The Estimation of the Heath-Jarrow-Morton Model by Use of Kalman Filtering Techniques

Ramaprasad Bhar
Carl Chiarella

ISSN: 1036-7373
THE ESTIMATION OF THE HEATH-JARROW-MORTON
MODEL BY USE OF KALMAN FILTERING TECHNIQUES

by

Ramaprasad Bhar and Carl Chiarella

School of Finance & Economics
University of Technology, Sydney
PO Box 123
Broadway, NSW 2007
Australia
PH: +61 2 330 3626
FAX: +61 2 330 3636

Abstract

A fairly flexible functional form for the forward rate volatility is applied in the Heath-Jarrow-Morton model of the term structure of interest rates to reduce the system dynamics to Markovian form. The resulting stochastic dynamic system is cast into a form suitable for estimation by use of nonlinear filter methodology. The technique is applied to 90 day bank bill and 3 year treasury bond data in the Australian market.
1. **Introduction**

The modern theory of the term structure of interest rates and the pricing of interest sensitive contingent claims involves expressing the bond price dynamics in terms of an equivalent probability measure which renders them arbitrage free. As a consequence both the drift and the diffusion coefficients of the stochastic differential equations describing bond price dynamics become functions of a maturity dependent volatility function. The genesis of this approach is the work of Heath, Jarrow and Morton (HJM, 1992a).

Analogously to the Black-Scholes equity option pricing model, the main input in the HJM approach is the entire term structure of the underlying asset and its volatility. Given the current term structure, the volatility function describes how it subsequently evolves over time. HJM (1992b) describe how this model can be used to price and hedge the entire interest derivative book of a financial institution and thus offer a consistent approach to the management of interest rate exposure. At the very heart of this technique is the choice of the volatility function. In the HJM context, the volatility function describes the fluctuations of each part of the term structure. Thus the modelling of the term structure of volatility becomes of crucial importance in modelling term structure of interest rates and in contingent claims pricing. Many of the forms of the volatility functions reported in the literature have been chosen for analytical convenience rather than being based on any empirical evidence. Apart from the study of HJM (1990), there has not been a great deal of empirical research into the appropriate form of the volatility function to be used in the arbitrage free class of models.

This paper seeks to address this gap. The parameters of a postulated functional form of the volatility function are estimated using the observed bond prices by applying Kalman filtering methodology. This requires transformation of the HJM model to finite-dimensional state space form. The complexity of this step is determined by the functional form of the volatility function and, in particular, whether it is dependent on the state variable.

The data sets used for the empirical study are the 90-day bill futures contracts trading on the Sydney Futures Exchange (SFE) and the prices of 3-year Australian Treasury bonds traded between banks. The former instrument, for all intents and purposes, is the coupon-free bond of the theoretical development and thus any errors introduced by coupon stripping procedures are avoided.

The remainder of the paper is organised as follows. In section 2, the essential framework of HJM model relevant to this study is reviewed. Section 3 reviews how the model may be expressed in finite dimensional state space form when certain functional forms for the volatility function are adopted. Section 4 shows how the state space form may be used to apply Kalman filtering methodology to estimate the parameters of the volatility function. In section 5, the data set is described in detail, whereas section 6 describes the empirical results. Section 7 concludes the paper.
2. The Basic Framework

Based on the general framework of HJM (1992a) we assume that the forward rate \( f(t, T) \) (the rate contracted at time \( t \) for instantaneous borrowing at time \( T \geq t \)) is driven by the one factor stochastic integral equation

\[
    f(t, T) - f(0, T) = \int_0^t \alpha(v, T) dv + \int_0^t \sigma(v, T) dW(v).
\]

Here \( \alpha(t, T), \sigma(t, T) \) are respectively the drift and diffusion coefficients at time \( t \) for an instrument maturing at time \( T \), \( f(0, T) \) is the initial forward rate curve (observed from the market determined yield curve) and \( dW(v) \) are the increments of the single Wiener process driving the stochastic fluctuations of the forward rate curve.

HJM show (see Bhar and Chiarella (1995) for a succinct summary) that if the economy is arbitrage free then the drift cannot be chosen independently of the volatility function. As a consequence the stochastic process for the forward rate becomes

\[
    f(t, T) = f(0, T) + \int_0^t \sigma(v, T) \int_0^v \sigma(y, y) dy dv + \int_0^t \sigma(v, T) d\tilde{W}(v),
\]

which may also be written as the stochastic differential equation

\[
    df(t, T) = \left( \sigma(t, T) \int_t^T \sigma(s, T) ds \right) dt + \sigma(t, T) d\tilde{W}(t),
\]

\[
    = \nu_f(t, T) dt + \sigma(t, T) d\tilde{W}(t).
\]

Here \( d\tilde{W}(v) \) are the increments of a Wiener process from the equivalent martingale probability measure which arises in the HJM theory by application of Girsanov's theorem.

