A MAXIMUM LIKELIHOOD APPROACH TO ESTIMATION OF HEATH-JARROW-MORTON MODELS

RAMAPRASAD BHAR[1], CARL CHIARELLA[2], AND THUY-DUONG TÔ[3]

[1]*School of Banking and Finance*
University of New South Wales
Sydney NSW2052
Australia
R.Bhar@unsw.edu.au

[2,3]*School of Finance and Economics*
University of Technology, Sydney
PO Box 123, Broadway NSW2007
Australia
[2]Carl.Chiarella@uts.edu.au
[3]Thuy.D.To@uts.edu.au

ABSTRACT. Research on the Heath-Jarrow-Morton (1992) term structure models so far has focused on the class having time-deterministic instantaneous forward rate volatility. In this case the forward rate is Markovian, even if the spot rate process is not. However, this Markovian feature can only be used under the historical measure, involving two unsatisfactory assumptions: one on market price of risk, usually made for pure mathematical tractability, the other to use futures yields as a proxy for the instantaneous forward rate, which may results in estimation bias. This paper circumvents both of these assumptions. First, the bias is quantified and shown to be non-negligible. Then futures contracts are treated as derivative instruments written on forward rates to derive the full information maximum likelihood estimator for observable futures prices, using both time series and cross-sectional data, without the need to assume and estimate any functional forms for the market price of interest rate risk. The derivation involves the likelihood transformation method of Duan (1994). The method is then applied to the estimation of a humped forward rate volatility model for Eurodollar futures series traded on the Chicago Mercantile Exchange.

Key words: Term structure; Heath-Jarrow-Morton; Time-deterministic forward volatility; Humped forward volatility model; Full information maximum likelihood

JEL classifications: C51, E43, G12, G13

1. INTRODUCTION

Interest rate modelling has long been of interest to researchers and practitioners. The arbitrage-free approach to modelling the term structure of interest rates has its origin in Ho and Lee (1986), and is most clearly articulated in Heath, Jarrow, Morton (1992), (hereafter HJM). The model is based on the specification of the term structure of forward rates in terms of the initial forward rate curve and the forward rate volatility. The condition that rules out arbitrage opportunities determines uniquely the drift of the instantaneous forward rate from the forward rate volatility via the market price of interest rate risk. Under the equivalent measure, this market price of risk disappears. The dynamics of the instantaneous spot rate is then developed from the forward rate evolution.

The HJM approach, therefore, has many advantages over the earlier approaches such as Vasicek (1977), Brennan and Schwartz ((1979, 1982)), Cox et al. (1985). First, the model matches the current term structure by construction. Second, there is no need for any assumptions on investors' preferences. Third, the model offers a parsimonious representation of the market dynamics and requires only specification of the form of the forward rate volatility function. Despite these advantages, there have been very few empirical studies of the HJM model. This is due to the fact that in its most general form, the resulting instantaneous spot rate evolution is not path-independent, ie. it is non-Markovian, and the entire history of the term structure has to be carried forward, thus increasing the computational complexity.

In one approach to the empirical study of the HJM model, researchers have relied on implied volatility, most notably Amin and Morton (1994) and Amin and Ng (1997). Under this approach, each day, the volatility parameters are backed out from market prices of derivative instruments, for example, by finding the set of parameters that minimizes the sum of squared errors. The implied volatility approach gives estimates of the model parameters that change every day. This approach is useful from the perspective of market practitioners who need to calibrate the model daily to prevailing market conditions in order to ensure accurate pricing and hedging strategies.

The focus of this paper will rather be on estimation of the (fixed) parameters of a volatility specification across an estimation period, for example to find the "best" from a family of possible volatility specifications. The resulting functional forms could of course then be used by market practitioners in their calibration procedures.

The approach to estimation so far adopted relies on reducing the system to Markovian form under some particular functional specification of the forward rate volatility. Theoretical work on reduction-to-Markovian form can be found in Bjork and Svensson (2001), Bliss and Ritchken (1996), Bhar and Chiarella (1997*b*), Chiarella and Kwon (2001a, 2001b), De Jong and Santa-Clara (1999), Inui and Kijima (1998), Ritchken and Sankarasubramanian (1995). Within these classes of models, empirical work lags behind the cited theoretical developments.

