ \vdots & & 0 \\ 0 & \dots 0 & \sigma_K^2 \end{bmatrix}$. Factors $\mathbf{z}_{t+1}$ follow a VAR(1) process in eq. (6).

$$\mathbf{z}_{t+1} = \varphi \mathbf{z}_t + \boldsymbol{\varepsilon}_{t+1}. \qquad (6)$$

Coefficient matrix $\varphi = \begin{bmatrix} \varphi_{11} & 0 \dots & 0 \\ \varphi_{21} & \varphi_{22}\ 0 \dots & \vdots \\ \vdots & \ddots & 0 \\ \varphi_{K1} & \varphi_{K2} \ \dots & \varphi_{KK} \end{bmatrix}$ is a lower triangular matrix.

Because the logged nominal pricing kernel m_{t+1}^* is $m_{t+1} - \pi_{t+1}$, from eq. (4) it must equal

$$m_{t+1}^* = -(\bar{r} + \boldsymbol{\gamma}' \mathbf{z}_t) - \frac{\Lambda_t' \Omega \Lambda_t}{2} - \boldsymbol{\Lambda}_t' \Omega^{\frac{1}{2}} \boldsymbol{\varepsilon}_{t+1} - \pi_{t+1}. \qquad (7)$$

Following Duffie and Kan (1996), Joyce *et al*. (2010) derived the solutions for the real and nominal yields as affine functions of latent factors in eqs. (8) and (9).

$$y_t^{n,R} = -\frac{1}{n}\{A_n + \mathbf{B}_n'\mathbf{z}_t\} \quad (8)$$

$$y_t^{n,N} = -\frac{1}{n}\{A_n^* + \mathbf{B}_n^{*\prime}\mathbf{z}_t\}, \quad (9)$$

where $A_0 = A_0^* = 0.00$ and $\mathbf{B}_0 = \mathbf{B}_0^*$ are (Kx1) zero vectors. Coefficients $A_{n>0}$ and $A_{n>0}^*$ and vectors $\mathbf{B}_{n>0}$ and $\mathbf{B}_{n>0}^*$ are determined sequentially with respect to the systems of equations (10).

$$A_n = -\bar{r} + A_{n-1} - \mathbf{B}_{n-1}'\Omega\boldsymbol{\lambda} + \frac{1}{2}\mathbf{B}_{n-1}'\Omega\mathbf{B}_{n-1} \quad (10.1)$$

$$\mathbf{B}_n' = -\boldsymbol{\gamma}' + \mathbf{B}_{n-1}'(\varphi - \Omega\beta) \quad (10.2)$$

and

$$A_n^* = -\bar{r} - \mu_\pi + A_{n-1}^* - \mathbf{B}_n^{*\prime}\Omega\boldsymbol{\lambda}^* + \frac{1}{2}\mathbf{B}_{n-1}^{*\prime}\Omega\mathbf{B}_{n-1}^* + \frac{\sigma_1^2}{2} + \sigma_1^2\lambda_1 \quad (10.3)$$

$$\mathbf{B}_n^{*\prime} = -(\boldsymbol{\gamma}' + \varphi_1) + \mathbf{B}_{n-1}^{*\prime}(\varphi - \Omega\beta) + \boldsymbol{\iota}'\Omega\beta, \quad (10.4)$$

where $\varphi_1 = [\varphi_{11} \quad 0 \ldots \quad 0]$, $\boldsymbol{\iota}' = [1 \quad 0 \ldots \quad 0]$ and $\boldsymbol{\lambda}^* = \boldsymbol{\lambda} + \boldsymbol{\iota}$. μ_π is the unconditional mean of the inflation. The specifications (10.3) and (10.4) are specific to the perfect correlation assumption of factor $z_{1,t}$ with inflation π_t. Modification needs be made under a different assumption for π_t.

Model Estimation

Because factors $\mathbf{z}_t$ are latent, the econometrician will have to relate them with observed variables. From eq. (9), the measurement equations for day t are given by

$$\begin{bmatrix} y_t^{n_1,N} \\ \vdots \\ y_t^{n_H,N} \end{bmatrix} = \begin{bmatrix} -\frac{1}{n_1}A_{n_1}^* \\ \vdots \\ -\frac{1}{n_H}A_{n_H}^* \end{bmatrix} + \begin{bmatrix} -\frac{1}{n_1}\mathbf{B}_{n_1}^{*\prime} \\ \vdots \\ -\frac{1}{n_H}\mathbf{B}_{n_H}^{*\prime} \end{bmatrix}\mathbf{z}_t + \begin{bmatrix} \omega_{n_1,t} \\ \vdots \\ \omega_{n_H,t} \end{bmatrix} \quad (11)$$

$y_t^{n_h,N}$ is the daily nominal yield with an n_h-day maturity. With respect to Piazzesi (2010), a month of 21 trading days is assumed. So, n_h is 21h and 252h days for h-month and h-year maturities respectively. $\omega_{n_h,t}$ is the measurement error due to, for example, bid-ask spreads and zero-curve interpolation.

Khanthavit (2014a, b, c) propose a linear projection approach to estimate the model on a daily basis even though inflation is reported monthly. Latent factors $\mathbf{z}_t$ can be projected linearly by a set of η observed projection variable $\mathbf{q}_t' = [q_{0,t} = 1, q_{1,t}, \dots, q_{\eta-1,t}]$. The projection equation is given by

$$\mathbf{z}_t = \mathbf{b}'\mathbf{q}_t + \mathbf{v}_t, \tag{12}$$

where $\mathbf{b}' = \begin{bmatrix} b_{1,0}, b_{1,1}, \dots, b_{1,\eta-1} \\ \vdots \\ b_{K,0}, b_{K,1}, \dots, b_{K,\eta-1} \end{bmatrix}$ is the matrix of projection coefficients and $\mathbf{v}_t' = [v_{1,t}, \dots, v_{K,t}]$ are projection errors. The linear projection approach follows Mishkin (1981) who estimated unobserved real yields by information variables. When $\mathbf{b}'\mathbf{q}_t + \mathbf{v}_t$ is substituted for $\mathbf{z}_t$ in eq. (11), eq. (13) is obtained.

$$\begin{bmatrix} y_t^{n_1,N} \\ \vdots \\ y_t^{n_H,N} \end{bmatrix} = \boldsymbol{\alpha}^{\mathbf{T}}\mathbf{q_t} + \mathbf{u}_t, \tag{13}$$

where the $\boldsymbol{\alpha}$ matrix has non-linear functional relationships with the model's parameter vector $[\bar{r}, \pi, \lambda_1, \dots, \lambda_K, \beta_{11}, \beta_{12}, \dots, \beta_{KK}, \varphi_{11}, \dots, \varphi_{11}, \sigma_2^2, \dots, \sigma_K^2]'$ and the projection coefficient vector $[b_{1,0}, b_{2,1}, \dots, b_{K,\eta-1}]'$. $\mathbf{u}_t = \begin{bmatrix} \omega_{n_1,t} - \frac{1}{n_1}\mathbf{B}_{n_1}^{*\prime}\mathbf{v}_t \\ \vdots \\ \omega_{n_H,t} - \frac{1}{n_H}\mathbf{B}_{n_H}^{*\prime}\mathbf{v}_t \end{bmatrix}$.

