Centre for Computational Finance and Economic Agents

Working Paper Series

WP037-09

J. Castellanos and N. Constantinou

Comparison of Two Factor CIR and Essentially Affine Models for the UL Term Structure: from Black Wednesday to the 2008 Credit Crisis

November 2009

University of Essex

CCFEA

www.ccfea.net

Comparison of Two Factor CIR and Essentially Affine Models for the UK Term Structure: from Black Wednesday to the 2008 Credit Crisis*

Jenny Castellanos† and Nick Constantinou‡§

November 2, 2009

Abstract

This paper explores the differences and relative goodness of fit of two term structure models, the CIR two factor model and a two factor essentially affine model, $EA_1(2)$. The latter model generates correlation between the factors and time-varying risk premia. However these characteristics increase the complexity of the model and is therefore computationally more intensive. The comparison is made using data that incorporates both Black Wednesday and the 2008 credit crisis, covering the widest range of short term rates in recent UK history. It is found that the added complexity of the essentially affine model only marginally improves the fit to the UK term structure for the time period studied. Both models provide a good fit to the observed yield curve.

1 Introduction

Over the past few years, researchers and finance practitioners have focused on the accurate modelling of the term structure. One of the main goals has

*We would like to acknowledge valuable discussions with Wing Lon Ng.

†Centre for Computational Finance and Economic Agents (CCFEA), University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK. jbcast@essex.ac.uk

‡Essex Bussiness School,University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK. nconst@essex.ac.uk

§corresponding author

been to try to understand the dynamics and to create an accurate model that fits the observed yield curve both at the long and short end. The first models employed were the so called affine models where the bond yields were assumed to be affine (constant plus linear) functions of some state variable. These types of models were favored due to their tractability and flexibility.

The initial impetus was provided by Vasicek (1977) and Cox, Ingersoll and Ross (1985) (CIR hereafter), who assumed that the instantaneous interest rate (or short rate) was an affine function of a single factor, the state variable. The short rate was modelled as a diffusion where both the drift and the diffusions terms were affine functions of the short rate itself. Other models were later developed by Ho and Lee (1986), Hull and White (1990) and Longstaff and Schwartz (1992) who introduced stochastic volatility, among other features.

The Vasicek and CIR models remain popular as both models have the advantage of being affine and tractable as well as yielding a simple close form solution for the bond price. The CIR model has an advantage over the Vasicek model as it assures positive rates. However as highlighted by Canabarro (1985), single factor models, when estimated empirically, do not accurately fit the observed yield curves which led to multifactor generalisations.

The multifactor CIR model has been investigated by various authors, Chen and Scott (2003), De Jong (1998), Duan and Simonato (1999), Geyer and Pichler (1999) and De Jong and Santa Clara (1999) focusing on US data and Nath and Nowman (2001) on UK data. The general finding is the poor fit of the single factor CIR model, for example Chen and Scott (2003) reported measurement errors greater than 100 basis points for the single factor model while for the two factor model the errors did not excedeed 35 basis points providing a significant improvement.

The CIR model and its multifactor generalisations lie within the category of completely affine models (Duffee 2002). The market price of risk is given by a constant parameter that implies affine dynamics in both the real and risk neutral measure. Duffee (2002) went further and introduced essentially affine models. In these models, the price of risk is not completely defined by the variation of the volatility of the yields, implying that the price of risk can vary independently. The simplest multifactor model with correlated factors and essentially affine characteristics (denoted by "E") is, the $\mathrm{EA}_1(2)$ model. Following the notation of Dai and Singleton (2000), the model has two factors but only one of them determines the conditional variance of the factors or state variables.

Duffee's insight was to allow for time varying risk premia and with this added flexibility generated an improved fit to the U.S. Treasury term structure. However this is at the cost of an increase in mathematical complexity and the number of parameters (e.g. the $EA_1(2)$ model has four more parameters than the two factor CIR model which has eight). Dai and Singleton (2000) also concluded the need for correlation among the state variables and time varying volatility. Cheridito et al. (2007) compared different market price of risk specifications and using likelihood ratio statistics showed that their extended affine specification fitted better than the completely affine model. Driessen (2005) also employed the $EA_1(2)$ model to fit the term structure of interest rates for his study on the decomposition of corporate bond spreads. These studies focused on US data and support the use of the $EA_1(2)$ model to fit the term structure.

The term structure models studied in this paper are calibrated using a panel approach as implemented by Chen and Scott (2003). This approach avoids the disadvantages of the cross section and time series approaches by taking into account both the dynamics of the model and the observed yields. The actual implementation is achieved by representing the problem in a state space form and combining the Kalman Filter (described in Section 2.4) with a maximum likelihood (or a quasi maximum likelihood) approach. Alternative methodologies have been used in the literature to estimate or model the term structure based on the CIR model. Some examples include the Efficient Method of Moments used by Dai and Singleton (2000), and the Maximum Likelihood Method used by Chen and Scott (1993).

Different calibration methods have different disadvantages, some use proxies, others are less efficient or are computationally expensive. Specific to the Kalman Filter, it has been noted in the literature that for non Gaussian models such as the CIR model, the Kalman filter can not give an exact maximum likelihood, instead an approximation is done to generate a quasi maximum likelihood. Chen and Scott (2003), Lund (1997) and Duan and Simonato (1999) validated the methodology by showing in Monte Carlo experiments that although there is a bias in the estimators, it is negligible. These results imply that the quasi maximum likelihood is an appropriate approach to use.

The motivation for this paper is to compare the relative fitness that both the CIR two factor and $EA_1(2)$ model give to the UK term structure from February 1992 until April 2009. The UK term structure has seen a sharp decrease in rates following actions to help alleviate the recent credit crisis. The 6-months rate has decreased from 5% to 0.6% between 2007/08 and across the whole period, the yield has ranged between 10.59% and 0.6%. This paper focuses on modelling the UK term structure for a time period that includes Black Wednesday (16th September 1992) and the credit crisis of 2008.

The rest of this paper is organized as follows: Section 2 presents the two factor CIR and $EA_1(2)$ model, the state space formulation and the Kalman filter. Section 3 consists of the empirical results and analysis, and Section 4 concludes.

2 Theoretical Framework

This section outlines the general theoretical framework followed by a description of the two factor CIR and the $EA_1(2)$ model. The section concludes with a description of the Kalman filter methodology and in particular the transformation of the two models to state space form.