The HJM class that has time-deterministic instantaneous forward rate volatility is regarded as relatively easy to implement. This is because the instantaneous forward rate is Markovian, and therefore there is no apparent need to Markovianize the spot rate of interest[1]. However, if this Markovian forward rate dynamics is used directly in estimation, there are two main issues to consider: how to handle the market price of risk and the approximation problem.

The market price of risk is a difficult quantity to work with. Empirical studies so far have relied on analytical formulae of market variables (e.g. bond price or futures price formulae) so that there is no need to consider the market price of risk in estimation. There is only a limited number of volatility specifications that allows such analytical results. Examples include the constant volatility model in Flesaker (1993), the exponential volatility model and exponential square-root model[2] in Bhar and Chiarella (1997*a*) and Raj et al. (1997). In general, one will have to work with the forward rate in the historical measure, and have to assume a functional form for the market price of risk, which is taken for mathematical tractability rather than any economic justification. For example, De Jong and Santa-Clara (1999) assume that the market price of risk is proportional to the square root of the spot rate in order to estimate one affine representation of their class of models, whereas Bühler et al. (1999) do not consider the link between drift and volatility at all when they attempt the Maximum Likelihood Approach to test models where the instantaneous volatility is either constant or a linear function of the forward rate.

The second issue concerns the un-observability of the instantaneous forward rate, since using a fixed-maturity futures rate as a proxy for it may result in estimation bias. It turns out that for the class of HJM models where forward volatility function is time deterministic, the evolution of the futures price can be derived from the forward rate evolution. In section 2, the bias due to using fixed-maturity futures yields as an approximation to instantaneous forward rates is quantified. In particular, the bias is decomposed into two components, maturity bias and convexity bias. The maturity bias arises from approximating an instantaneous forward rate by a fixed-maturity forward rate, and is negligible if the fixed-maturity is small. The convexity bias, which is not negligible, arises from using a fixed-maturity futures yield to approximate the fixed-maturity forward rate.

This paper takes advantage of the link between forward and futures rate evolution (due to the time deterministic forward rate volatility specification) to derive the exact likelihood function for the time series of futures prices observed in the market,

[1]It should be noted that although there are important classes of HJM with time-deterministic forward volatility whose spot rate can be placed into a Markovianized system, the question of whether *all* HJM models with time-deterministic volatility have this property is still an open question for research.

[2]In the exponential square-root model, the instantaneous forward volatility is the product of an exponential function of time to maturity and the square-root of instantaneous spot rate of interest.

rather than treating the short maturity futures rate as a proxy for the instantaneous forward rate. A similar approach has been used by Pearson and Sun (1994) in estimating the Cox, Ingersoll, Ross model, and by Ho et al. (2001) in estimating the one factor HJM model with exponential forward volatility function. These studies rely on the closed-form solution for bond prices and futures prices to estimate the unobservable instantaneous spot rate and forward rate respectively. The key advance in our approach is that we recognize the observable futures rate as a derivative instrument driven by the same source of uncertainty as that driving the underlying unobservable forward rate. Therefore, despite the fact that we cannot establish a closed-form formula for the futures price on forward rate, we are able to derive the exact likelihood function for all model specifications that have deterministic volatility forms, albeit the likelihood will be different in its degree of complexity.

The major contribution of our paper, as a consequence, is a systematic method to estimate a rich class of HJM models, where the forward volatility function is time deterministic, and the spot rate may or may not be Markovian, without the need to assume and estimate any functional forms for the market price of interest rate risk. An additional important improvement in our estimation approach is that we recognize that futures prices are less than perfectly correlated with each other under the stochastic setting. Therefore, we apply the full information maximum likelihood method to pooled time series and cross-sectional futures price data to estimate our model. By incorporating cross-sectional data, we can exploit the full information content along a yield curve.