The corresponding expression for the instantaneous spot rate of interest \( r(t) = f(t, t) \) is given by the stochastic integral equation

\[
    r(t) = f(0, t) + \int_0^t \sigma(v, t) \int_0^v \sigma(y, y) dy dv + \int_0^t \sigma(v, t) d\tilde{W}(v),
\]

or the equivalent stochastic differential equation
\[ dr(t) = \left[ f_2(0,t) + \frac{\partial}{\partial t} \int_0^t \sigma(v,t) \int_0^v \sigma(v,y) dy dy + \int_0^t \sigma_2(v,t) d\tilde{W}(v) \right] \]
\[ + \sigma(t,t) d\tilde{W}(t). \] (4)

Furthermore it can be shown by application of Ito's lemma that as a consequence, the price at time \( t \) of a pure discount bond maturing at time \( T \), \( P(t,T) \), is driven by the stochastic differential equation,

\[
\frac{dP(t,T)}{P(t,T)} = r(t) dt + \left[ - \int_t^T \sigma(t,v) dv \right] d\tilde{W}(t) \]
\[= r(t) dt + \sigma_b(t,T) d\tilde{W}(t). \] (5)

Note that \( B(t,T) = \ln P(t,T) \) by Ito's lemma satisfies the stochastic differential equation,

\[
dB(t,T) = \left[ r(t) - \frac{1}{2} \sigma_b^2(t,T) \right] dt + \sigma_b(t,T) d\tilde{W}(t) \]
\[= \mu_b(t,T) dt + \sigma_b(t,T) d\tilde{W}(t). \] (5')

The principle difficulty in application of the HJM model lies in the third term in the expression for the drift in equation (4'). This expression involves an integral over the history of the shock process up to time \( t \) and so makes the stochastic process driving bond prices non-Markovian. Most applications of the HJM approach seek some simplification which renders the driving dynamics Markovian.

3. Reduction to a Markovian System

It is shown in Bhar and Chiarella (1995) that by assuming for the forward rate volatility the specific functional form,

\[ \sigma(t,T) = [a_0 + a_1 (T-t) + \ldots + a_n (T-t)^n] e^{-\lambda(T-t)} \frac{G(r(t))}{\gamma}, \]

where \( G \) is a sufficiently well behaved function, the dynamics driving \( r(t) \) and \( B(t,T) \) turn out to be Markovian. However the dimension of this Markovian system increases rapidly with \( n \). Hence in this paper we consider the special case
\[ \sigma(t, T) = a_0 e^{-\lambda(T-t)} r(t)^\gamma, \] (6)

so that

\[ \sigma_B(t, T) = \frac{a_0}{\lambda} [e^{-\lambda(T-t)} - 1] r(t)^\gamma, \] (7)

where \( a_0, \lambda \) are parameters to be determined and \( \gamma \) will be set at 0.5 or 1 in the empirical analysis below.

It is also shown in Bhar and Chiarella (1995) that the non-Markovian stochastic dynamics system (4'), (5') for \( r(t), B(t, T) \), with \( \sigma(t, T) \) given by (6) can be expressed as the Markovian stochastic differential system

\[ dS(t) = [J(t) + H(S(t), t)] S(t) dt + V(S(t), t) d\tilde{W}(t), \] (8)

where

\[ S(t) = [B(t, T), r(t), \phi(t)]^T, \] (9a)

\[ J(t) = [0, f_2(0, t) + \lambda f_1(0, t), 0]^T, \] (9b)

\[ V(S(t), t) = [a_0 r(t)^\gamma (e^{-\lambda(T-t)} - 1)/\lambda, a_0 r(t)^\gamma, 0]^T, \] (9c)

\[ H(t) = \begin{bmatrix} 0 & h_{12} & 0 \\ 0 & -\lambda & a_0^2 \\ 0 & h_{32} & -2\lambda \end{bmatrix}, \] (9d)

with

\[ h_{12} = 1 - \frac{1}{2} r(t)^{2\gamma-1} [e^{-\lambda(T-t)} - 1]^2/\lambda^2, \]

\[ h_{32} = r(t)^{2\gamma-1}, \]

and
\[ \phi(t) = \int_0^t r(u)^2 e^{-2\lambda(t-u)} du. \] (9e)

The state variable \( \phi(t) \), which summarises characteristics of the path history of the instantaneous spot rate process is not observable. The instantaneous spot rate \( r(t) \) itself is also assumed not to be directly observable in this study. The shortest rate that could be used in Australia would be a 30-day rate. Data on overnight rates exists but are considered too "noisy" for reliable empirical work. So the only element of the state vector \( S(t) \) which is considered observable is the log bond price \( B(t, T) \), hence we have the observation vector (in this case a scalar)

\[ Y(t) = CS(t), \] (10)

where

\[ C = [1, 0, 0]. \]

It should be pointed out that similar reductions to a Markovian system have been found by Carverhill (1994) and Ritchken and Sankarasubramanian (1995).

4. **Nonlinear Filter Estimation**

Under the assumed specification of the volatility function, the problem of estimating the unknown parameters \( (a_0, \lambda) \) essentially implies solving the stochastic differential system represented by equation (8). Since, this system contains the unobserved variables \( r(t) \), \( \phi(t) \) and the only observed variable \( B(t, T) \) enters the system through the measurement equation (10), the approach to solving this problem is based on nonlinear filtering.

For ease of exposition, the equation (9) is expressed as,

\[ dS(t) = F(S(t); \theta) dt + V(S(t); \theta) d\tilde{W}(t), \] (11)

where

\[ \theta = [a_0, \lambda]. \]

In general, \( F(\cdot) \) and \( V(\cdot) \) will be non-linear in both the state variables as well as the parameters. Estimation of the parameter \( \theta \) in equation (11) will involve some form of discretisation from which the conditional moments over successive time intervals can be calculated. A number of approaches are possible and here we employ local linearisation and discretisation using the Euler-Maruyama and Milstein schemes.
4.1 Local Linearisation

The method of local linearisation is based upon first order Taylor series expansion about each observation. A version of the algorithm that we present here has been used by Ozaki (1992).