The projection variables in Khanthavit's (2014a, b, c) studies are a constant and four 1-day lagged Bjork and Christensen (1999) beta shape factors. I am aware that the beta shape factors are not readily available to practitioners. Hence this choice for projection variables is not very practical. I propose 1-day lagged nominal yields as the alternative. Practitioners already have nominal yields and hence their lags. Moreover, as I will show below, considering lagged nominal yields as projection variables can reduce the number of parameters to be estimated substantially from that of Khanthavit (2014a, b, c), so that the resulting empirical model is less complicated and less computationally intensive.

To proceed, I separate the H nominal yields being considered in the analysis into two groups. Group one contains $K < H$ nominal yields $\mathbf{Y}_t^{1,N} = \begin{bmatrix} y_t^{11,N} \\ \vdots \\ y_t^{1K,N} \end{bmatrix}$ whose 1-day lags will serve as the projection variables. The projection variables are $\mathbf{q_t} = \begin{bmatrix} 1 \\ y_{t-1}^{11,N} \\ \vdots \\ y_{t-1}^{1K,N} \end{bmatrix}$. Group two are the remainders $\mathbf{Y}_t^{2,N} = \begin{bmatrix} y_t^{21,N} \\ \vdots \\ y_t^{2(H-K),N} \end{bmatrix}$. From eq. (11), $\mathbf{z}_t$ is related with $\mathbf{Y}_t^{1,N}$ by

$$\mathbf{z}_t = \mathcal{B}_1^{*-1}\left(\mathbf{Y}_t^{1,N} - \mathcal{A}_1^*\right) - \mathcal{B}_1^{*-1}\boldsymbol{\omega}_{1,t} \tag{14.1}$$

so that

$$\mathbf{z}_{t-1} = \mathcal{B}_1^{*-1}\left(\mathbf{Y}_{t-1}^{1,N} - \mathcal{A}_1^*\right) - \mathcal{B}_1^{*-1}\boldsymbol{\omega}_{1,t-1}, \tag{14.2}$$

where for brevity $\mathcal{B}_1^* = \begin{bmatrix} -\frac{1}{n_{1,1}}\mathbf{B}_{1,1}^{*\prime} \\ \vdots \\ -\frac{1}{n_{1,K}}\mathbf{B}_{1,K}^{*\prime} \end{bmatrix}$ and $\mathcal{A}_1^* = \begin{bmatrix} -\frac{1}{n_{1,1}}A_{1,1}^* \\ \vdots \\ -\frac{1}{n_{1,K}}A_{1,K}^* \end{bmatrix}$. $\boldsymbol{\omega}_{1,t} = \begin{bmatrix} \omega_{11,t} \\ \vdots \\ \omega_{1K,t} \end{bmatrix}$. By substituting and rearranging terms in eqs. (6), (13) and (14), the regression equations of $\mathbf{Y}_t^{1,N}$ on $\mathbf{q_t} = \begin{bmatrix} 1 \\ y_{t-1}^{11,N} \\ \vdots \\ y_{t-1}^{1K,N} \end{bmatrix}$ can be written as in eq. (15),

$$\begin{aligned}\mathbf{Y}_t^{1,N} &= \left(\mathcal{A}_1^* - \mathcal{B}_1^*\varphi\mathcal{B}_1^{*-1}\mathcal{A}_1^*\right) + \left(\mathcal{B}_1^*\varphi\mathcal{B}_1^{*-1}\right)\mathbf{Y}_{t-1}^{1,N} \\ &\quad + \left(\mathcal{B}_1^*\boldsymbol{\varepsilon}_t + \boldsymbol{\omega}_{1,t} - \mathcal{B}_1^*\varphi\boldsymbol{\omega}_{1,t-1}\right),\end{aligned} \tag{15}$$

and those of $\mathbf{Y}_t^{2,N}$ on $\mathbf{q_t}$ can be written as in eq. (16),

$$\begin{aligned}\mathbf{Y}_t^{2,N} &= \left(\mathcal{A}_2^* - \mathcal{B}_2^*\varphi\mathcal{B}_1^{*-1}\mathcal{A}_1^*\right) + \left(\mathcal{B}_2^*\varphi\mathcal{B}_1^{*-1}\right)\mathbf{Y}_{t-1}^{1,N} \\ &\quad + \left(\mathcal{B}_2^*\boldsymbol{\varepsilon}_t + \boldsymbol{\omega}_{2,t} - \mathcal{B}_2^*\varphi\boldsymbol{\omega}_{1,t-1}\right),\end{aligned} \tag{16}$$

where $\mathcal{B}_2^* = \begin{bmatrix} -\frac{1}{n_{2,1}}\mathbf{B}_{2,1}^{*\prime} \\ \vdots \\ -\frac{1}{n_{2,(H-K)}}\mathbf{B}_{2,(H-K)}^{*\prime} \end{bmatrix}$ and $\mathcal{A}_2^* = \begin{bmatrix} -\frac{1}{n_{2,1}}A_{2,1}^* \\ \vdots \\ -\frac{1}{n_{2,(H-K)}}A_{2,(H-K)}^* \end{bmatrix}$. $\boldsymbol{\omega}_{2,t} = \begin{bmatrix} \omega_{21,t} \\ \vdots \\ \omega_{2(H-K),t} \end{bmatrix}$.

It is important to note that the parameters of the empirical model in eqs. (15) and (16) reduce to only $[\mu_\pi, \bar{r}, \lambda_1, \dots, \lambda_K, \beta_{11}, \dots, \beta_{KK}, \varphi_{11}, \dots, \varphi_{KK}, \sigma_1^2, \dots, \sigma_K^2]'$. The projection coefficients need not be estimated.

Although eqs. (15) and (16) and those in Hamilton and Wu (2012) look similar, their derivations, properties and implications differ in two important ways. Firstly, Hamilton and Wu (2012) relate $\mathbf{Y}_t^{1,N}$ with $\mathbf{Y}_{t-1}^{1,N}$ by the theoretical relationship in eq. (9), while I do by the empirical relationship in eq. (11). Secondly, Hamilton and Wu (2012) relate $\mathbf{Y}_t^{2,N}$ with $\mathbf{Y}_t^{1,N}$, while I relate $\mathbf{Y}_t^{2,N}$ with $\mathbf{Y}_{t-1}^{1,N}$. Our different ways of relating variables lead us to consider different econometric techniques to estimate the regression coefficients and their covariances. In this study, from eqs. (15) and (16) because $\mathbf{Y}_{t-1}^{1,N}$ is endogenous and the regression errors are autocorrelated, I will have to use instrumental variable (IV) regressions as opposed to simple least squared regressions in Hamilton and Wu (2012). The application of IV regressions will be discussed below.