The paper is organized as follows. Section 2 reviews the HJM model, discusses the futures rate evolution given the forward rate evolution where forward rate volatility is deterministic. This section will also discuss the bias in using futures rates to approximate forward rates. Section 3 then presents the likelihood transformation method, utilizing the results of Duan (1994) to simplify the likelihood calculation. The full information likelihood is derived by transforming market variables to state variables whose density can be found by analytically solving the Kolmogorov partial differential equation, subject to appropriate boundary conditions, as proposed by Lo (1988). Data and models considered are described in section 4. We discuss the parameter estimate in section 5. Section 6 concludes the paper.

2. Forward and Futures Link within HJM framework

Under the time deterministic forward rate volatility, HJM model assumes that instantaneous T-maturity forward rate $f(t,T)$ (for $t \leq T \in \mathbb{R}^+$) evolves according to

$$f(t,T) = f(0,T) + \int_0^t \mu(u,T,\cdot)du + \sum_{i=1}^{I} \int_0^t \sigma_i(u,T)dW_i(u), \tag{2.1}$$

where the $W_i(t)$ are standard Wiener processes under the historical measure $\mathcal{Q}$, and $\mu(t,T,\cdot)$ and the $\sigma_i(t,T)$ are respectively the drift and the set of diffusion coefficients for the instantaneous forward rate to maturity T.

HJM show that the elimination of arbitrage opportunities implies that the drift is uniquely determined by the volatility function via the market prices of interest rate risk $\phi_i(t)$ according to

$$\mu(t,T,\cdot) = -\sum_i \sigma_i(t,T)\left[\phi_i(t) - \int_t^T \sigma_i(t,s)ds\right]. \tag{2.2}$$

The forward rate evolution can then be described under the equivalent measure $\widetilde{\mathcal{Q}}$, where the market price of risk is absorbed into the Wiener process under $\widetilde{\mathcal{Q}}$, as the stochastic integral equation

$$f(t,T) = f(0,T) + \sum_i \left[\int_0^t \sigma_i(u,T)\int_u^T \sigma_i(u,s)ds\,du + \int_0^t \sigma_i(u,T)d\widetilde{W}_i(u)\right], \tag{2.3}$$

or in the more familiar form of a stochastic differential equation as

$$df(t,T) = \sum_i \left[\sigma_i(t,T)\int_t^T \sigma_i(t,s)ds + \sigma_i(t,T)d\widetilde{W}(t)\right]. \tag{2.4}$$

The evolution of the instantaneous spot rate of interest can be derived accordingly from (2.3) by setting $T = t$, thus,

$$r(t) = f(0,t) + \sum_i \left[\int_0^t \sigma_i(u,t)\int_u^t \sigma_i(u,s)ds\,du + \int_0^t \sigma_i(u,t)d\widetilde{W}_i(u)\right].$$

The corresponding stochastic differential equation for the instantaneous spot rate of interest is

$$dr(t) = \left[f_2(0,t) + \sum_i \left(\frac{\partial}{\partial t}\int_0^t \sigma_i(u,t)\int_u^t \sigma_i(u,s)ds\,du + \int_0^t \frac{\partial \sigma_i(u,t)}{\partial t}d\widetilde{W}_i(u)\right)\right]dt + \sum_i \sigma_i(t,t)d\widetilde{W}_i(t). \tag{2.5}$$

Any derivative instruments can then be priced under the risk neutral measure. A futures contract is a derivative instrument written on a bond, and therefore, its price today is just the expectation of the future payoff under the risk neutral measure.

Let $F(t, T_F, T_B)$ be the price at time t of a futures contract maturing at time $T_F(> t)$. The contract is written on a pure discount instrument which has a face value of \$1 and matures at time $T_B(> T_F)$.