Assuming that the variable \( B(t, T) \) is observed at intervals \( t_i, i \in \{1, \ldots, n\} \), and concentrating on the sub-interval \( \delta_k = t_k - t_{k-1} \) and approximating the volatility vector over this sub-interval by its value at the beginning of the sub-interval, the following linear stochastic differential equation is obtained,

\[
\frac{dS(t)}{dt} = (a_k + A_k S(t)) dt + V_k d\bar{W}(t)
\]

where,

\[
A_k = \left\{ \frac{\partial F_i}{\partial S_j} \bigg| S_k \right\}, \quad i, j = 1, 2, \ldots , l
\]

(l is the length of the vector \( S(t) \))

\[
a_k = F_k - A_k S_k,
\]

\[
F_k = F(S(t_k); \theta), \quad S_k = S(t_k), \quad V_k = V(S_k, t_k).
\]

Since not all elements of \( S_k \) are directly observable we need to approximate \( S_k \) in the above equations by the most recent filter estimate. That is, in terms of the notation introduced below, we make the approximation

\[
S_k = \hat{S}_{k|k}.
\]

The crucial point here is that this approach leaves \( V_k \) constant within the sub-interval \( \delta_k \). In this sub-interval, since equation (12) is linear in \( S(t) \), the solution is given by,

\[
S_{k+1} = \exp [A_k \delta_k] S_k + \int_0^\delta \exp [A_k (\delta_k - \nu)] a_k d\nu + \int_0^\delta \exp [A_k (\delta_k - \nu)] V_k d\bar{W}(\nu).
\]

Therefore, the equations (10) and (16) represent the linear measurement equation and functionally non-linear state transition equation, respectively. By use of recursive definitions (e.g. in Harvey (1989)), it is now possible to generate the prediction error decomposition form of the likelihood function. Maximisation of this likelihood function will generate estimates of the parameter vector \( \theta \).

Since \( \bar{W} \) is a Wiener process, \( S_{k+1} \) has a conditionally normal distribution within \( \delta_k \).
and the first two conditional moments are,

\[ E(S_{k+1} \mid S_k) = \exp[A_k \delta_k] S_k + \int_0^\delta \exp[A_k(\delta_k - \nu)] a_k d\nu, \]  

(17)

\[ \text{Cov}[S_{k+1} \mid S_k] = \int_0^\delta \exp[A_k \nu] V_k V_k' \exp[A_k' \nu] d\nu. \]  

(18)

It should be noted that in this approach no measurement error is introduced, although it is possible to argue for a measurement error due to bid-ask spread or thin trading etc. Examination of this issue is left for later research.

From equation (17) the best forecast of \( S \) at \( t_{k+1} \) made at \( t_k \) (knowing \( Y_k \) at \( t_k \)) is,

\[ \hat{S}_{k+1 \mid k} = \exp[A_k \delta_k] \hat{S}_{k \mid k} + \int_0^\delta \exp[A_k(\delta_k - \nu)] a_k d\nu, \]  

(19)

and the best forecast of variance of \( S_{k+1} \) is,

\[ P_{k+1 \mid k} = \exp[A_k \delta_k] P_{k \mid k} \exp[A_k' \delta_k]' + Q_{k+1}, \]  

(20)

where, \( Q_{k+1} \) is given by (18). The estimation error is, therefore, given by

\[ Y_{k+1} = C \hat{S}_{k+1 \mid k}, \]  

(21)

and the variance of the estimation error is,

\[ v_{k+1} = CP_{k+1 \mid k} C^T. \]  

(22)

The updating equation for the state vector is

\[ \hat{S}_{k+1 \mid k+1} = \hat{S}_{k+1 \mid k} + K_{k+1} [Y_{k+1} - C \hat{S}_{k+1 \mid k}], \]  

(23)

where \( K_{k+1} \), the Kalman gain matrix, is given by

\[ K_{k+1} = P_{k+1 \mid k} C^T v_{k+1}^{-1}. \]  

(24)

(Note that \( v_{k+1} \) in this case is a scalar).

The recursion for the error covariance completes the specification of the Kalman filter updating equations,
\[ P_{k+1|k+1} = [I - K_{k+1} C] P_{k+1|k} [I - K_{k+1} C]^T. \] (25)

Under the assumption of normal distribution as incorporated in the equations (17) and (18), the transition probability density function for the state vector \( S_k \) to \( S_{k+1} \) can be written for a given set of observations \( T \) with the help of the updating equations (23) - (25). Following the argument in Harvey (1989) the prediction error decomposition form of the likelihood function is given by,

\[ \log L = -\frac{n}{2} \log 2\pi - \frac{1}{2} \sum_{k=1}^{n} \log |\nu_k| - \frac{1}{2} \sum_{k=1}^{n} e_k^2 \nu_k^{-1}, \] (26)

where, \( e_k = y_k - \hat{y}_{k|k-1} \).

To estimate the parameter vector \( \theta \) of the volatility function of the forward rates, the likelihood function in (26) can be maximised using a suitable numerical optimisation procedure. Before this can be implemented, a numerical technique to evaluate the expressions (19) - (22) has to be explored. This essentially means efficient computation of integrals of matrix exponentials. The method described by Van Loan (1978) has been adopted here.