Identification of Interesting Parameters

Similar to Hamilton and Yu (2012), Khanthavit (2014b, c) acknowledge that all the model parameters need not be estimated jointly but sequentially in steps. The expectation μ_π for daily inflation can be inferred from monthly inflation data in step one. Once μ_π is obtained, it can be employed together with daily nominal yield data to identify the remaining parameters in step two.

At this point, the remaining parameters that must be estimated are $= [\bar{r}, \lambda_1, \cdots, \lambda_K, \beta_{11}, \ldots, \beta_{KK}, \varphi_{11}, \ldots, \varphi_{KK}, \sigma_1^2, \ldots, \sigma_K^2]'$. One way to proceed in the second step is to follow Khanthavit (2014b) to estimate them by nonlinear SURE using daily nominal yield curves and the μ_π estimate from the first step. But as Hamilton and Wu (2012) and Khanthavit (2014c) pointed out, the estimation is a numerical challenge because the objective surfaces are highly nonlinear and they behave badly. In order to lessen the computation time, I will follow Khanthavit (2014c) to apply Rothenberg's (1973) minimum-chi-square estimation for the problem.

Consider the following system of linear regression equations of daily nominal yields on the projection variables.

$$\begin{bmatrix}\mathbf{Y}_t^{1,N}\\ \mathbf{Y}_t^{2,N}\end{bmatrix} = \begin{bmatrix}\begin{bmatrix}c_{11,0}\\ \vdots\\ c_{1K,0}\end{bmatrix}\\ \begin{bmatrix}c_{21,0}\\ \vdots\\ c_{2(H-K),0}\end{bmatrix}\end{bmatrix} + \begin{bmatrix}\begin{bmatrix}c_{11,1}, \ldots, c_{11,K}\\ \vdots\\ c_{1K,1}, \ldots, c_{1K,K}\end{bmatrix}\\ \begin{bmatrix}c_{21,1}, \ldots, c_{21,K}\\ \vdots\\ c_{2(H-K),1}, \ldots, c_{2(H-K),K}\end{bmatrix}\end{bmatrix}\mathbf{Y}_{t-1}^{1,N} + \begin{bmatrix}\mathbf{W}_t^{1,N}\\ \mathbf{W}_t^{2,N}\end{bmatrix}, \quad (17)$$

where $\begin{bmatrix}\begin{bmatrix}c_{11,0}\\ \vdots\\ c_{1K,0}\end{bmatrix}\\ \begin{bmatrix}c_{21,0}\\ \vdots\\ c_{2(H-K),0}\end{bmatrix}\end{bmatrix}$ is the vector of intercepts and $\begin{bmatrix}\begin{bmatrix}c_{11,1}, \ldots, c_{11,K}\\ \vdots\\ c_{1K,1}, \ldots, c_{1K,K}\end{bmatrix}\\ \begin{bmatrix}c_{21,1}, \ldots, c_{21,K}\\ \vdots\\ c_{2(H-K),1}, \ldots, c_{2(H-K),K}\end{bmatrix}\end{bmatrix}$ is the matrix of slope coefficients. $\begin{bmatrix}\mathbf{W}_t^{1,N}\\ \mathbf{W}_t^{2,N}\end{bmatrix}$ is the vector of regression errors.

Define $\mathbf{C} = Vech\left(\begin{matrix}\begin{bmatrix}c_{11,0}\\ \vdots \\ c_{1K,0}\end{bmatrix}\\ \vdots \\ \begin{bmatrix}c_{21,1}, \cdots, c_{21,K}\\ \vdots \\ c_{2(H-K),1}, \cdots, c_{2(H-K),K}\end{bmatrix}\end{matrix}\right)$ as the vector of regression coefficients and R is the covariance matrix of $\mathbf{C}$. $\mathbf{g}(\boldsymbol{\Theta}) = Vech\begin{pmatrix}\mathcal{A}_1^* - \mathcal{B}_1^*\varphi\mathcal{B}_1^{*-1}\mathcal{A}_1^*\\ \vdots \\ \mathcal{B}_2^*\varphi\mathcal{B}_1^{*-1}\end{pmatrix}$ is the vector of functions $\mathbf{g}(\boldsymbol{\Theta})$ of the remaining parameters $\boldsymbol{\Theta}$ that describe the shape of nominal curves. Rothenberg (1973) shows that the remaining parameters can be estimated by minimizing the chi-square statistic χ^2 in eq. (18) with respect to $\boldsymbol{\Theta}$.

$$\chi^2 = \tau[\mathbf{C} - \mathbf{g}(\boldsymbol{\Theta})]'\mathrm{R}^{-1}[\mathbf{C} - \mathbf{g}(\boldsymbol{\Theta})]. \tag{18}$$

where τ is the number of observations. The minimizers $\widehat{\boldsymbol{\Theta}}$ have the property $\sqrt{\tau}(\widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}) \rightarrow Normal\left(0, \left[\left(\frac{\partial \mathbf{g}(\boldsymbol{\Theta})}{\partial \boldsymbol{\Theta}'}\right)' \mathrm{R}^{-1} \left(\frac{\partial \mathbf{g}(\boldsymbol{\Theta})}{\partial \boldsymbol{\Theta}'}\right)\right]^{-1}\right)$.

The Econometrics

The study must estimate the coefficient vector $\mathbf{C}$ and its covariance matrix R from daily $\begin{bmatrix}\mathbf{Y}_t^{1,N}\\ \mathbf{Y}_t^{2,N}\end{bmatrix}$ and $\mathbf{Y}_{t-1}^{1,N}$ data. Because the regressors $\mathbf{Y}_{t-1}^{1,N}$ are endogenous, conventional OLS regressions will give biased and inconsistent $\mathbf{C}$. To obtain unbiased and consistent $\mathbf{C}$, I choose IV regression estimation. In general, it is difficult to find IVs that are highly correlated with the regressors and orthogonal to the regression errors. In this study, I consider 2-day lagged $\mathbf{Y}_{t-2}^{1,N}$ as being the IVs. I have two reasons. Firstly, previous studies such as Khanthavit (2013) reported that nominal yields were strongly autocorrelated. So, the regressors $\mathbf{Y}_{t-1}^{1,N}$ and the IVs $\mathbf{Y}_{t-2}^{1,N}$ must have strong correlations. Secondly, from eqs. (11), (15) and (16), the IVs $\mathbf{Y}_{t-2}^{1,N}$ are orthogonal to the regression errors $\left(\mathcal{B}_1^*\boldsymbol{\varepsilon}_t + \boldsymbol{\omega}_{1,t} - \mathcal{B}_1^*\varphi\boldsymbol{\omega}_{1,t-1}\right)$ and $\left(\mathcal{B}_2^*\boldsymbol{\varepsilon}_t + \boldsymbol{\omega}_{2,t} - \mathcal{B}_2^*\varphi\boldsymbol{\omega}_{1,t-1}\right)$.