Proposition 2.1. *The evolution of $F(t, T_F, T_B)$ is given by the stochastic integral equation*

$$F(t, T_F, T_B) = F(0, T_F, T_B) \exp\left[-\frac{1}{2} \sum_i \int_0^t \left(\int_{T_F}^{T_B} \sigma_i(u, s) ds \right)^2 du - \sum_i \int_0^t \left(\int_{T_F}^{T_B} \sigma_i(u, s) ds \right) d\widetilde{W}_i(u) \right],$$

or equivalently by the stochastic differential equation

$$\frac{dF(t, T_F, T_B)}{F(t, T_F, T_B)} = -\sum_i \left(\int_{T_F}^{T_B} \sigma_i(t, s) ds \right) d\widetilde{W}_i(t). \tag{2.6}$$

Proof. The proof is a straight forward extension of the model in Musiela et al. (1992) where only one noise term is considered. Details can be found in Appendix A. □

Let $y(t, T_F, T_B)$ be the corresponding "futures yield", ie. the quantity defined according to

$$F(t, T_F, T_B) = \frac{1}{1 + y(t, T_F, T_B)(T_B - T_F)}. \tag{2.7}$$

Application of Itô lemma gives

$$dy(t, T_F, T_B) = \sum_i \left(\int_{T_F}^{T_B} \sigma_i(t, s) ds \right)^2 \left(\frac{1}{T_B - T_F} + y(t, T_F, T_B) \right) dt + \sum_i \left(\int_{T_F}^{T_B} \sigma_i(t, s) ds \right) \left(\frac{1}{T_B - T_F} + y(t, T_F, T_B) \right) d\widetilde{W}_i(t). \tag{2.8}$$

Under the equivalent measure $\widetilde{\mathcal{Q}}$, the forward rate $f(t, T_F, T_B)$ is distributed normally, whereas the futures yield $y(t, T_F, T_B)$ is not distributed normally. The resulting variances of the two processes are different, depending on the maturity of the futures contract (ie. $T_B - T_F$) and the specification of the volatility function. Since the variance structure is preserved under the transformation from the historical measure to the equivalent measure (see Figure 1), using futures rates as a proxy for forward rates (under the historical measure) will impose a wrong variance on the distribution, and therefore, distort the estimation results.

FIGURE 1. The change of measure

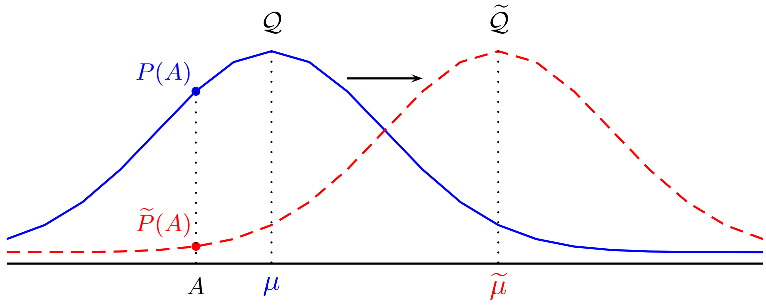
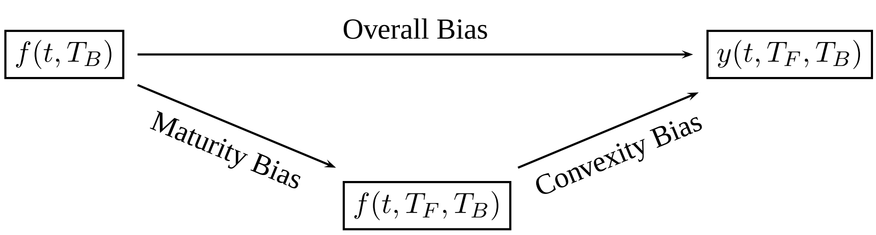