Maximising \( L \) with respect to \( \theta \) yields consistent and asymptotically efficient estimators \( \hat{\theta} \). That is (see Lo (1988)) \( \overset{p}{\lim}_{n \to \infty} \hat{\theta} = \theta \) and

\[ \sqrt{n}(\hat{\theta} - \theta) \sim N(0, I^{-1}(\theta)), \] where the asymptotic covariance matrix \( I^{-1}(\theta) \) is the inverse of the information matrix \( I(\theta) \) given by

\[ I(\theta) = \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} E \left[ \frac{\partial^2}{\partial \theta \partial \theta'} \log L \right]. \]

4.2 Discretisation Using Euler-Maruyama and Mihelstein Schemes

Instead of employing local linearisation we may proceed directly and discretise equation (11) using any of a number of the schemes for the numerical solution of stochastic differential equations, outlined in Kloeden and Platen (1992). Here we shall use the Euler-Maruyama and Mihelstein strong schemes. Both of these schemes also involve approximating the volatility vector over each subinterval by its most recent filter estimate.

The discretisation for the Euler-Maruyama scheme is

where \( \hat{\xi}_k \sim N(0,1) \).
\[ S_{k+1} = S_k + F(S_k; \theta) \delta_k + V(S_k; \theta) \sqrt{\delta_k \zeta_k}, \]  

(27)

As in the case of local linearisation \( S_{k+1} \) is conditionally normal within \( \delta_k \) and the first two conditional moments are easily calculated to be

\[ E(S_{k+1} | S_k) = S_k + F(S_k; \theta) \delta_k, \]  

(28)

\[ \text{Cov}(S_{k+1} | S_k) = V(S_k; \theta) V^T(S_k; \theta) \delta_k. \]  

(29)

These latter two quantities are used in the Kalman filter updating equations (20) - (25) in place of those given by equations (17), (18) for the local linearisation case.

The Milstein scheme, which is an order 1.0 strong Taylor scheme (see Kloeden and Platen for details) discretises the stochastic differential equation (11) as

\[ S_{k+1} = [S_k + F(S_k; \theta) \delta_k] + V(S_k; \theta) \sqrt{\delta_k \zeta_k} \]

\[ + \frac{1}{2} V'(S_k; \theta) V(S_k; \theta) \delta_k [(\zeta_k)^2 - 1]. \]  

(30)

Here \( V' \) is the matrix of derivatives of the elements of the vector \( V \) with respect to elements of the state vector \( S \). Using \( V^{(i)} \) to denote the \( i \)th element of \( V \) then

\[
V' = \begin{bmatrix}
\frac{\partial V^{(1)}}{\partial B} & \frac{\partial V^{(1)}}{\partial r} & \frac{\partial V^{(1)}}{\partial \phi} \\
\frac{\partial V^{(2)}}{\partial B} & \frac{\partial V^{(2)}}{\partial r} & \frac{\partial V^{(2)}}{\partial \phi} \\
\frac{\partial V^{(3)}}{\partial B} & \frac{\partial V^{(3)}}{\partial r} & \frac{\partial V^{(3)}}{\partial \phi}
\end{bmatrix}
= \begin{bmatrix}
0 & a_0 (e^{-\lambda(T-n)} - 1) \gamma r^{\gamma-1}/\lambda & 0 \\
0 & 0 & a_0 \gamma r^{\gamma-1} \\
0 & 0 & 0
\end{bmatrix}.
\]  

(31)

From equation (30) we still obtain

\[ E(S_{k+1} | S_k) = S_k + F(S_k; \theta) \delta_k, \]

however the conditional covariance term now involves higher powers of \( \delta_k \) i.e.
\[ \text{Cov}(S_{k+1} | S_k) = V(S_k; \theta)V^T(S_k; \theta)\delta_k \]
\[ + \frac{1}{2}(V'VV^T + VV^TV^T)\delta_k^{3/2}E(\zeta_k^3) \]
\[ + \frac{1}{4}V'VV^TV'\delta_k^2(E(\zeta_k^4)-1). \]

(32)

The only change from use of the Euler-Maruyama scheme is the more precise estimate of \( \text{Cov}(S_{k+1} | S_k) \) in equation (32) which needs to be incorporated into the Kalman filter updating equations.

5. The Data Set

Our 90-day bank bill futures data is based on the Australian Bank Bill Market. The physical bank accepted bills market is an important short term financial market in Australia. These bills are issued on a discount basis and those for 90 days are the most actively traded. Periods of up to 180 days are also traded on the exchange. These bills have a face value of $500,000 and are usually traded in lots of $5 million. Discount rates are quoted as an index derived from (100 - yield p.a.) and the convention of 365 days in a year is used.

The futures market in bank bills started trading on the Sydney Futures Exchange (SFE) in 1979 and is the most important futures contract traded in the Australian market. The futures contract is for delivery of bills of face value $500,000 and of maturity of 85-95 days and the delivery months are March, June, September and December. Contracts for delivery for up to 36 months are not uncommon. Discount rates are also quoted as an index as for the physical bank bills. Settlement dates for these futures contracts are the second Friday of the delivery months, although trading terminates on Wednesday prior to this second Friday.

The data set used in this study consists of end-of-day bank bill futures prices reported by the SFE for the contracts deliverable in March quarter of 1991. There are 250 observations for this contract in the sample. For the purpose of this study all discount bank bill futures prices have been standardised to $1 of face value and the relationship between the yield and the price per dollar face value of these contracts is,

\[ \text{price} = \frac{1}{1 + \text{yield} \times \frac{90}{365}}. \]

In addition to this short-term contract, long-term (i.e. 3 year Australian Treasury
bonds) bond data\(^1\) has also been used in the estimation of the coefficients of the volatility function. Since these bonds are coupon paying bonds, a coupon stripping procedure outlined in Hunt (1994) is first applied to produce an equivalent zero coupon bond price series. Details of the coupon bond pricing mechanism are also described in that paper. This data set consists of weekly observations and thus there are 156 observations in this set.