2.1 Term Structure

A common group of models used to characterize the term structure are the so-called affine models. Bond yields are linearly related to the underlying state variables. Different states create different bond prices implying that the dynamics of the term structure depends on the evolution of the state variables. The mathematical framework is outlined below. Hereafter the superscript Q is used to differenciate the risk neutral measure from the P or 'real-world' measure.

The time t price of a default free zero coupon bond expiring at time T (Duffie and Kan (1996)), is given by the following expression:

$$P(t,\tau) = E_t^Q \left[\exp\left(-\int_t^T r(s)\, ds \right) \right] \tag{1}$$

where r is the (instantaneous) risk free rate and E^Q is the risk neutral expectation. Given that the instantaneous rate is known it is sufficient information to be able to characterize the whole term structure. Duffie and Kan(1996) show that equation (1) can also be expressed in the following way, where the price $P(t,\tau)$ of a discount bond that matures in $\tau = (T-t)$ years can be given by:

$$P(t,\tau) = \exp\left(A(\tau) - B(\tau)\, r(t)\right). \tag{2}$$

In equation (2) $r(t)$, also known as the short rate, can be modelled within a single or a multifactor framework. In both cases, the short rate is an affine function of unobservable N-state variables, represented by the vector $\mathbf{Y}(t)$. In its most general form, the short rate is defined as

$$r(t) = \delta_0 + \boldsymbol{\delta}_y' \mathbf{Y}(t) \tag{3}$$

where $\mathbf{Y}'(t) = (Y_1(t), Y_2(t), ..., Y_N(t))$, δ_0 represents a constant term and $\boldsymbol{\delta}_y' = (\delta_1, \delta_2, ..., \delta_N)$, a vector of state variable loadings. The vector $\mathbf{Y}(t)$

follows the following process (Dai and Singleton (2000))

$$d\mathbf{Y}(t) = \mathbf{K}^Q\left(\boldsymbol{\theta}^Q - \mathbf{Y}(t)\right)dt + \boldsymbol{\Sigma}\sqrt{\mathbf{S}(t)}\mathbf{dW}^Q(t). \tag{4}$$

In the above, $\mathbf{K}^Q$ and $\boldsymbol{\Sigma}$ are $N \times N$ matrices, $\boldsymbol{\theta}^Q$ is a N-vector, $\mathbf{W}^Q$ is a N-dimensional independent Wiener process, while $\mathbf{S}(t)$ is a $N \times N$ diagonal matrix with diagonal elements (conditional variances) given by

$$[S(t)]_{ii} = \alpha_i + \boldsymbol{\beta}_i'\mathbf{Y}(t) \qquad for \;\; i = 1, N \tag{5}$$

where α_i is a constant and $\boldsymbol{\beta}_i$ is a N-vector.

Given the dynamics of the short rate and the expressions for $A(\tau)$ and $B(\tau)$, the bond price may be calculated using equation(2) where $A(\tau)$ and $B(\tau)$ satisfy the following ordinary differential equations

$$\frac{dA(\tau)}{d\tau} = -(\boldsymbol{\theta}^Q)'(\mathbf{K}^Q)'\mathbf{B}(\tau) + \frac{1}{2}\sum_{i=1}^{N}\left[\boldsymbol{\Sigma}'\mathbf{B}(\tau)\right]_i^2 \alpha_i - \delta_0 \tag{6}$$

$$\frac{d\mathbf{B}(\tau)}{d\tau} = -(\mathbf{K}^Q)'\mathbf{B}(\tau) - \frac{1}{2}\sum_{i=1}^{N}\left[\boldsymbol{\Sigma}'\mathbf{B}(\tau)\right]_i^2 \boldsymbol{\beta}_i + \boldsymbol{\delta}_y \tag{7}$$

with initial conditions $A(0) = 0$ and $\mathbf{B}(0) = \mathbf{0}_{N\times 1}$ (see Dai and Singleton 2000). In general the ODEs have to be solved numerically however for some models such as the multifactor CIR model, analytical expressions can be derived.

For empirical studies equation (4) must be expressed under the real or objective measure P. To change the measure some assumption about the market price of risk is necessary. Duffee(2002) characterizes completely affine models as those where the price of risk vector $\boldsymbol{\Lambda}(t)$ is given by $\boldsymbol{\Lambda}(t) = \sqrt{\mathbf{S}(t)}\boldsymbol{\lambda}$ where $\boldsymbol{\lambda}$ is a N-vector of constants. From Girsanov's Theorem

$$\mathbf{dW}^Q(t) = \mathbf{dW}(t) + \boldsymbol{\Lambda}(t)dt, \tag{8}$$

replacing equation (8) in equation (4) yields the P dynamics which are given by (see Dai and Singleton 2000)

$$d\mathbf{Y}(t) = \mathbf{K}(\boldsymbol{\theta} - \mathbf{Y}(t))dt + \boldsymbol{\Sigma}\sqrt{\mathbf{S}(t)}\mathbf{dW}(t) \tag{9}$$

where

$$\mathbf{K} = \mathbf{K}^Q - \boldsymbol{\Sigma}\boldsymbol{\Phi}$$

and

$$\boldsymbol{\theta} = \mathbf{K}^{-1}\left(\mathbf{K}^Q\boldsymbol{\theta}^Q + \boldsymbol{\Sigma}\boldsymbol{\Psi}\right).$$

The *ith* row of $\boldsymbol{\Phi}$ is given by $\lambda_i\boldsymbol{\beta}_i'$ and $\boldsymbol{\Psi}$ is an N-vector with elements $\lambda_i\alpha_i$.

2.2 Two Factor CIR Model

This section reviews the formulation and specification of the uncorrelated two factor ($N = 2$) CIR model (see Chen and Scott 2003). The model is presented under the P measure in the following SDEs:

$$dY_i(t) = \kappa_i \left(\theta_i - Y_i(t)\right) dt + \sigma_i \sqrt{Y_i(t)} dW_i(t) \; for \;\; i = 1, 2. \tag{10}$$

The two state variables are independent, $W_1(t)$ and $W_2(t)$ are independent Wiener processes. The parameters $\kappa_1$and κ_2 represent the speed of mean reversion of the short rate for each individual process, while $\theta_1$and θ_2 represent the risk neutral long run mean, and σ_1 and σ_2 the volatility of the factors. The form of the CIR diffusion restricts the resulting rates to be positive and only when the condition $2\kappa_i\theta_i > \sigma_i^2$does not hold, rates can become zero.