FIGURE 2. Bias Decomposition



To be more precise, from (2.3), the variance of the instantaneous forward rate is

$$\operatorname{var}\big(f(t, T_B)\big) = \sum_i \int_{t_0}^{t} \sigma_i^2(u, T_B)du, \tag{2.9}$$

whereas the variance of the fixed-maturity futures rate is (see Appendix B)

$$\begin{aligned} \operatorname{var}\big(y(t, T_F, T_B)\big) = & \left(\frac{1}{T_B - T_F} + y(0, T_F, T_B)\right)^2 \\ & \times \exp\left(2\sum_i \int_{t_0}^{t} \left(\int_{T_F}^{T_B} \sigma_i(u, s)ds\right)^2 du\right) \\ & \times \left[\exp\left(\sum_i \int_{t_0}^{t} \left(\int_{T_F}^{T_B} \sigma_i(u, s)ds\right)^2 du\right) - 1\right]. \end{aligned} \tag{2.10}$$

The difference between the two variance measures is the overall bias, which can be decomposed into two components, maturity bias and convexity bias, as illustrated in Figure 2.

The maturity bias component, which arises from approximating an instantaneous forward rate by a fixed-maturity forward rate, is given by [3]

$$\text{Maturity Bias} = \frac{1}{(T_B - T_F)^2} \sum_i \int_{t_0}^{t} \left(\int_{T_F}^{T_B} \sigma(u,s)ds \right)^2 du - \sum_i \int_{t_0}^{t} \sigma_i^2(u, T_B)du. \tag{2.11}$$

This bias component is negligible when the fixed-maturity is short (ie. $\tau = T_B - T_F \to 0$). This is in agreement with Chapman et al. (1999) who study the bias induced by using short rates as a proxy for the instantaneous spot rate. They also conclude that the bias is not economically significant in the class of linear short rate models, to which the HJM with deterministic volatility belongs.

The convexity bias component, which arises from approximating the fixed-maturity forward rate by a fixed-maturity futures rate, is given by

$$\text{Convexity Bias} = \left(\frac{1}{T_B - T_F} + y(0, T_F, T_B) \right)^2 e^{2\overline{\sigma}_f^2} \left(e^{\overline{\sigma}_f^2} - 1 \right) - \frac{1}{(T_B - T_F)^2} \overline{\sigma}_f^2, \tag{2.12}$$

where

$$\overline{\sigma}_f^2 = \sum_i \int_{t_0}^{t} \left(\int_{T_F}^{T_B} \sigma_i(u,s)ds \right)^2 du,$$

which is non-negligible due to the existence of the initial futures yield value and the convexity of the exponential function. The difference between forward rates and futures rates results from the difference between forward contract prices and futures contract prices. The marking-to-market feature of futures contracts causes their prices to differ from forward contract prices under a stochastic interest rate environment. Even without this daily marking-to-market feature, Amin and Morton (1994) (p. 152) argue qualitatively that "an unambiguous forward price does not exist which corresponds to the [Eurodollar] futures price". This is because the terminal futures price is based on the yield, which is not a linear function of the price of the traded asset, and therefore, standard arbitrage argument cannot be applied.

From (2.6), it can be seen that the futures contract is a derivative instrument written on the instantaneous forward rate, and therefore, the futures price is driven by the same source of uncertainty as that driving the instantaneous forward rate. Knowing the structure of the uncertainty source allows us to derive the likelihood function for the observable futures prices, without the need to assume and estimate the market price of risk. However, it is proved in Lo (1988) that a naive discretization of the

[3] The derivation for the conditional variance of the fixed-maturity forward rate is given in Appendix C

continuous Itô process, which guarantees the convergence of the discretized sample paths, may not necessarily guarantee the convergence of the discretized estimators to the true parameters of interest, ie. mis-specification of the true likelihood function will lead to inconsistent estimators. In the next section, we will revise the likelihood transformation technique, and derive the likelihood function for quoted futures prices via a state variable whose likelihood function is readily available.