The system of IV regression equations can be large when the study considers $\begin{bmatrix}\mathbf{Y}_t^{1,N}\\ \mathbf{Y}_t^{2,N}\end{bmatrix}$ of various tenors. So for a practical purpose I will estimate the regression coefficients $\mathbf{C}_{\{j,i\}} = \begin{bmatrix}c_{\{j=(1 or\ 2)\}\{i=1)\},0}\\ \vdots \\ c_{\{j=(1\ or\ 2)\}\{i=(K\ or\ (H-K)\},K}\end{bmatrix}$ for each $\{j,i\}$ tenor separately by a single-IV-regression equation. Because the regressors are the same across regression equations, the resulting $\mathbf{C}$ from single-IV-regression equations is the same as the one from the system of IV regression equations. The IV estimator $\mathbf{C}_{\{j,i\}}^{IV}$ for $\mathbf{C}_{\{j,i\}}$ is

$$\mathbf{C}_{\{j,i\}}^{IV} = \left[\left[\mathbf{1} \quad \mathbf{Y}_{-2}^{1,N}\right]'\left[\mathbf{1} \quad \mathbf{Y}_{-1}^{1,N}\right]'\right]^{-1}\left[\left[\mathbf{1} \quad \mathbf{Y}_{-2}^{1,N}\right]'\mathbf{Y}^{\{j,i\},N}\right], \tag{19}$$

where $\mathbf{1}$ is a $(\tau \times 1)$ vector of 1's, $\mathbf{Y}_{-1}^{1,N}$ is a $(\tau \times K)$ matrix of 1-day lagged $\mathbf{Y}_{t-1}^{1,N}$, $\mathbf{Y}_{-2}^{1,N}$ is a $(\tau \times K)$ matrix of 2-day lagged $\mathbf{Y}_{t-2}^{1,N}$ and $\mathbf{Y}^{\{j,i\},N}$is a $(\tau \times 1)$ vector of the nominal yield of $\{j,i\}$ tenor.

The covariance matrix of $\mathbf{C}_{\{j,i\}}^{IV}$ can be estimated by a consistent estimator $\Sigma_{\{j,i\}}^{IV}$ in eq. (20).

$$\begin{aligned}\Sigma_{\{j,i\}}^{IV} = &\left[\left[\mathbf{1} \quad \mathbf{Y}_{-2}^{1,N}\right]'\left[\mathbf{1} \quad \mathbf{Y}_{-1}^{1,N}\right]'\right]^{-1}\left[\mathbf{1} \quad \mathbf{Y}_{-2}^{1,N}\right]' \times \Sigma_{\{j,i\}} \\ &\times \left[\mathbf{1} \quad \mathbf{Y}_{-1}^{1,N}\right]\left[\left[\left[\mathbf{1} \quad \mathbf{Y}_{-2}^{1,N}\right]'\left[\mathbf{1} \quad \mathbf{Y}_{-1}^{1,N}\right]'\right]^{-1}\right]',\end{aligned} \tag{20}$$

where $\Sigma_{\{j,i\}}$ is the $(\tau \times \tau)$ covariance matrix of the regression errors for tenor $\{j,i\}$. To estimate $\Sigma_{\{j,i\}}$, turn first to tenor $\{j=2,i\}$. From eq. (16) because the errors are homoskedastic and uncorrelated, $\Sigma_{\{j=2,i\}}$ is $\sigma_{\{j=2,i\}}^2 \times I$, where $\sigma_{\{j=2,i\}}^2$ can be estimated conveniently by the variance of the IV regression errors for tenor $\{j=2,i\}$.

Turn next to tenor $\{j=1,i\}$. From eq. (15) it is important to note that the errors are autocorrelated of order 1. This fact constitutes $\Sigma_{\{j=1,i\}}$, whose elements have the following properties. The diagonal elements are the same and equal to $\sigma_{\{j=1,i\}}^2$--the variance of the errors. The off-diagonal elements $(s,s-1)$ and $(s-1,s)$ are the same and equal to $\sigma_{\{j=1,i\}}(s,s-1)$—the first-order autocovariance of the regression errors. The remaining elements are identically zero. It is not difficult to estimate the variance $\sigma_{\{j=1,i\}}^2$ and autocovariance $\sigma_{\{j=1,i\}}(s,s-1)$ statistics. From the IV regression, once the econometrician obtains the regression errors, the estimates are the errors' variance and autocovariance.

Practicality Problems and Improvements

The regression coefficients $\mathbf{C}$ and their covariance matrix R are needed for estimating the parameters $\mathbf{\Theta}$ by minimum chi-square estimation in eq. (18). Despite the fact that the covariance matrices $\Sigma^{IV}_{\{j,i\}}$'s are obtained for all $\mathbf{C}^{IV}_{\{j,i\}}$'s in $\mathbf{C}$, these matrices are not sufficient to construct R because the covariances among $\mathbf{C}^{IV}_{\{j,i\}}$'s are not available.

It is difficult, especially for practitioners, to estimate R when regression errors are autocorrelated and the system of IV regression equations is large because the estimation involves extremely large matrices of stacked nominal yields, regressors and IVs and the autocorrelation-consistent covariance matrix of stacked regression errors. To proceed, I recall that in minimum chi-square estimation the weighting matrix needs not be R^{-1}. Any positive semi-definite matrix can give unbiased and consistent estimates for $\boldsymbol{\Theta}$. The choice for efficient R is to enhance efficiency of the estimation.

If the approach is not practical, it is not useful to practitioners. So, I sacrifice the use efficient R for the use of a practical, qualified but less efficient, positive-semi-definite matrix. In this study, I propose a diagonal matrix $\mathcal{R}$ whose diagonal elements are the consistent variances of $\mathbf{C}$. These variances can be obtained from the diagonal elements of $\Sigma^{IV}_{\{j,i\}}$'s.

It is interesting to ask how important it is to discard the off-diagonal elements of matrix R when it is substituted for by matrix $\mathcal{R}$. Because the diagonal elements of matrices R and $\mathcal{R}$ are the same, matrix $\mathcal{R}$ offers the same degree of efficiency if the regression coefficients $\mathbf{C}$ are uncorrelated. However, in reality the regression coefficients $\mathbf{C}$ are hardly uncorrelated. To analyze the question in a more realistic case, let us consider the chi-square objective function in eq. (18). This structure compares well with a non-linear least-square regression in which the regression errors are heteroskedastic and autocorrelated. So, substituting $\mathcal{R}$ for R can be thought of as considering only the errors' heteroskedasticity but ignoring autocorrelation in the non-linear regression analysis. In the literature, for some reasons heteroskedasticity or autocorrelation is at times ignored even though it probably affects reliability of the estimates. See, for example, Hamel *et al.* (2012). Here, the effects should not be so severe because the approach maintains unbiasedness and consistence properties of the estimates.