Equation (10) can also be expressed under the Q measure, by introducing the market price of risk (λ). The CIR model is a completely affine model where the transformation of measures is straightforward since (see Nawalkha 2007)

$$\Lambda(t)_i = \lambda_i \sqrt{Y_i(t)}/\sigma_i.$$

This transforms equation (10) to the following expression for the risk neutral dynamics of the two factor CIR model:

$$dY_i(t) = \left(\kappa_i\theta_i - \left(\kappa_i + \lambda_i\right) Y_i(t)\right) dt + \sigma_i \sqrt{Y_i(t)} dW_i^Q(t). \tag{11}$$

The bond price formula in equation (1) is given by

$$P\left(t, \tau\right) = \prod_{i=1}^{2} A_i\left(\tau\right) \exp\left\{ - \sum_{i=1}^{2} B_i\left(\tau\right) Y_i\left(t\right) \right\}. \tag{12}$$

The ODEs in equations (6) and (7) can be solved analytically for the CIR model giving a close form solution for the bond price and are given by the following equations (see e.g. Geyer and Pichler 1999)

$$A_i\left(\tau\right) = \left[\frac{2\gamma_i \exp\left(\frac{1}{2}\left(\kappa_i + \lambda_i + \gamma_i\right)\tau\right)}{2\gamma_i + \left(\kappa_i + \lambda_i + \gamma_i\right)\left(\exp\left(\gamma_i\tau\right) - 1\right)} \right]^{\frac{2\kappa_i\theta_i}{\sigma_i^2}} \tag{13}$$

$$B_i\left(\tau\right) = \frac{2\left(\exp\left(\gamma_i\tau\right) - 1\right)}{2\gamma_i + \left(\kappa_i + \lambda_i + \gamma_i\right)\left(\exp\left(\gamma_i\tau\right) - 1\right)} \tag{14}$$

$$\gamma_i = \sqrt{\left(\kappa_i + \lambda_i\right)^2 + 2\sigma_i^2}. \tag{15}$$

The above equations determine the bond price under the two factor CIR model, with parameters to be determined from the empirical data via the Kalman filter method.

2.3 The $EA_1(2)$ model

This section outlines the framework for the more general two factor essentially affine model. The $EA_1(2)$ notation follows from Dai and Singleton (2000) which denotes that the model has two factors but only one of them determines the conditional variance of $\mathbf{Y}$. The $EA_1(2)$ model is represented by the following matrix equation

$$\begin{bmatrix} dY_1(t) \\ dY_2(t) \end{bmatrix} = \begin{bmatrix} \kappa_{11} & 0 \\ \kappa_{21} & \kappa_{22} \end{bmatrix} \begin{bmatrix} \theta_1 - Y_1(t) \\ -Y_2(t) \end{bmatrix} dt + \begin{bmatrix} \sqrt{Y_1(t)} & 0 \\ 0 & \sqrt{1+\beta_{21} Y_1(t)} \end{bmatrix} \begin{bmatrix} dW_1(t) \\ dW_2(t) \end{bmatrix} \quad (16)$$

which is a canonical representation of equation (9)(see e.g. Driessen 2005). For an arbitrary choice of parameters, the general specification of equation (9) is not admissible[1], and constraints to the drift parameter ($\mathbf{K}$ and $\boldsymbol{\theta}$) and diffusion coefficients $\boldsymbol{\beta}$ need to be imposed ($\boldsymbol{\Sigma}$ is the identity matrix under the canonical representation). Dai and Singleton (2000) specified additional parametric restrictions to ensure admissibility, for the essentially affine $EA_1(2)$ the restrictions:

$$\delta_2 \geq 0, \quad \kappa_{21} \leq 0, \quad \theta_1 > 0, \quad \beta_{21} \geq 0, \quad \kappa_{11} > 0. \quad (17)$$

Equation (16) illustrates the first factor follows a similar specification as the CIR model however the second factor is coupled to the first factor. For this reason this model is considered to be correlated as both processes have common factors; the Wiener processes however are still independent.

The $EA_1(2)$ processes are expressed under the P measure but bond prices are estimated under the risk neutral measure. The transformation between one measure and the other is determined by the form of the bond risk premia as discussed before. For the $EA_1(2)$ model presented in this paper the price of risk follows (Duffee (2002))

$$\boldsymbol{\Lambda}(t) = \sqrt{\mathbf{S}(t)} \boldsymbol{\lambda}_1 + \mathbf{S}(t)^{-} \boldsymbol{\lambda}_2 \mathbf{Y}(t)$$

and explicitely given by the following equation:

$$\begin{aligned} \boldsymbol{\Lambda}(t) = & \begin{bmatrix} \sqrt{Y_1(t)} & 0 \\ 0 & \sqrt{1+\beta_{21} Y_1(t)} \end{bmatrix} \begin{bmatrix} \lambda_{1(11)} \\ \lambda_{1(21)} \end{bmatrix} + \\ & \begin{bmatrix} 0 & 0 \\ 0 & (1+\beta_{21} Y_1(t))^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ \lambda_{2(21)} & \lambda_{2(22)} \end{bmatrix} \begin{bmatrix} Y_1(t) \\ Y_2(t) \end{bmatrix} \end{aligned} \quad (18)$$

[1] Dai and Singleton (2000) defined an admissible model as a specification of equation(9) where the resulting values of $S(t)_{ii}$ are strictly positive for all i.

In the above equation the first part of the equation corresponds to the completely affine specification ($\mathbf{\Lambda}(t) = \sqrt{S(t)}\boldsymbol{\lambda}$) where there is only one market risk parameter per factor. The second part of the equation corresponds to the additional flexibility that the essentially affine model allows.