6. **The Empirical Results**

We recall that the functional form assumed for \(\sigma(t, T)\) is that given in equation (6). Two different values of \(\gamma\) have been adopted in this study, namely \(\gamma = 0.5\) and \(\gamma = 1\).

The estimation process also requires specification of the initial forward rate function \(f(0, T)\). This is done using a polynomial fitting and is described in detail in Bhar and Hunt (1993). The initial value of the spot rate is obtained from \(f(0, T)\) by setting \(T=0\).

The maximisation of the forecast error decomposition form of the likelihood function is carried out in GAUSS\(^{TM}\) using the secant method BFGS and the variation of the golden section BRENT as the step search method.

Table 1 and the Figures 1 and 2 refer to the local linearisation scheme. Table 1 shows the estimated coefficients together with the value of the log likelihood functions. The standard errors are also reported however these should be treated with caution because of the small sample size. Below we discuss a Monte-Carlo bootstrap approach to an analysis of the statistical confidence intervals. The sample statistics of the implied instantaneous rate of interest are also included in that table\(^2\). These sample statistics for the 90-day Bill futures contracts for both values of \(\gamma\) are quite different but this is not so for the 3-year bond data. The implied mean of around 11\% p.a. is close to the mean of the 30-day rate observed during the sample period and suggests that a value of \(\gamma\) around 0.5 might be a reasonable fit for this parameter for 90-day bank bills futures market. For 3-year bonds both values of \(\gamma\) seem reasonable. However the question of the value of \(\gamma\) that fits the data best is left for future research. The plot of the estimated volatility functions for the forward rate as well as of the bonds are shown in Figures 1 and 2. As expected the bond price volatility falls towards zero with falling time to maturity. With respect to Figure 1 there is almost a doubling in the variability of bank-bill price volatility between the two values of \(\gamma\). The range for \(\gamma = 0.5\) is closer to the observed bond price volatility over this period (calculated on the basis of intraday high-low prices), again suggesting \(\gamma = 0.5\) might be a more appropriate value for the 90-day bank bill futures

---

\(^1\)The data was collected partially from the Australian Financial Review (which publishes a Reserve Bank of Australia collation of prices from bond market dealers) and partially from Australian Gilt Securities Pty Ltd.

\(^2\)These statistics are simply the mean and standard deviation of the time path for \(r(t)\) generated by the state equations with the estimated value of \(\theta\).
market. With respect to Figure 2 the variability of bond price volatility is not very marked between the two values of $\gamma$, again indicating that for the 3-year bond data both values of $\gamma$ seem reasonable.

Table 2 and Figures 3 and 4 refer to the Euler-Maruyama discretisation scheme. The contents in these tables and figures are similar to those for the local linearisation scheme and are self explanatory. Similarly, Table 3 and Figures 5 and 6 refer to the Mihlstein discretisation scheme. It should be noted that the estimated parameters are almost the same for these two discretisation approaches but differ quantitatively from those obtained from the local linearisation scheme. Note however that the sample statistics of the implied instantaneous rate of interest and the observations on the variability of bond price volatility are much the same as for the local linearisation case. The superiority of the statistical identification via these more sophisticated discretisation schemes is yet to be established.

In order to be able to comment about the statistical confidence intervals of the estimated parameters by the three different methods we employ the Monte-Carlo bootstrap technique to determine the distribution of the estimated parameters. All of these distributions have been calculated for the 3-year Australian Treasury bond data in the $\gamma = 0.5$ case.

The bootstrap samples are generated as follows:

(i) Using the original data set, the parameters are estimated as explained in the paper by maximising the prediction error decomposition form of the likelihood function.

(ii) The prediction error vector is stored.

(iii) The prediction error vector is sampled via uniformly distributed random numbers and added back to the original data to generate one instance of the bootstrap sample.

(iv) Using the simulated data obtained in the third step, the parameters are estimated as in step one.

(v) The steps three and four are repeated 1000 times and a histogram of the distribution of the estimated parameters is obtained.

In the Figures 7 - 8 the distributions of the parameters estimated using the local linearisation scheme are shown. The quantiles distribution are also given as a measure for the confidence interval. Figures 11, 12 give the corresponding distribution when Mihlstein's scheme is used. In both cases the estimated value of $a_0$ lies within

---

3The skewed distributions seen in the Figures 7, 8, 11 and 12 appear to confirm similar findings by Stoffer and Wall (1991) in their study to investigate the properties of parameter estimates from state-space models by Kalman filter.
the peak block of the distribution for \( a_0 \) whereas the estimated value of \( \lambda \) lies one standard deviation to the left of the peak block of the distribution for \( \lambda \).