It should be noted that the contribution of $\begin{bmatrix} c_{11,0} \\ \vdots \\ c_{1K,0} \end{bmatrix}$ to the chi-square objective function is to compare it with $\begin{bmatrix} \mathcal{A}_1^* - \mathcal{B}_1^* \varphi \mathcal{B}_1^{*-1} \mathcal{A}_1^* \\ \mathcal{A}_2^* - \mathcal{B}_2^* \varphi \mathcal{B}_1^{*-1} \mathcal{A}_1^* \end{bmatrix}$. This results from the fact that the regressors are $\mathbf{Y}^{1,N}_{t-1}$. Recall the pricing formula in eq. (9) and the zero expected z_t in eq. (6). If I substitute demeaned $\mathbf{Y}^{1,N}_{t-1}$ for $\mathbf{Y}^{1,N}_{t-1}$ in the regression eq. (17), the contribution of $\begin{bmatrix} c_{11,0} \\ \vdots \\ c_{1K,0} \end{bmatrix}$ to the chi-square objective function is to compare it with $\begin{bmatrix} \mathcal{A}_1^* \\ \mathcal{A}_2^* \end{bmatrix}$. Because the $\begin{bmatrix} \mathcal{A}_1^* \\ \mathcal{A}_2^* \end{bmatrix}$ functions are much less complicated than the $\begin{bmatrix}\mathcal{A}_1^* - \mathcal{B}_1^*\varphi\mathcal{B}_1^{*-1}\mathcal{A}_1^* \\ \mathcal{A}_2^* - \mathcal{B}_2^*\varphi\mathcal{B}_1^{*-1}\mathcal{A}_1^*\end{bmatrix}$ functions, in the analysis I will use the demeaned $\mathbf{Y}_{t-1}^{1,N}$ and compare $\begin{bmatrix}c_{11,0} \\ \vdots \\ c_{1K,0}\end{bmatrix}$ with $\begin{bmatrix}\mathcal{A}_1^* \\ \mathcal{A}_2^*\end{bmatrix}$ instead.

The contribution of $Vech\left(\begin{bmatrix}c_{21,1}, \dots, c_{21,K} \\ \vdots \\ c_{2(H-K),1}, \dots, c_{2(H-K),K}\end{bmatrix}\right)$ to the objective function is to compare it with $Vech\begin{pmatrix}\mathcal{B}_1^*\varphi\mathcal{B}_1^{*-1} \\ \mathcal{B}_2^*\varphi\mathcal{B}_1^{*-1}\end{pmatrix}$. The comparison most likely suffers from operational problems because $\mathcal{B}_1^{*-1}$ needs be computed in each step of the gradient search in the minimization. If $\mathcal{B}_1^*$ is singular, the inversion fails and the search stops. To avoid the failure of $\mathcal{B}_1^*$ inversion, I multiply the slope coefficients in $\mathbf{C}$ and in $\mathbf{g}(\boldsymbol{\Theta})$ on the right by $\mathcal{B}_1^*$ to obtain $Vech\left(\begin{bmatrix}[c_{21,1}, \dots, c_{21,K}]\mathcal{B}_1^* \\ \vdots \\ [c_{2(H-K),1}, \dots, c_{2(H-K),K}]\mathcal{B}_1^*\end{bmatrix}\right)$ and $Vech\begin{pmatrix}\mathcal{B}_1^*\varphi \\ \mathcal{B}_2^*\varphi\end{pmatrix}$ so that $\mathcal{B}_1^{*-1}$ is not required any longer in the analysis.

It is important to note that when $Vech\left(\begin{bmatrix}[c_{21,1}, \dots, c_{21,K}]\mathcal{B}_1^* \\ \vdots \\ [c_{2(H-K),1}, \dots, c_{2(H-K),K}]\mathcal{B}_1^*\end{bmatrix}\right)$ and $Vech\begin{pmatrix}\mathcal{B}_1^*\varphi \\ \mathcal{B}_2^*\varphi\end{pmatrix}$ are considered, their corresponding elements of matrix $\mathcal{R}$ should be adjusted accordingly.

To adjust the corresponding diagonal elements of matrix $\mathcal{R}$, suppose the diagonal matrix $Diag([\mathcal{R}_{ji,1}, \dots, \mathcal{R}_{ji,K}]')$ for $[c_{ji,1}, \dots, c_{ji,K}]$ is considered. When $[c_{ji,1}, \dots, c_{ji,K}]$ is adjusted by $\mathcal{B}_1^*$, the diagonal matrix should be modified to $\mathcal{D}_{ji} = \mathcal{B}_1^{*\prime} Diag([\mathcal{R}_{ji,1}, \dots, \mathcal{R}_{ji,K}]')\mathcal{B}_1^*$. In order to maintain practicality of the analysis and the diagonal matrix nature of $\mathcal{R}$, I propose to substitute the diagonal elements of $\mathcal{D}_{ji}$ for $[\mathcal{R}_{ji,1}, \dots, \mathcal{R}_{ji,K}]'$.

The estimation can be improved one step further by standardizing and constraining certain parameters. The standardization helps to reduce the number of parameters to be estimated and the constraints help to limit the areas for parameter search. I follow Hamilton and Wu (2012) to standardize the volatilities of latent factors such that $\sigma_1 = \cdots = \sigma_K$. But instead of setting them equal to 1.00, I set them equal to 0.0001 which is in the same magnitude as the ones found for Thailand by Khanthavit (2014a, b, c). Next I follow Dai and Singleton (2000) to constrain $\boldsymbol{\gamma}' = [\gamma_1, \dots, \gamma_K]$ to be non-negative.

The Data

The study applies the minimum-chi-square technique to estimate daily real yields of up to 15-year maturity in Thailand's bond market. The nominal zero-coupon yield data begin July 30, 2013 and end August 8, 2014. They are constructed by the Thai Bond Market Association. The sample period provides 250 daily observations. I choose this particular sample size because it is the size commonly chosen by practitioners and is accepted by regulators such as the Bank of Thailand (2003) as being a sufficiently large sample size for daily observations. The inflation is the log monthly inflation, computed using the headline consumer price index from the Bureau of Trade and Economic Indices, Ministry of Commerce. The consumer price index data begin May 2000 and end July 2014. The expected daily inflation is set to monthly average inflation divided by 21.