To express equation (16) in the risk neutral measure (for bond pricing purposes), the P measure drift parameters need to be transformed into their corresponding Q measure given the form of the price of risk given in equation (18); the following are the transformed parameters (Duffee 2002):

$$\mathbf{K}^Q = \mathbf{K} + \mathbf{\Phi} + \mathbf{I}^{-}\boldsymbol{\lambda}_2 \tag{19}$$

$$\boldsymbol{\theta}^Q = \left[\mathbf{K}^Q\right]^{-1}(\mathbf{K}\boldsymbol{\theta} - \mathbf{\Psi}), \tag{20}$$

where $\mathbf{\Phi}$ is a 2×2 matrix with the i^{th} row is defined by $\lambda_{1i}\boldsymbol{\beta}'_i$, and $\mathbf{\Psi}$ is a 2×1 vector with the i^{th} element given by $\lambda_{1i}\alpha_i$. $\mathbf{I}^{-}$ is a 2×2 diagonal matrix with $\mathbf{I}^{-}_{ii} = 1$ if $\mathbf{S}^{-}_{ii} \neq 0$, $\mathbf{I}^{-}_{ii} = 0$ if $\mathbf{S}^{-}_{ii} = 0$. The various parameters are estimated using the Kalman Filter described in the following section.

2.4 The Kalman Filter

The Kalman Filter methodology combined with a quasi maximum likelyhood approach is used to estimate the model parameters. This methodology is useful as it treats the short rate as the unobservable variable and uses its relationship with observable yields to estimate model parameters and a time series of the factors. The Kalman Filter is a recursive estimator that has been used in the context of affine term structure modelling in the studies of Duan and Simonato (1999), Lund (1997), and De Jong (1998) among others. This technique is well suited for this class of problem as the underlying state variables are not observable. The state variables are filtered using the relationship between the observed variable (bond yields) and the state variables (measurement equation) and the dynamics through time of the state variables (transition equation).

The measurement equation is used in the updating phase of the filter while the transition equation is used in the prediction phase. Once the equations are established the standard recursive steps are implemented to generate, through the Kalman gain, a minimum mean square error of the state variables. The recursive inferences are used to construct and maximize a log-likelihood function to find the optimal parameter set. To incorporate the above into the CIR and $\mathrm{EA}_1(2)$ models it is necessary to express the models in a state space form. The following section outlines this form for the different models studied in this paper.

For both models the instantaneous nominal interest rate is assumed to be the sum of two state variables as observed in equation (3). The specification of the CIR model implies that $\delta_0 = 0$ and $\boldsymbol{\delta}_y = (1, 1)$, i.e. the short rate is just the sum of the individual factors. For the $\mathrm{EA}_1(2)$ model, δ_0 and $\boldsymbol{\delta}_y$ are not predetermined, instead these parameters are estimated by the Kalman filter.

2.4.1 Transition Equation

The CIR and $\mathrm{EA}_1(2)$ models are continuous time models. However for the state variable estimation they need to be time discretized (for convenience the time dependence will be denoted by subscripts). The transition equation is recursive and defines how the state variable evolves through time. As demonstrated by Bolder (2001) and Chen and Scott (2003), the transition equation can be expressed in the following way:

$$\mathbf{Y}_t = \mathbf{C} + \mathbf{H}\mathbf{Y}_{t-1} + \boldsymbol{\varepsilon}_t \tag{21}$$

$\mathbf{Y}_t$, represents the state variable vector for time t. In the above, $\mathbf{C}$ (2×1 vector) and $\mathbf{H}$ (2×2 matrix) follow from the mean of the transition density of $r(t)$ over a discrete time interval and $\boldsymbol{\varepsilon}_t$ denotes a vector of noise terms. The details of the derivation can be found in Bolder (2001). For the CIR model elements of $\mathbf{C}$ are given by

$$C_i = \theta_i \left(1 - \exp\left(-\kappa_i \Delta t\right)\right) \quad for\ i = 1, 2 \tag{22}$$

where Δt denotes the time step of the discretization ($\Delta t = 1/52$ for weekly data). For the CIR model $\mathbf{H}$ is a diagonal matrix where the diagonal elements are given by:

$$H_{ii} = \exp\left(-\kappa_i \Delta t\right) \tag{23}$$

The last term in the transition equation is a noise term that is assumed to follow a normal distribution, $\varepsilon_{it} \sim N(0, \mathbf{Q}_t)$ where $\mathbf{Q}_t$ is a $N \times N$ matrix. The elements of $\mathbf{Q}_t$ represent the variances of the transition densities of the factors which are employed by the Kalman Filter. For the CIR model these values depend on the previous estimate (filtered value) of the factor denoted by $\widehat{Y}_{i,t-1}$. As the CIR two factor model is uncorrelated, $\mathbf{Q}_t$ is a diagonal matrix with elements given by Bolder (2001):

$$q_{ii} = \frac{\theta_i \sigma_i^2}{2\kappa_i}\left(1 - \exp\left(-\kappa_i \Delta t\right)\right)^2 + \frac{\sigma_i^2}{\kappa_i}\left(\exp\left(-\kappa_i \Delta t\right) - \exp\left(-2\kappa_i \Delta t\right)\right)\widehat{Y}_{it-1}. \tag{24}$$

For the $\mathrm{EA}_1(2)$ model the expression for the transition equation is modified (see Duffee 2002) as

$$\mathbf{Y}_t = \left(\mathbf{I} - \exp\left(-\mathbf{K}\tau\right)\right)\boldsymbol{\theta} + \exp\left(-\mathbf{K}\tau\right)\mathbf{Y}_{t-1} + \boldsymbol{\varepsilon}_t \tag{25}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{K}$ is a 2×2 matrix. The conditional variance for the $\mathrm{EA}_1(2)$ model is more complex than for the CIR model. The $\mathbf{Q}_t$ matrix is given by (Duffee 2002):

$$\mathbf{Q}_t = \mathbf{N}\mathbf{b}_0\mathbf{N}' + \sum_{i=1}^{N}\left(\sum_{j=1}^{N}\mathbf{N}\mathbf{b}_j\mathbf{N}'N_{i,j}^{-1}\right)Y_{i,t} \tag{26}$$

The specific forms of the different elements of equation (26) are given in Duffee (2002) and summarized in the appendix.