3. The Likelihood Transformation Technique

3.1. **State variables.** Assume that for each underlying pure-discount interest rate instrument, there are K futures contracts maturing at times T_{Fk} ($k = 1, 2, \dots, K$). The (observable) quoted futures price in the market is $G(t, T_{Fk}, T_{Bk})$, which is linked with $F(t, T_{Fk}, T_{Bk})$ via a function γ

$$F(t, T_{Fk}, T_{Bk}) \equiv \eta\big(G(t, T_{Fk}, T_{Bk})\big). \tag{3.1}$$

The link between F and G depends on the quoting convention of each exchange. For example, Eurodollar futures prices traded on the Chicago Mercantile Exchange are quoted as

$$\begin{aligned} F(t, T_{Fk}, T_{Bk}) &= \frac{1}{1 + \left(1 - \frac{G(t, T_{Fk}, T_{Bk})}{100}\right)(T_{Bk} - T_{Fk})} \\ &\equiv \eta\big(G(t, T_{Fk}, T_{Bk})\big). \end{aligned} \tag{3.2}$$

We are considering the case in which all of the futures contracts are written on the same underlying instrument, and therefore the time to maturity of the underlying contract is $T_{Bk} - T_{Fk} = \tau$ constant for all $k \in [0, K]$.

Assuming that there is a measurement error in the market (for example, due to bid-ask spread), we introduce into the evolution of $F(t, T_{Fk}, T_{Bk})$ a new Wiener processes $\widetilde{\varepsilon}_k$ which is independent of the processes driving the uncertainty of forward rates. We further assume that the market errors for the return on futures with different maturities are uncorrelated with each other. The stochastic differential equation for $F(t, T_{Fk}, T_{Bk})$ becomes [4]

$$\frac{dF(t, T_{Fk}, T_{Bk})}{F(t, T_{Fk}, T_{Bk})} = -\sum_i \left(\int_{T_{Fk}}^{T_{Bk}} \sigma_i(t, s) ds\right) d\widetilde{W}_i(t) + \sigma_\varepsilon d\widetilde{\varepsilon}_k. \tag{3.3}$$

[4]In practice, the value of σ_ε should be small (in order and magnitude) in comparison with forward rate volatility σ, so that any attempt to set up an arbitrage portfolio to trade on this uncertainty source will not result in profits after bid-ask spread and transaction costs are taken into account.

Let

$$\begin{aligned} X(t, T_{Fk}, T_{Bk}) &= \ln(F(t, T_{Fk}, T_{Bk})) \\ &\equiv \zeta(F(t, T_{Fk}, T_{Bk})). \end{aligned} \tag{3.4}$$

Application of Itô lemma gives

$$\begin{aligned} dX(t, T_{Fk}, T_{Bk}) = &- \frac{1}{2} \left[\sum_i \left(\int_{T_{Fk}}^{T_{Bk}} \sigma_i(t,s) ds \right)^2 + \sigma_\varepsilon^2 \right] dt \\ &- \sum_i \int_{T_{Fk}}^{T_{Bk}} \sigma_i(t,s) ds \, d\widetilde{W}(t) + \sigma_\varepsilon d\widetilde{\varepsilon}_k. \end{aligned} \tag{3.5}$$

If we know the likelihood function for X, then we can use the transformation technique twice to derive first the likelihood function for F and then the market quoted variable G. In the next section, we will revise the likelihood transformation technique, and utilize Duan's (1994) result to simplify it. In the subsequent section, we will write out the exact likelihood function for the quoted futures price, pooled time series and cross sectional data.

3.2. **Likelihood transformation formula.** Let $X_{j\,k} \equiv X(t_j, T_{Fk}, \tau) \equiv X(t_j, T_{Fk}, T_{Bk})$[5] be an unobservable state variable k ($k = 1, 2, \ldots, K$) occurring at time $t_j < T_F$ ($j = 0, 1, \ldots, J$).

Denote by $\boldsymbol{x}_j$ the vector of unobservable state variables occurring at the t_j, ie. $\boldsymbol{x}_j = \big(X(t_j, T_1, \tau), X(t_j, T_2, \tau), \ldots, X(t_j, T_K, \tau)\big)$. Denote by $\boldsymbol{x}$ the unobservable state vector of size $K(J+1) \times 1$ at time t_J, ie.