Table 1 reports the descriptive statistics of nominal yields and inflation. The average inflation is 2.6300% when it is scaled to annual rate. This level is within the 0.00-to-3.5 percent band being monitored by the Bank of Thailand under its inflation targeting policy. The average term structure of nominal yields has a normal shape, while its volatility structure has a "U" shape. This finding is similar to what Khanthavit (2014c) reported earlier. However, it is important to note that, due to our different sample periods and sizes, the average levels in my study are about 20 basis points lower for short yields and about 60 basis points lower for long yields than the ones in that study.

Table 1 Descriptive Statistics

Variables	Average	Max	Min	Std.	Skew.	E. Kurt.	JB St.	AR(1)
Inflation	2.6300	25.8264	-36.7878	6.3601	-1.2921	10.1489	729.5896***	0.3258
1M	2.2456	2.5316	1.9988	0.2064	0.1669	-1.6176	28.4160***	0.9963
3M	2.2714	2.5694	2.0145	0.2113	0.2045	-1.6062	28.6149***	0.9967
6M	2.3052	2.6182	2.0530	0.2183	0.2584	-1.5869	29.0165***	0.9964
1Y	2.3286	2.6291	2.0750	0.2113	0.2854	-1.5584	28.6925***	0.9962
2Y	2.5724	3.1071	2.2277	0.2726	0.3849	-1.3539	25.2668***	0.9951
3Y	2.8422	3.3432	2.3857	0.2907	0.1085	-1.4758	23.1790***	0.9958
4Y	3.2001	3.6814	2.8956	0.1934	0.4855	-0.7005	14.9346***	0.9841
5Y	3.4229	3.8172	2.9882	0.2245	-0.0748	-1.2440	16.3531***	0.9972
6Y	3.5538	4.0376	3.1187	0.2387	0.2853	-0.9597	12.9860***	0.9884
7Y	3.7987	4.2369	3.3962	0.2016	0.1829	-0.5000	3.9975	0.9915
8Y	3.8313	4.3151	3.4049	0.2080	0.2684	-0.4160	4.8045*	0.9830
9Y	3.9251	4.4452	3.5102	0.2044	0.4912	-0.0987	10.1529***	0.9790
10Y	4.0090	4.5562	3.4563	0.2440	0.1053	-0.4801	2.8625	0.9829
11Y	4.1204	4.6327	3.7091	0.2119	0.3672	-0.1985	6.0273**	0.9882
12Y	4.1780	4.6529	3.8664	0.1873	0.5665	0.2509	14.0255***	0.9896
13Y	4.1855	4.6612	3.8563	0.1917	0.3750	-0.1591	6.1243**	0.9925
14Y	4.2179	4.7072	3.8585	0.2008	0.1021	-0.5567	3.6625	0.9956
15Y	4.3132	4.7797	3.8915	0.2150	-0.1169	-0.5837	4.1178	0.9989

Note: The statistics for inflation are based on monthly data, while those for nominal interest rates are based on daily data. *, ** and *** = significance at 90%, 95% and 99% confidence levels, respectively. The monthly sample for inflation is from May 2000 to July 2014 (170 monthly observations) and the daily sample for nominal yields is from July 30, 2013 to August 8, 2014 (250 daily observations).

The study tests for normality of the nominal yields. The Jarque-Bera (JB) tests reject the assumptions for most of the yields in the sample. Nevertheless, non-normality of the nominal yields affects neither biasedness nor consistence of the IV estimates.

Finally, in the last column Table 1 reports the first-order autocorrelation coefficients of the inflation and nominal yields. The AR(1) coefficients of nominal yields are positive, very high and close to 1.00. This finding supports the use of 2-day lagged $\mathbf{Y}_{t-2}^{1,N}$ as IVs because they are highly correlated with the 1-day lagged $\mathbf{Y}_{t-1}^{1,N}$ regressors. But it may raise concern as to whether the nominal yields are I(1) variables. As for this particular time-series property, Khanthavit (2013) reported that the nominal yields were not I (1) variables but they followed long-memory processes.

In order to identify the number of latent factors, Khanthavit (2014a, b, c) employed long time-series data of more than 10 years in the principal component analyses and found that the first two principal components could explain about 98% of the yields' variation. Here, I use a sample of approximately one year. I perform the principal component analysis based on the one-year recent data of nominal yields to reexamine the number of factors. The result is reported in Table 2. I find similar results. The first two principal components can explain 96.70% of the variation. With respect to this finding, I conclude that the number K of latent factor is 2.

Table 2 Principal Component Analysis

Principal Component	Contribution	Accumulated Contribution
1	88.8606%	88.8606%
2	7.8473%	96.7079%
3	1.7342%	98.4421%
4 and Beyond	1.5579%	100.0000%

Because there are $K = 2$ factors, I will have to choose two nominal yields to serve as the regressors and IVs. With respect to the relationship in eqs. (14.1) and (14.2), nominal yields of any two tenors are equivalent. I choose 3-year and 7-year tenors for two reasons. Firstly, the 3-year and 7-year tenors are considered benchmark tenors in Thailand's bond market. The 3-year tenor can represent bonds of shorter tenors and the 7-year tenor can represent bonds of longer tenors. Secondly, Kiatnakin Bank compiled cumulative trading values of the bonds in each tenor from October 2012 to July 2013 and found that the value for the 3-year tenor was the highest and that for the 7-year tenor is the second highest.

Empirical Results

Table 3 reports the IV regression coefficients of nominal yields on a constant 1-day lagged demeaned 3-year and 7-year yields. All the coefficients are significant at a 99% confidence level. As was pointed out by Khanthavit (2014a, b), highly significant regression coefficients result from the long memory property of nominal yields. The IV regression coefficients and their consistent standard errors will be used for parameter estimation based on the minimum chi-square objective.

Table 3 IV Regression Coefficients

Maturity	Constant	Demeaned, Lagged 3Y Yield	Demeaned, Lagged 7Y Yield
1M	8.91E-05***	0.7732***	-0.1565***
3M	9.01E-05***	0.7985***	-0.1669***
6M	9.15E-05***	0.8745***	-0.2483***
1Y	9.24E-05***	0.8696***	-0.2724***
2Y	1.02E-04***	1.0574***	-0.2368***
3Y	1.13E-04***	1.0060***	-0.0195***
4Y	1.27E-04***	0.4241***	0.3180***
5Y	1.36E-04***	0.3631***	0.5917***
6Y	1.41E-04***	0.3680***	0.6330***
7Y	1.51E-04***	0.0359***	0.9373***
8Y	1.52E-04***	0.0886***	0.8610***
9Y	1.56E-04***	-0.2162***	1.2281***
10Y	1.59E-04***	-0.3569***	1.5910***
11Y	1.64E-04***	-0.2606***	1.3398***
12Y	1.66E-04***	-0.1875***	1.1234***
13Y	1.66E-04***	-0.1124***	1.0483***
14Y	1.67E-04***	-0.0615***	1.0134***
15Y	1.71E-04***	0.0044***	0.9829***

Note: The instrumental variables are constant, demeaned lagged-two 3Y yield and demeaned lagged-two 7Y yield, respectively. *** = significant at a 99% confidence level, computed using autocorrelation consistent standard errors.