2.4.2 Measurement Equation

The second equation that is key to the Kalman filter methodology is the measurement equation which is identical for both the CIR and $\mathrm{EA}_1(2)$ model. It is derived from the relation expressed in equation (2). It describes the interaction between the observable variable, M-zero coupon yields $Z_t = (z_{1,}...z_M)$ with term to maturity $\tau = (\tau_1, ..., \tau_M)$ and the unobservable variable (the short rate). The measurement equation is obtained from the following relationship, where the subscript M denotes a specific maturity

$$Z_{M,t} = -\frac{\ln(P_M(t,\tau))}{\tau_M} = \sum_{i=1}^{2}\left\{-\frac{\ln A_i(\tau_M)}{\tau_M} + \frac{B_i(\tau_M)Y_{t,i}}{\tau_M}\right\} \tag{27}$$

where $P(t,\tau)$ is given by equation (12). From equation (27) the measurement equation is derived and it is given by

$$\mathbf{Z}_t = \widetilde{\mathbf{A}} + \widetilde{\mathbf{B}}\mathbf{Y}_t + \mathbf{V}_t \tag{28}$$

where $\mathbf{Z}_t$ represents a $M \times 1$ vector of the continuously compounded yields extracted from the yield curve. $\widetilde{\mathbf{A}}$ and $\widetilde{\mathbf{B}}$ correspond to the two last expressions in equation (27). The additional term, $\mathbf{V}_t$, is a M-vector of noise terms. This term reflects the assumption of small measurement errors in the bonds prices. For the purpose of this paper the error terms are assumed to be uncorrelated across different maturities following Chen and Scott (2003).

The noise term is assumed to follow $\mathbf{V}_t \sim N(0, \mathbf{R})$, where $\mathbf{R}$ is a diagonal $M \times M$ matrix where the variance of the measurement errors are denoted by r_M^2. These elements are part of the set of parameters that are estimated with the Kalman filter and the maximum likelihood method. The standard deviation of the errors are an indication of the goodness of fit of the model. The transition equation for six different maturities and two factors employed in this work is set out below:

$$\begin{bmatrix} Z_{t,1} \\ Z_{t,2} \\ Z_{t,3} \\ Z_{t,4} \\ Z_{t,5} \\ Z_{t,6} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{2} \frac{-\ln A_i(\tau_1)}{\tau_1} \\ \sum_{i=1}^{2} \frac{-\ln A_i(\tau_2)}{\tau_2} \\ \sum_{i=1}^{2} \frac{-\ln A_i(\tau_3)}{\tau_3} \\ \sum_{i=1}^{2} \frac{-\ln A_i(\tau_4)}{\tau_4} \\ \sum_{i=1}^{2} \frac{-\ln A_i(\tau_5)}{\tau_5} \\ \sum_{i=1}^{2} \frac{-\ln A_i(\tau_6)}{\tau_6} \end{bmatrix} + \begin{bmatrix} \frac{B_1(\tau_1)}{\tau_1} & \frac{B_2(\tau_1)}{\tau_1} \\ \frac{B_1(\tau_2)}{\tau_2} & \frac{B_2(\tau_2)}{\tau_2} \\ \frac{B_1(\tau_3)}{\tau_3} & \frac{B_2(\tau_3)}{\tau_3} \\ \frac{B_1(\tau_4)}{\tau_4} & \frac{B_2(\tau_4)}{\tau_4} \\ \frac{B_1(\tau_5)}{\tau_5} & \frac{B_2(\tau_5)}{\tau_5} \\ \frac{B_1(\tau_6)}{\tau_6} & \frac{B_2(\tau_6)}{\tau_6} \end{bmatrix} \begin{bmatrix} Y_{t,1} \\ Y_{t,2} \end{bmatrix} + \begin{bmatrix} V_{t,1} \\ V_{t,2} \\ V_{t,3} \\ V_{t,4} \\ V_{t,5} \\ V_{t,6} \end{bmatrix}. \tag{29}$$

Once the measurement and transition equations are established the standard Kalman Filter recursions are implemented to generate and estimate the state variables. For each point in time a measurement prediction error and a prediction error covariance matrix were estimated to construct the log-likelihood function. The optimization (given the restrictions in equation (17)) of the log-likelihood estimators were achieved using a Genetic Algorithim which has the advantage of not requiring initial values and is well suited in situations where search spaces are rough.

3 Empirical Analysis and Results

3.1 Data Description

The data used for this paper consists of weekly UK zero coupon yields from 12 February 1992 to 29 April 2009 shown in Figure 1. The data was obtained from the Bank of England's Statistics Section website[2].

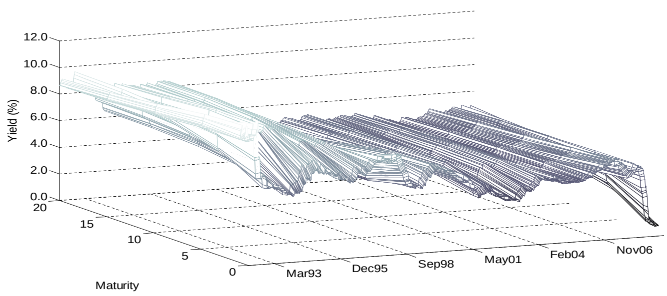

Figure 1: UK Term Structure (Feb 92 - Apr 09).

[2] http://www.bankofengland.co.uk/statistics/yieldcurve/index.htm

Following Nath and Nowman (2001) the observations were sampled on Wednesdays to avoid the weekday effects or missing data due to bank holidays. To characterize both the short and long end of the yield curve, six maturities were considered: 0.5, 1, 5, 10, 15 and 20 years. For each maturity there are 899 data points. For the six month maturity the complete data set was not available for few days previous to January 1997. Given the current methodology the Bank of England uses, the reason for unavailable data could be either there were no bonds with this maturity or the bonds were not liquid enough to be included in the Bank's yield estimation. Regression analysis was used to estimate the missing points (0.94% of the data). Table 1 shows the statistical summary of the data.

Maturity (years)	Mean (%)	Standard Deviation (%)
0.5	5.3028	1.4595
1	5.3447	1.4505
5	5.7586	1.5342
10	5.8432	1.5919
15	5.8248	1.6341
20	5.7573	1.6760

Table 1: UK Data Summary

3.2 Two Factor CIR and $EA_1(2)$ Model Results

The results of the parameter estimation for the two factor CIR model are presented in Table 2. As time is measured in years, all the parameter values are expressed on an annual basis. The variance parameters are both significant. The market price of risk parameters are both negative implying positive excess returns. Investors demand higher expected excess returns in compensation for holding extra risk, however neither of these parameters are statistically significant. As for $\widehat{\theta}$ and $\widehat{\kappa}$ both are not significant for the second factor as well as the values for $\widehat{r}_1$, $\widehat{r}_3$, $\widehat{r}_5$. These results are consistent with the results of Chen and Scott (2003).