On this basis we perhaps could claim that the estimates are statistically significant and that there seems little advantage in adopting the more sophisticated Mihlstein's scheme. However these Monte-Carlo bootstrap distributions for the estimates are subject to median bias as discussed by Effron (1987). This bias problem can be overcome by considering, amongst a number of other methods, the pivotal quantity. This quantity is in essence a distribution of the t-statistic of the estimate. For an estimate of a parameter \( \theta \) (= \( a_0 \) or \( \lambda \)) it is defined as

\[
\text{Pivotal Quantity (}\theta\text{)} = \text{PQ}(\theta) = (\theta^b - \theta)/\text{SE}(\theta^b)
\]

where

\[
\theta = \text{the postulated estimate (i.e. the estimates in Tables 1 or 3).}
\]
\[
\theta^b = \text{the mean of the bootstrap sampling distribution.}
\]

and

\[
\text{SE}(\theta^b) = \text{the standard error of the bootstrap sampling distribution.}
\]

The Figures 9 - 10 and 13-14 display the bootstrap distribution of the pivotal statistics as explained in the footnotes. Consider first the local linearisation scheme. Referring to Figure 7 we calculate PQ(\( a_0 \)) to be 0.185 which according to Figure 9 lies within the 95% confidence interval confirming the statistical significance of our estimate. Turning to Figure 8 we calculate that PQ(\( \lambda \)) is 33.8 which according to Figure 10 is in the 90% confidence interval.

Turning now to the Mihlstein scheme. From Figure 11 we calculate PQ(\( a_0 \)) to be 0.24 which according to Figure 13 is within the 95% confidence interval, whilst from figure 12 we calculate PQ(\( \lambda \)) to be -0.1 which according to figure 14 is in the 95% confidence interval.

Thus the pivotal quantity confirms the statistical significance of the estimates in the case \( \gamma = 0.5 \) for both the local linearisation and the Mihlstein method. Furthermore, on the basis of the results reported here there is nothing to distinguish between the different ways of discretising the stochastic differential equation system (11) at least as far as parameter estimation is concerned.

7. Conclusion

We have discussed how the assumption for the forward rate volatility of a certain deterministic function of time multiplied by a function of the spot rate allows the reduction of the (non-Markovian) dynamics of the Heath-Jarrow-Morton model to a Markovian system. We have expressed the resulting Markovian system in state-space form and shown how the problem of estimating the parameters of the forward rate
volatility function can be viewed as a nonlinear filtering problem. By using discretisations based on local linearisation, the Euler-Maruyama method and the Mihlstein scheme we have approached the estimation problem by means of the extended Kalman filter.

In this study we estimate a forward rate volatility function involving two parameters using 90 day bank bill futures and 3 year Australian Treasury bond data from the Australian markets.

In the case of the 3-year Australian Treasury bond data we have used Monte-Carlo bootstrap simulation of the parameter estimates as well as of their pivotal quantities to determine that the estimates are statistically significant for both the local linearisation method and Mihlstein schemes.

We conclude in favour of the viability the general approach which we advocate. Namely that of choosing forward rate volatility functions which allow a reduction of the Heath-Jarrow-Morton system to Markovian form and then approaching the parameter estimation problem by use of the extended Kalman filter.

Further research needs to consider a number of issues. Firstly extend the results to larger data sets to determine the stability of the parameter estimates. Secondly estimate the parameter $\gamma$ (see equation(6)) and examine its stability. Thirdly the significance of the incorporation of more maturity dependent terms in the expression for the forward rate volatility (i.e. more $a_j$ coefficients). Fourth recast the system dynamics in terms of market quoted periodic rates (weekly, monthly etc) rather than the idealised instantaneous rates of the theoretical development of HJM. This is possible by appropriate use of Ito's lemma. Fifth, use nonlinear filtering methods based on density approximation techniques as discussed for example by Tanizaki (1993) to overcome approximation errors and biases which may have been introduced by the various discretisation schemes considered here. Sixth, we have assumed for the forward rate volatility functional forms dependent upon the instantaneous spot rate itself. Unfortunately in this case reduction of the system dynamics to a Markovian system is not possible. It then becomes necessary to consider infinite dimensional filtering problems. The early work of Falb (1967) may provide a starting point to this difficult problem.
Table 1
Estimated Coefficients and Other Statistics
( Local Linearisation )

<table>
<thead>
<tr>
<th></th>
<th>a₀</th>
<th>λ</th>
<th>Log L</th>
<th>Implied Spot Rate</th>
<th>Mean</th>
<th>Stdv</th>
</tr>
</thead>
<tbody>
<tr>
<td>90-Day Bill</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(γ=0.5)</td>
<td>0.029779</td>
<td>0.025591</td>
<td>-1171.55</td>
<td>0.111994</td>
<td>0.00825</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.9244E-6)</td>
<td>(1.7133E-6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(γ=1.0)</td>
<td>0.186720</td>
<td>0.063397</td>
<td>-1304.54</td>
<td>0.173190</td>
<td>0.06221</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000252)</td>
<td>(0.013242)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-Year Bonds</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(γ=0.5)</td>
<td>0.003885</td>
<td>-0.047371</td>
<td>-891.10</td>
<td>0.111024</td>
<td>0.00827</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(6.7199E-6)</td>
<td>(0.002749)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(γ=1.0)</td>
<td>0.009993</td>
<td>-0.014385</td>
<td>-899.03</td>
<td>0.111596</td>
<td>0.00833</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.6540E-7)</td>
<td>(0.000856)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Number of data points in 90-Day Bill dataset is 250, and in 3-Year Bond dataset is 156. The numbers in parentheses under the coefficients represent standard errors. These result from the covariance matrix obtained by the inversion of the last Hessian computed during the optimisation process.
Figure 1
Volatility Function of Forward Rate and Bond Price
(Local Linearisation)

Volatility functions for the 90-Day Bill data
Figure 2
Volatility Function of Forward Rate and Bond Price
(Local Linearisation)