Table 4 reports parameters $[\mu_\pi, \bar{r}, \lambda_1, \lambda_2, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}, \varphi_{11}, \varphi_{21}, \varphi_{22}, \sigma_1^2, \sigma_2^2]'$ of the model. $\mu_\pi = 2.6300\%$ is the monthly average inflation multiplied by 12. The value $\sigma_1 = \sigma_2 = 0.0001$ is fixed for standardization. The remaining parameters are from minimum chi-square estimation. It is found that the parameter estimates are not very close to the ones reported by Khanthavit (2014c) who uses the same minimum chi-square estimation technique. The differences are not surprising but should be expected because Khanthavit (2014c) considers a much longer sample period and a different set of projection variables. Moreover, when this study considers lagged nominal yields as projection variables, the projection coefficients need not be estimated but are absorbed in the parameters of the theoretical model.

Table 4 Parameter Estimates

Parameters	**Value**
$\bar{r} \times 25200$	4.1729
γ_1	13.0020
γ_2	0.0479
λ_1	1345.5976
λ_2	-1579.1845
β_{11}	1.6503
β_{12}	-829.9405
β_{21}	4.2856
β_{22}	-770.2675
φ_{11}	0.0032
φ_{21}	0.9998
φ_{22}	0.9999
σ_1	0.0001
σ_2	0.0001
$\mu_{\pi} \times 25200$	2.6300

I follow Ang *et al.* (2008) to conduct specification tests for the model. If the model fits, the moments of sample and fitted nominal yields should not differ. Comparison of the means, standard deviations, skewnesses and excess kurtoses are in Table 5. The numbers in the first lines are for fitted yields and those in the second lines are their deviations from the sample moments. Significance is based on the White (2000) procedure.

The deviations are small and not significant for all the moments and across maturities, except for the standard deviations of almost all maturities. These findings may result from the fact that the unbiased and consistent but less efficient weighting matrix is substituted for the efficient one in minimum chi-square estimation. The resulting unbiasedness and consistence are indicated by the small and insignificant deviations for the first moments.

Table 5 Specification Tests

Maturity	Descriptive Statistics			
	Mean	**Std.**	**Skew.**	**E. Kurt**
1M	2.2738	0.3585	-0.5954	-0.6094
	0.0282	0.1521***	-0.7623	1.0082
3M	2.1812	0.3628	-0.5953	-0.6042
	-0.0902	0.1514**	-0.7998	1.002
6M	2.2245	0.3636	-0.5953	-0.6029
	-0.0806	0.1453**	-0.8537	0.984
1Y	2.3615	0.3636	-0.5953	-0.6023
	0.0328	0.1524***	-0.8806	0.9562
2Y	2.6436	0.3629	-0.5952	-0.602
	0.0712	0.0903*	-0.9801	0.7519
3Y	2.9063	0.3619	-0.5952	-0.6019
	0.0641	0.0712	-0.7038	0.874
4Y	3.1453	0.3609	-0.5952	-0.6018
	-0.0548	0.1675***	-1.0808	0.0987
5Y	3.3601	0.3599	-0.5952	-0.6018
	-0.0628	0.1354**	-0.5204	0.6422
6Y	3.5505	0.3589	-0.5952	-0.6018
	-0.0033	0.1201**	-0.8805	0.358
7Y	3.7169	0.3609	-0.5952	-0.6018
	-0.0818	0.1594**	-0.7781	-0.1018
8Y	3.8594	0.3568	-0.5952	-0.6017
	0.0281	0.1488**	-0.8636	-0.1857
9Y	3.9783	0.3558	-0.5952	-0.6017
	0.0533	0.1515***	-1.0864	-0.503
10Y	4.074	0.3548	-0.5952	-0.6017
	0.0649	0.1108**	-0.7005	-0.1216
11Y	4.1466	0.3538	-0.5952	-0.6017
	0.0262	0.1419**	-0.9624	-0.4032
12Y	4.1964	0.3528	-0.5952	-0.6017
	0.0184	0.1655***	-1.1617*	-0.8526
13Y	4.2238	0.3518	-0.5952	-0.6017
	0.0384	0.1601***	-0.9703	-0.4426

14Y	4.2291 0.0112	0.3508 0.1499**	-0.5952 -0.6973	-0.6017 -0.045
15Y	4.2125 -0.1007	0.3498 0.1348**	-0.5952 -0.4783	-0.6017 -0.0180

Note: *, ** and *** = significance at 90%, 95% and 99% confidence levels, respectively. The statistics on the upper lines are those of the fitted yields and the ones on the lower lines are the deviations from sample statistics.

The significant deviations for the standard deviations may be induced by the use of less efficient weighting matrix. However, the significance should not cause poorer performance of the estimation when it is compared with that of competing approaches. The reasons are as follows. The significance of standard deviations was also reported for most specifications of the Ang *et al.* (2008) model. Moreover, when it is compared with Khanthavit (2014c) who uses an efficient weighting matrix, Khanthavit (2014c) reports 13 significance cases while this study does 17 significance cases.

In Panel 6.1 of Table 6, the term structure of Thailand's real yields is time varying. Its average has a normal shape. The averages for 1-month up to 6-month maturities are negative. They turn positive and rising for a 1-year maturity and over. When compared to those in Khanthavit (2014c), the average curve is much lower. This result is expected because our sample periods differ and the average nominal curve in this study is much lower than that in Khanthavit (2014c).

Finally, the figure in Panel 6.2 shows the real curve for August 8, 2014, which is day t=0 or the current date in the estimation. The figure demonstrates to practitioners that the real curve can be updated daily by an up-to-date sample and model re-estimation.