The values of the standard deviation of the measurement errors (r_M) show that there is a better fit for the 15 year yield as its value is the lowest, 4 basis points. The largest is for the 5 year yield with 36 basis points. In respect to the specific value of the parameters the first factor is characterized by a very small long run mean and speed of mean reversion. On the contrary, the second factor has a mean level of 3.7% and a much faster speed. The volatility values are very similar just short of 7%. The predominance of one factor's speed has been found in previous studies as well (Chen and Scott 2003).

Parameters	θ_1	θ_2	σ_1	σ_2	κ_1	κ_2	λ_1
Estimate	0.0001	0.0376	0.0662	0.0690	0.0001	0.3641	-0.0471
Std Error	0.0472	0.0051	0.0118	0.0159	0.0675	0.0735	0.0591
Parameters	λ_2	r_1	r_2	r_3	r_4	r_5	r_6
Estimate	-0.0249	0.0019	0.0014	0.0036	0.0018	0.0004	0.0012
Std Error	0.0530	0.0023	0.00001	0.0092	0.0005	0.0010	0.0003

Table 2: Two Factor CIR Results

The results of the parameter estimation for the $EA_1(2)$ model are presented in Table 3. As discussed by Cheridito et al. (2007), the state variables in the $EA_1(2)$ are not easily interpreted. The intuitive interpretation of the CIR model is lost as the state variables are only indirectly related to the term structure through the bond pricing formula. The sum of the state variables do not constitute the short rate as in the CIR model but the sum of a constant and the scaled factors. However the standard deviation of the measurement errors are still a measure of relative goodness of fit. As shown in Table 3 these parameters are between less than 1 basis point (for both the 1 and 15 years yield) and 33 basis points. The estimates for the constant term of the instantaneous rate, $\widehat{\delta}_0$, is 1.79%, while the values of the coefficients $\widehat{\delta}_1$ and $\widehat{\delta}_2$ are quite small. All parameters are statistically significant except for $\widehat{\kappa}_{11}$, $\widehat{\kappa}_{21}$, $\widehat{r}_2$, $\widehat{r}_5$.

Parameters	Estimate	Std. Error	Parameter	Estimate	Std. Error
κ_{11}	0.0002	0.0003	$\lambda_{1(21)}$	0.0878	0.0002
κ_{21}	-0.0001	0.0042	$\lambda_{2(21)}$	-4.4497	0.0086
κ_{22}	0.2321	0.0035	$\lambda_{2(22)}$	0.0223	0.0035
θ_1	29.6943	0.7542	r_1	0.0022	4.08E-05
δ_0	0.0179	0.0003	r_2	1.74E-05	0.0001
δ_1	-0.0011	2.02E-05	r_3	0.0033	0.0001
δ_2	0.0011	3.29E-07	r_4	0.0016	2.21E-05
β_{12}	39.5291	0.0971	r_5	1.01E-05	3.81E-05
$\lambda_{1(11)}$	-0.0067	0.0004	r_6	0.0010	1.29E-06

Table 3: Parameter estimates for the $EA_1(2)$ Model

To further investigate the results and as a comparative metric, different errors were measured for each of the maturities; root mean square error (RMSE), mean error (ME) and mean absolute error (MAE). These errors were calculated comparing the filtered yields with the observed ones. The filtered yields were estimated using the parameters given by the maximum likelihood estimation and using the filtered factors in the measurement equation. The standard deviation and mean absolute deviation of the errors were also estimated and shown in Table 5.

	Two Factor CIR model		$EA_1(2)$ model	
Maturities (years)	MAE	RMSE	MAE	RMSE
0.5	0.0013	0.0017	0.0017	0.0022
1	0.0008	0.0013	0.0001	0.0005
5	0.0026	0.0036	0.0022	0.0032
10	0.0012	0.0018	0.0011	0.0016
15	0.0001	0.0002	1.66E-05	0.0001
20	0.0009	0.0011	0.0008	0.0010

Table 4: Mean Errors (in sample)

Table 4 shows in accordance to the standard deviation of the measurement errors, that the best fit in the Two Factor CIR model was for the 15 year yield. The RMSE is only 2 basis points while for the other maturities the errors are between 11 and 36 basis points. These errors are quite small and comparable with the results of Nath and Bowman (2001) and therefore the two factor CIR model can be said to be appropriate to fit the UK term structure for the period studied in this paper. As for $EA_1(2)$ model the RMSE showed the best fitted yield is also 15 year followed by the 1 year yield, while the others show errors between 10 and 32 basis points. Comparing the RMSE for both the models studied, the $EA_1(2)$ model shows slightly lower RMSE errors across most maturities. An average across maturities shows a difference of 2 basis points between the two models. Table 5 shows the standard deviation and the mean absolute deviation of the errors for both models. For most of the maturities the error standard deviation is lower for the $EA_1(2)$ model.

Although testing rigously the goodness of fit of any of the models used in this paper is not a trivial problem, a measure that can be used, in addition to the pricing errors of the yields used in the estimation, are the out of sample errors. The latter are given in Table 6 and support the good fit of both models. The out of sample maturities used are 2,4,7,9 12 and 18 years.

	Two Factor CIR model		$EA_1(2)$ model	
Maturities (years)	Std	MADev	Std	MADev
0.5	0.0016	0.0013	0.0022	0.0017
1	0.0013	0.0008	0.0005	0.0001
5	0.0034	0.0024	0.0031	0.0022
10	0.0016	0.0012	0.0016	0.0011
15	0.0002	0.0001	0.0001	3.24E-05
20	0.0011	0.0009	0.0010	0.0008

Table 5: Error statistics

	Two Factor CIR model		$EA_1(2)$ model	
Maturities (years)	MAE	RMSE	MAE	RMSE
2	0.0013	0.0018	0.0017	0.0025
4	0.0008	0.0012	2.30E-05	0.0003
7	0.0009	0.0013	0.0005	0.0008
9	0.0006	0.009	0.0004	0.0006
12	0.0001	0.0002	1.15E-05	0.0001
18	0.0010	0.0013	0.0008	0.0011

Table 6: Mean Errors (Out of sample)

For the out of sample comparison again the $EA_1(2)$ model is marginally better than the two factor CIR model. The RMSE are slightly lower for the $EA_1(2)$ model except for the two year maturity. The error statistics are also consistent with the previous result.