Volatility functions for the 3-Year Treasury Bond data
<table>
<thead>
<tr>
<th></th>
<th>$a_0$</th>
<th>$\lambda$</th>
<th>Log L</th>
<th>Implied Spot Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>90-Day Bill</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma = 0.5$</td>
<td>0.017189</td>
<td>0.014599</td>
<td>-1168.07</td>
<td>0.109635</td>
</tr>
<tr>
<td></td>
<td>(6.39E-07)</td>
<td>(0.000691)</td>
<td></td>
<td>0.00882</td>
</tr>
<tr>
<td>$\gamma = 1.0$</td>
<td>0.175446</td>
<td>0.065338</td>
<td>-1294.18</td>
<td>0.181372</td>
</tr>
<tr>
<td></td>
<td>(0.000162)</td>
<td>(0.010451)</td>
<td></td>
<td>0.06854</td>
</tr>
<tr>
<td><strong>3-Year Bonds</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma = 0.5$</td>
<td>0.003937</td>
<td>-0.047981</td>
<td>-891.38</td>
<td>0.110250</td>
</tr>
<tr>
<td></td>
<td>(6.73E-07)</td>
<td>(0.002789)</td>
<td></td>
<td>0.01208</td>
</tr>
<tr>
<td>$\gamma = 1.0$</td>
<td>0.010162</td>
<td>-0.014577</td>
<td>-899.89</td>
<td>0.110691</td>
</tr>
<tr>
<td></td>
<td>(0.003918)</td>
<td>(0.005856)</td>
<td></td>
<td>0.01211</td>
</tr>
</tbody>
</table>

Number of data points in 90-Day Bill dataset is 250, and in 3-Year Bond dataset is 156. The numbers in parentheses under the coefficients represent standard errors. These result from the covariance matrix obtained by the inversion of the last Hessian computed during the optimisation process.
Volatility Function of Forward Rate and Bond Price
(Euler-Maruyama Discretisation)

Volatility functions for the 90-Day Bill data
Figure 4
Volatility Function of Forward Rate and Bond Price
(Euler-Maruyama Discretisation)

Volatility functions for the 3-Year Treasury Bond data
<table>
<thead>
<tr>
<th></th>
<th>$a_0$</th>
<th>$\lambda$</th>
<th>Log L</th>
<th>Implied Spot Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td><strong>90-Day Bill</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>($\gamma=0.5$)</td>
<td>0.017189</td>
<td>0.014599</td>
<td>-1168.07</td>
<td>0.110073</td>
</tr>
<tr>
<td></td>
<td>(4.69E-06)</td>
<td>(0.000687)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>($\gamma=1.0$)</td>
<td>0.175446</td>
<td>0.065338</td>
<td>-1294.18</td>
<td>0.182097</td>
</tr>
<tr>
<td></td>
<td>(2.50E-06)</td>
<td>(0.010265)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>3-Year Bonds</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>($\gamma=0.5$)</td>
<td>0.003937</td>
<td>-0.047981</td>
<td>-891.38</td>
<td>0.111405</td>
</tr>
<tr>
<td></td>
<td>(1.23E-06)</td>
<td>(0.002790)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>($\gamma=1.0$)</td>
<td>0.010162</td>
<td>-0.014577</td>
<td>-899.89</td>
<td>0.110962</td>
</tr>
<tr>
<td></td>
<td>(1.88E-08)</td>
<td>(0.000872)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Number of data points in 90-Day Bill dataset is 250, and in 3-Year Bond dataset is 156. The numbers in parentheses under the coefficients represent standard errors. These result from the covariance matrix obtained by the inversion of the last Hessian computed during the optimisation process.
Figure 5
Volatility Function of Forward Rate and Bond Price
(Mihistein’s Discretisation)

Volatility functions for the 90-Day Bill data
Figure 6
Volatility Function of Forward Rate and Bond Price
(Mihlstein's Discretisation)

Volatility functions for the 3-Year Treasury Bond data
**Figure 7**

**Bootstrap Sampling Distribution (a,)**

**(Local Linearisation - γ = 0.5)**

<table>
<thead>
<tr>
<th>Summary Statistics</th>
<th>Quantiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.00681</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.01625</td>
</tr>
<tr>
<td>Skewness</td>
<td>4.38410</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>52.45060</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Frequency distribution of the parameter from 1000 bootstrap resampling using 3-year Treasury bond data. Highly nonuniform distribution indicating median bias. Possible improvements may be achieved by examining the distribution of a pivotal statistic e.g. bootstrap-t and/or applying bias correction due to Efron (1987).
Figure 8

Bootstrap Sampling Distribution (λ)
(Local Linearisation - γ = 0.5)

Summary Statistics                      Quantiles

Mean          -0.05898                  0.025                  -0.10785
Standard Deviation  0.12875         0.050                  -0.08832
Skewness        -7.16174                 0.100                 -0.07885
Kurtosis        52.12661                 0.500                 -0.04630
                 0.900                   -0.00498
                 0.950                   0.00769
                 0.975                   0.01060

Frequency distribution of the parameter from 1000 bootstrap resampling using 3-year Treasury bond data. Highly nonuniform distribution indicating median bias. Possible improvements may be achieved by examining the distribution of a pivotal statistic e.g. bootstrap-t and/or applying bias correction due to Efron (1987).
Figure 9
Bootstrap Sampling Distribution (Pivotal Quantity)
(Local Linearisation)

Quantiles (* E+04)

<table>
<thead>
<tr>
<th>Quantile</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.025</td>
<td>-6.48</td>
</tr>
<tr>
<td>0.050</td>
<td>-2.70</td>
</tr>
<tr>
<td>0.100</td>
<td>-1.34</td>
</tr>
<tr>
<td>0.500</td>
<td>0.56</td>
</tr>
<tr>
<td>0.900</td>
<td>3.91</td>
</tr>
<tr>
<td>0.950</td>
<td>8.39</td>
</tr>
<tr>
<td>0.975</td>
<td>10.14</td>
</tr>
</tbody>
</table>

Frequency distribution of the pivotal quantity defined below from 1000 bootstrap resampling using 3-year Treasury bond data.