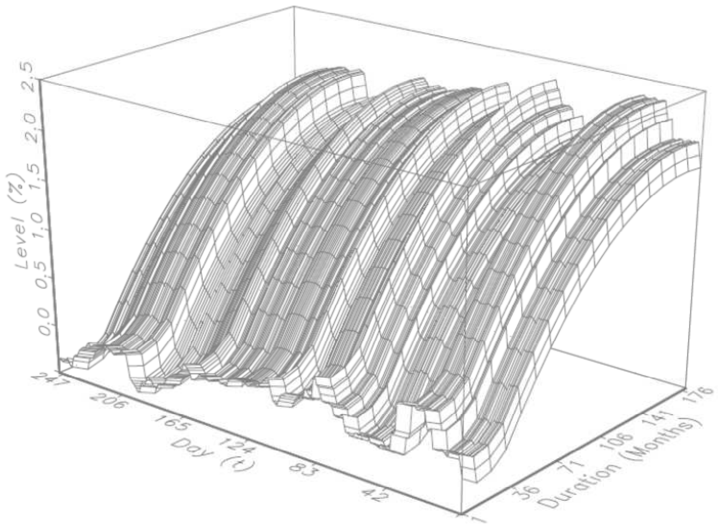
Conclusion

Certain approaches that can be applied to estimate Thailand's daily real yields are not very useful to practitioners because they are complicated, data-intensive or numerically challenging. This study proposes a simple and practical approach which practitioners can actually use. It is based on minimum chi-square estimation. But the efficient weighting matrix is replaced by a qualified but less efficient diagonal positive-semi-definite weighting matrix. Here, simplicity and practicality are demonstrated by parameter estimation and real-yield inference in Microsoft-Excel spreadsheets. It turns out the estimation and inference are successful and fast.

Although the proposed approach is intended for practitioners in Thailand, it is general and can be applied in those countries where the markets for inflation-linked bonds and inflation derivatives are inactive or inexistent but the data on daily nominal yields and monthly inflation are available. It is interesting to ask how practical the proposed approach is in other emerging markets. I leave this question for future research.

Table 6 Daily Term Structures

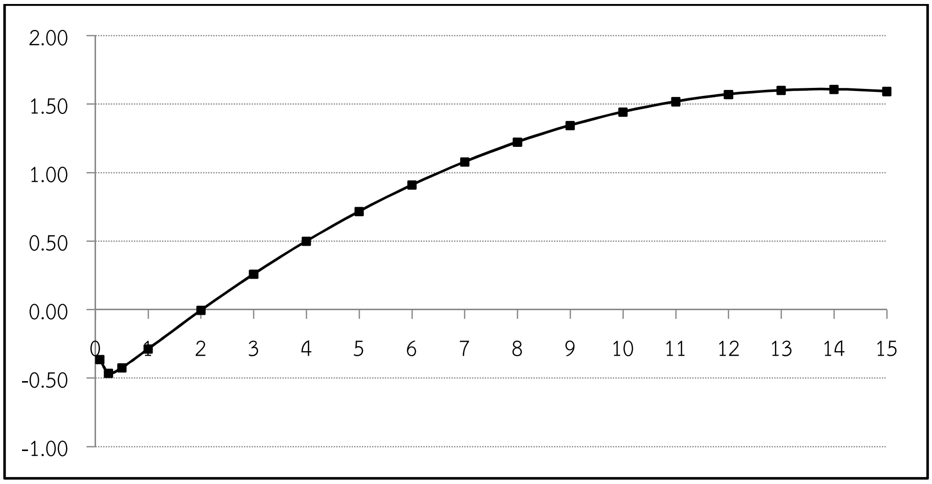
Panel 6.1 Real Yields



Maturity	Average	Max	Min	Std.
1M	-0.0176	-0.4130	0.3960	0.1818
3M	-0.1043	-0.4959	0.3053	0.1807
6M	-0.0595	-0.4500	0.3490	0.1807
1Y	0.0782	-0.3113	0.4858	0.1805
2Y	0.3610	-0.0273	0.7674	0.1801
3Y	0.6246	0.2372	1.0299	0.1797
4Y	0.8646	0.4783	1.2689	0.1793
5Y	1.0806	0.6952	1.4839	0.1789
6Y	1.2726	0.8881	1.6749	0.1784
7Y	1.4405	1.0570	1.8418	0.1780
8Y	1.5847	1.2021	1.9850	0.1776
9Y	1.7053	1.3237	2.1047	0.1771
10Y	1.8027	1.4220	2.2011	0.1767
11Y	1.8769	1.4971	2.2744	0.1763
12Y	1.9284	1.5495	2.3248	0.1759
13Y	1.9572	1.5793	2.3527	0.1754
14Y	1.9637	1.5867	2.3583	0.1750
15Y	1.9481	1.5720	2.3417	0.1746

Note: Day t=1 is July 30, 2013 and Day t=250 is August 8, 2014.

Panel 6.2 Estimated Real Yields on Day t=0, August 8, 2014



References

Ang, J., G. Bekeart, and M. Wei (2008). The term structure of real rates and expected inflation. *Journal of Finance*, *63*, 797-849.

Bank of Thailand. (2003). Market risk audit manual. Retrieved from http://www.bot.or.th/English/FinancialInstitutions/PruReg_HB/RiskMgt_Manual/DocLib_DocumentForDownload/RiskManagementExaminationManaul_04_MarketRisk.pdf

Bjork, T. and B. Christensen. (1999). Interest rate dynamics and consistent forward rate curves. *Mathematical Finance*, *9*, 323-348.

Cochrane, J. (2005). *Asset Pricing*. New Jersey: Princeton University Press.

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

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

Hamel, S., N. Yoccoz, and J. Gaillard. (2012). Statistical evaluation of parameters estimating autocorrelation and individual heterogeneity in longitudinal studies. *Methods in Ecology and Evolution*, *3*, 731-742.

Hamilton, J. D., and J. C. Wu. (2012). Identification and estimation of Gaussian affine term structure models. *Journal of Econometrics,* 186 (2), 315-331.

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

Khanthavit, A. (2013). Selecting a parsimoniously parametric model of Thailand's term structure of interest rates. *Journal of Business Administration*, *36*, 15-39.

Khanthavit, A. (2014a). Estimating daily real yields and expected inflations for Thailand's financial market. *Journal of Business Administration*, *37*, 29-52.

Khanthavit, A. (2014b). An improved linear projection approach to estimate daily real yields and expected inflations in a latent multifactor interest model, Unpublished Manuscript. Faculty of Commerce and Accountancy. Thammasat University. Bangkok. Forthcoming in *ABAC Journal*, *34*.

Khanthavit, A. (2014c). A fast minimum-chi-square estimation of Thailand's daily real yields. Unpublished Manuscript, Faculty of Commerce and Accountancy, Thammasat University, Bangkok. Forthcoming in *NIDA Business Journal*, *15*.

Mishkin, F. (1981). The real interest rate: An empirical investigation. *Carnegie-Rochester Conference Series on Public Policy*, *15*, 151-200.

Piazzesi, M. (2010). Affine term structure models. In Ait-Sahalia, Y., and Hansen, L., *Handbook of Financial Econometrics 1*: *Tools and Techniques* (pp. 691-766). The Netherlands: Elsevier.

Rothenberg, T. (1973). *Efficient Estimation with a Priori Information*. Connecticut: Yale University Press.

White, H. (2000). A reality check for data snooping. *Econometrica*, *68*: 1097-1126.

Acknowledgement

The author thanks the Faculty of Commerce and Accountancy, Thammasat University for the research grant and thanks the Thai Bond Market Association for the interest rate data.