As a by-product of the Kalman Filter methodology a time series of the estimated state variables and the short rate are available. The estimated short rate (Figure 3) shows two main drops. The first corresponds to the effect of Black Wednesday, this corresponds to September 16, 1992. On that day the interest rates increased as a response to the increase by the government of the base interest rates from 10% to 12% and 15%, only to be lowered a day later back to 12%. This event affected mostly the short end of the curve, this increase-fall period is reflected mainly in the time series of the second factor. The other factor does not react significantly while the second one has a sharp drop. The second drop corresponds to the decrease in rates by the Bank of England in response to the current crisis. Once again it is the second factor that makes the adjustment.

The characteristics of the individual factors are aligned with the results

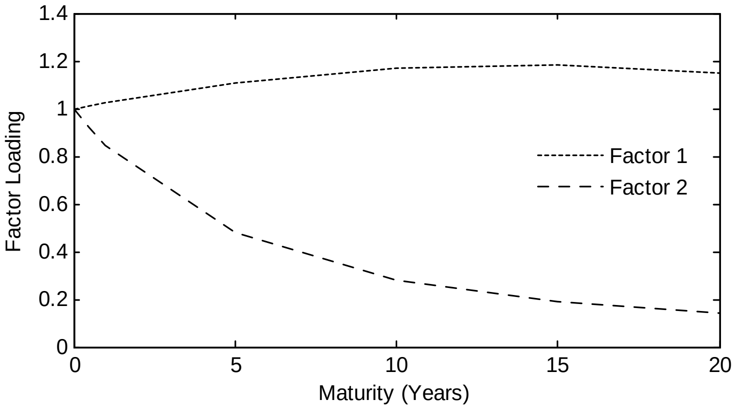

Figure 2: Factor Loadings (Two Factor CIR Model)

in Table 2. The first factor has a general trend to decrease towards a low mean while the second factor has a much faster speed and higher long run mean. An additional view of the two factors can be given by the factor loadings, the $\widetilde{\mathbf{B}}$ terms of the measurement equation (see equation (29)). As observed in other studies (Chen and Scott 2003, Geyer and Pichler 1999), the two factors correspond to a level and slope factors. The first factor slightly increases with time to maturity while for the second factor it is quite evident that the impact of this factor is greater for the short end of the curve as seen in Figure 2.

The short rate for the $EA_1(2)$ model can similarly be obtained given the parameters in Table 3 and the estimates of the two factors for the $EA_1(2)$ model. The filtered short rate given by the $EA_1(2)$ model follows a similar behaviour given by the two factor CIR model, the short rate decreases dramatically during Black Wednesday and again at the end of the period when the Bank of England decided to decrease the base rate to 0.5%, at this point the estimated short rate is 0.44%.

Figure 3 shows the comparison between the 6-month yield and the estimated short interest rate by both the two factor CIR and the $EA_1(2)$ model. The short rate estimated by both models follows closely the 6 month yield except for periods after sharp decreases in the 6 month yield such as after Black Wednesday and September 1998. It can be assumed that after extreme shocks the models take longer to adjust and therefore over or underestimate the short rate.

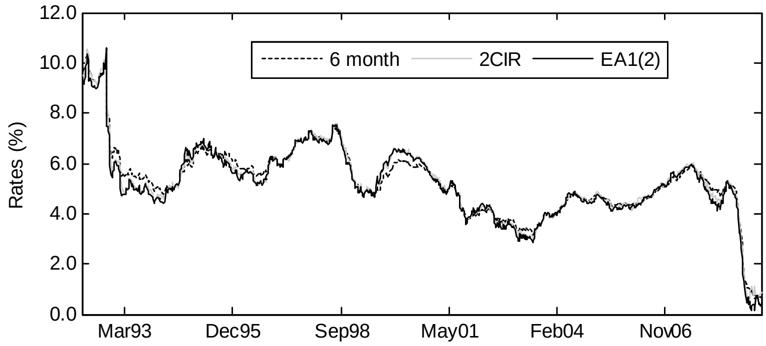

Figure 3: Comparison of 6-month yield and estimated short rate.

4 Conclusions

In this paper the UK term structure has been modelled using two factor CIR and essentially affine $EA_1(2)$ model. A panel approach was used to estimate the model parameters incorporating both the time series and cross sectional properties of the term structure. This approach was implemented using the Kalman Filter to construct a quasi maximum likelihood

The yield curve data fitted covers the period from Black Wednesday to the recent credit crisis, where yields have been as high as 10.5% and as low as 0.6%, the widest variation in recent times. It is demostrated that the more sophisticated $EA_1(2)$ model provides only a marginal improvement over the more tractable analytic two factor CIR model. Even though the essentially affine $EA_1(2)$ model achieved a higher loglikelihood the errors between the observed and estimated yields only favored the latter model on average by 2 basis points. The added flexibility given by the essentially affine specification is not substantially reflected in the fit of the UK term structure.

Both models are demonstrated to provide a good fit to the empirical term structure across all maturities as shown by the small magnitude of the root mean square errors (both in and out of sample) and standard deviation of the measurement errors.

5 References

Bolder D.J., 2001. Affine Term-Structure Models: Theory and Implementation. Bank of Canada Working Paper 2001-15.

Cairns, A.J.G, 2004. Interest Rate Models An Introduction. Princeton University Press.

Canabarro, E., 1995. Where do one factor interest rate models fail? The Journal of Fixed Income, September 1995, 31-52.

Chatterjee, R., 2004. Application of the Kalman Filter for Estimating Continuous Time Term Structure Models: Evidence from the UK and Germany. Society for Computational Economics, Computing in Economics and Finance, 346.

Chen, R., and Scott, L., 2003. Multi-factor Cox-Ingersoll-Ross models of the term-structure: estimates and tests from a Kalman Filter model. The Journal of Real Estate Finance and Economics, Vol 27, No.2, pp. 143-172.

Chen, R., and Scott, L., 1993. Maximum Likelihood Estimation for a Multifactor Equilibrium Model of the Term Structure of Interest Rates. Journal of Fixed Income, pp. 14 – 31.

Cheridito, P.,Filipovic, D., Kimmel, R., 2007. Market Price of Risk Specifications for Affine Models: Theory and Evidence. Journal of Financial Economics, 83, pp. 123-170.

Cox J.C., Ingersoll J.E. and Ross, S.A., 1985. A Theory of the Term Structure of Interest Rates. Econometrica, 53 pp., 385-407.