Pivotal quantity = Bootstrap-t = \( \frac{a^b - a_0}{SE(a^b)} \)

The superscript B denotes bootstrap estimate and SE stands for standard error.
Figure 10
Bootstrap Sampling Distribution (Pivotal Quantity)
(Local Linearisation)

Quantiles

<table>
<thead>
<tr>
<th>Probability</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.025</td>
<td>-8.50</td>
</tr>
<tr>
<td>0.050</td>
<td>-7.77</td>
</tr>
<tr>
<td>0.100</td>
<td>-6.62</td>
</tr>
<tr>
<td>0.500</td>
<td>0.40</td>
</tr>
<tr>
<td>0.900</td>
<td>87.00</td>
</tr>
<tr>
<td>0.950</td>
<td>127.00</td>
</tr>
<tr>
<td>0.975</td>
<td>176.00</td>
</tr>
</tbody>
</table>

Frequency distribution of the pivotal quantity defined below from 1000 bootstrap resampling using 3-year Treasury bond data.

Pivotal quantity = Bootstrap-t = \( \frac{(\lambda^B - \lambda)}{SE(\lambda^B)} \)

The superscript B denotes bootstrap estimate and SE stands for standard error.
Figure 11
Bootstrap Sampling Distribution ($a_n$)
(Mihlstein's Scheme)

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Range of Standard Deviations</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8</td>
<td>-5</td>
</tr>
<tr>
<td>0.6</td>
<td>-4</td>
</tr>
<tr>
<td>0.4</td>
<td>-3</td>
</tr>
<tr>
<td>0.2</td>
<td>-2</td>
</tr>
<tr>
<td>0.0</td>
<td>-1</td>
</tr>
<tr>
<td>0.0</td>
<td>0</td>
</tr>
<tr>
<td>0.2</td>
<td>1</td>
</tr>
<tr>
<td>0.4</td>
<td>2</td>
</tr>
<tr>
<td>0.6</td>
<td>3</td>
</tr>
<tr>
<td>0.8</td>
<td>4</td>
</tr>
<tr>
<td>0.0</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Summary Statistics</th>
<th>Quantiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.00772</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.01669</td>
</tr>
<tr>
<td>Skewness</td>
<td>6.81198</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>48.24813</td>
</tr>
</tbody>
</table>

Frequency distribution of the parameter from 1000 bootstrap resampling using 3-year Treasury bond data. Highly nonuniform distribution indicating median bias. Possible improvements may be achieved by examining the distribution of a pivotal statistic e.g. bootstrap-t and/or applying bias correction due to Efron (1987).
Figure 12
Bootstrap Sampling Distribution ($\lambda$)
(Mihlstein's Scheme)

<table>
<thead>
<tr>
<th>Summary Statistics</th>
<th>Quantiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-0.06087</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.13257</td>
</tr>
<tr>
<td>Skewness</td>
<td>-6.94590</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>48.86460</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Frequency distribution of the parameter from 1000 bootstrap resampling using 3-year Treasury bond data. Highly nonuniform distribution indicating median bias. Possible improvements may be achieved by examining the distribution of a pivotal statistic e.g. bootstrap-$t$ and/or applying bias correction due to Efron (1987).
Figure 13

Bootstrap Sampling Distribution (Pivotal Quantity)
(Mihlstein’s Scheme)

Quantiles (* E+04)

<table>
<thead>
<tr>
<th>Value</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.025</td>
<td>-5.13</td>
</tr>
<tr>
<td>0.050</td>
<td>-1.73</td>
</tr>
<tr>
<td>0.100</td>
<td>-0.72</td>
</tr>
<tr>
<td>0.500</td>
<td>0.52</td>
</tr>
<tr>
<td>0.900</td>
<td>3.67</td>
</tr>
<tr>
<td>0.950</td>
<td>8.03</td>
</tr>
<tr>
<td>0.975</td>
<td>10.95</td>
</tr>
</tbody>
</table>

Frequency distribution of the pivotal quantity defined below from 1000 bootstrap resampling using 3-year Treasury bond data.

Pivotal quantity = Bootstrap-t = \( (a_r^B - a_o) / SE(a_r^B) \)

The superscript B denotes bootstrap estimate and SE stands for standard error.
Figure 14
Bootstrap Sampling Distribution (Pivotal Quantity)
(Mihlstein's Scheme)

Quantiles

| 0.025 | -8.46 |
| 0.050 | -7.72 |
| 0.100 | -6.68 |
| 0.500 | 0.29  |
| 0.900 | 47.01 |
| 0.950 | 94.26 |
| 0.975 | 118.46 |

Frequency distribution of the pivotal quantity defined below from 1000 bootstrap resampling using 3-year Treasury bond data.

Pivotal quantity = Bootstrap-t = \( \frac{\lambda^B - \lambda}{SE(\lambda^B)} \)

The superscript B denotes bootstrap estimate and SE stands for standard error.
References