Dai, Q., and Singleton,K.J., 2000. Specification Analysis of Affine Term Structure Models. Journal of Finance, 55 pp, 1943-78.

De Jong, F. and Santa-Clara, P., 1999. The Dynamics of the Forward Interest Rate Curve: A Formulation with State Variables. The Journal of Financial and Quantitative Analysis, Vol. 34, No. 1., pp. 131-157.

De Jong, F., 1998. Time-Series and Cross-Section Information in Affine Term Structure Models. Journal of Business & Economic Statistics, Vol. 18, No. 3. pp. 300-314.

Driessen, J., 2005, Is Default Event Risk Priced in Corporate Bonds?, The Review of Financial Studies, Vol 18, No 1, 165-195.

Duan, J.C., and Simonato J.G., 1999. Estimating and Testing Exponential-Affine Term Structure Models by Kalman Filter, Review of Quantitative Finance and Accounting, 13, pp. 111-135.

Duffee, G., 2002. Term Premia and Interest Rate Forecasts in Affine Models. Journal of Finance, 57, 405-443.

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

Geyer, A. L. J., and Pichler, S., 1999. A State-Space Approach to Estimate and Test Multifactor Cox-Ingersoll-Ross Models of the Term Structure of Interest Rates. The Journal of Financial Research, Vol XXII, pp. 107-130.

Ho, T.S.Y, and Lee, S.B., 1986. Term Structure Movements and Pricing Interest Rate Contingent Claims. Journal of Finance, 41,1011-1029.

Hull, J and White, A., 1990. Pricing Interest Rate Derivative Securities. Review of financial Studies, 3,573-92.

Longstaff, F. A., and Schwartz, E. S., 1992. Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model. The Journal of Finance, XLVII, 1259-1282.

Lund, J., 1997. Econometric Analysis of Continuous-Time Arbitrage-Free models of the Term Structure of Interest Rates, Department of Finance, The Aarhus School of Business.

Nath, P. and Nowman,K.B., 2001. Estimates of the continuous time Cox-Ingersoll-Ross term structure model: further results for the UK gilt-edged market, Applied Economic Letters, 8, 85-88.

Nawalkha, S.K., Beliaeva, N.A. and Soto, G.M., 2007. Dynamic Term Structure Modeling. New Jersey: John Wiley & Sons, Inc.

Vasicek, O., 1977. An Equilibrium Characterization of the Term Structure, Journal of Financial Economics, 5, 177-188.

6 Appendix

6.1 Conditional Variance

The conditional variance for the $EA_1(2)$ model is derived by Duffee (2002) and its components are summarized for convenience in this appendix. The

$\mathbf{Q}_t$ matrix (conditional variance) for the Kalman Filter is as follows

$$\mathbf{Q}_t = \mathbf{N}\mathbf{b}_0\mathbf{N}' + \sum_{i=1}^{N}\left(\sum_{j=1}^{N}\mathbf{N}\mathbf{b}_j\mathbf{N}'N_{i,j}^{-1}\right)Y_{i,t} \tag{30}$$

where the matrix $\mathbf{N}$ is the results of the diagonalization of $\mathbf{K}$ as $\mathbf{K} = \mathbf{N}\mathbf{D}\mathbf{N}^{-1}$ with $\mathbf{D}$ as the diagonal matrix. The diagonal elements of $\mathbf{D}$ are denoted as $d_1, ..., d_N$. Before defining the other elements in the equation, new variables need to be defined

$$\begin{aligned}
\boldsymbol{\alpha}^* &= \boldsymbol{\alpha} \\
\boldsymbol{\theta}^* &= \mathbf{N}^{-1}\boldsymbol{\theta} \\
\boldsymbol{\Sigma}^* &= \mathbf{N}^{-1}\boldsymbol{\Sigma} \\
\boldsymbol{\beta}^* &= \boldsymbol{\beta}\mathbf{N}.
\end{aligned}$$

Given the new variables matrices $\mathbf{G}_0$ and $\mathbf{G}_i$ ($N \times N$ matrices) are constructed as

$$\begin{aligned}
\mathbf{G}_0 &= \boldsymbol{\Sigma}^*\mathbf{diag}\left(\boldsymbol{\alpha}^*\right)\boldsymbol{\Sigma}^{*\prime} \\
\mathbf{G}_i &= \boldsymbol{\Sigma}^*\mathbf{diag}\left(\boldsymbol{\beta}_i^*\right)\boldsymbol{\Sigma}^{*\prime}
\end{aligned}$$

where $\boldsymbol{\beta}_i^*$ is the i-th column of the matrix $\boldsymbol{\beta}^*$, $\mathbf{diag}(\boldsymbol{\alpha}^*)$ and $\mathbf{diag}\left(\boldsymbol{\beta}_i^*\right)$ represent a diagonal matrix with elements given by $\boldsymbol{\alpha}^*$and $\boldsymbol{\beta}_i^*$ respectively. Define $\mathbf{F}_0$, $\mathbf{F}_i$, and $\mathbf{H}_i$ $(for \quad i = 1, ..., N)$ as $N \times N$ matrices with typical elements (j, k), given by

$$\begin{aligned}
F_0^{(j,k)} &= \left(d_j + d_k\right)^{-1} G_0^{(j,k)} \left(1 - \exp\left(-\left(d_j + d_k\right)\Delta t\right)\right) && (31) \\
F_i^{(j,k)} &= \left(d_j + d_k\right)^{-1} G_i^{(j,k)} \left(1 - \exp\left(-\left(d_j + d_k\right)\Delta t\right)\right) && (32) \\
H_i^{(j,k)} &= \left(d_j + d_k - d_i\right)^{-1} G_i^{(j,k)} \left(\exp\left(-d_i\Delta t\right) - \exp\left(-\left(d_j + d_k\right)\Delta t\right)\right) && (33)
\end{aligned}$$

where Δt denotes the time step of the discretization ($\Delta t = 1/52$ for weekly data).Given the above definitions the elements of the $\mathbf{Q}$ matrix are constructed as follows

$$\begin{aligned}
\mathbf{b}_0 &= \mathbf{F}_0 + \sum_{i=1}^{N}\theta_i^*\left[\mathbf{F}_i - \mathbf{H}_i\right] && (34) \\
\mathbf{b}_j &= \mathbf{H}_j && (35)
\end{aligned}$$