$$\begin{aligned} \boldsymbol{x} &= \text{vec} \begin{pmatrix} \boldsymbol{x}_0 & \boldsymbol{x}_1 & \ldots & \boldsymbol{x}_J \end{pmatrix} \\ &= \text{vec} \begin{pmatrix} X(t_0, T_{F1}, \tau) & X(t_1, T_{F1}, \tau) & \ldots & X(t_J, T_{F1}, \tau) \\ X(t_0, T_{F2}, \tau) & X(t_1, T_{F2}, \tau) & \ldots & X(t_J, T_{F2}, \tau) \\ \vdots & \vdots & \ddots & \vdots \\ X(t_0, T_{FK}, \tau) & X(t_1, T_{FK}, \tau) & \ldots & X(t_J, T_{FK}, \tau) \end{pmatrix}, \end{aligned}$$

where vec is the standard matrix operator that, when applied to a matrix, transforms the matrix into a vector by stacking the columns of the matrix on top of each other.

Denote the density function of $\boldsymbol{X}$ by

$$p_{\boldsymbol{X}}(\boldsymbol{x}; \boldsymbol{\theta}) = p_{\boldsymbol{X}}(\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_J; \boldsymbol{\theta}),$$

where $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ is the parameter vector of interest.

[5]We can write $X(t_j, T_{Bk}, \tau) \equiv X(t_j, T_{Fk}, T_{Bk})$ because $T_{Bk} - T_{Fk} = \tau$ constant for all $k = 1, 2, \ldots, K$

Suppose that a transformation $\boldsymbol{\vartheta}$ exists, which applied to $\boldsymbol{X}$, produces a vector $\boldsymbol{Z}$ that is observable in the market

$$\boldsymbol{Z} = \boldsymbol{\vartheta}(\boldsymbol{X}; \boldsymbol{\theta}) : \mathbb{R}^{K(J+1)\times 1} \to \mathbb{R}^{K(J+1)\times 1},$$

where

$$\begin{aligned} \boldsymbol{z} &= \text{vec} \begin{pmatrix} \boldsymbol{z}_0 & \boldsymbol{z}_1 & \dots & \boldsymbol{z}_J \end{pmatrix} \\ &= \text{vec} \begin{pmatrix} Z_{01} & Z_{11} & \dots & Z_{J1} \\ Z_{02} & Z_{12} & \dots & Z_{J2} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{0K} & Z_{1K} & \dots & Z_{JK} \end{pmatrix}. \end{aligned}$$

Assume that this transformation is one-to-one for every $\boldsymbol{\theta} \in \boldsymbol{\Theta}$.

Since $\boldsymbol{\vartheta}$ is one-to-one, there exists an inverse $\boldsymbol{\vartheta}^{-1} = \boldsymbol{\zeta}(\boldsymbol{Z}; \boldsymbol{\theta})$. Applying the standard change of variable technique

$$p_{\boldsymbol{Z}}(\boldsymbol{z}, \boldsymbol{\theta}) = p_{\boldsymbol{X}}\left(\boldsymbol{\zeta}(\boldsymbol{z}; \boldsymbol{\theta})\right) \times \left| \text{J}\left(\boldsymbol{\zeta}(\boldsymbol{z}; \boldsymbol{\theta})\right) \right|,$$

where J is the Jacobian of the transformation from $\boldsymbol{X}$ to $\boldsymbol{Z}$, ie.

$$\text{J}\left(\boldsymbol{\zeta}(\boldsymbol{z}; \boldsymbol{\theta})\right) = \left| \frac{\partial \boldsymbol{\zeta}(\boldsymbol{z}; \boldsymbol{\theta})}{\partial \boldsymbol{z}'} \right|.$$

Duan (1994) proves that if the transformation is on an element-by-element basis, ie. $Z_{j\,k} = \vartheta_{j\,k}(X_{j\,k})$ (and $X_{j\,k} = \zeta_{j\,k}(Z_{j\,k})$) for all $j \in [0, J]$ and $k \in [1, K]$, then the first-derivative matrix is diagonal, therefore

$$\text{J}\left(\boldsymbol{\zeta}(\boldsymbol{z}; \boldsymbol{\theta})\right) = \prod_{j=0}^{J} \prod_{k=1}^{K} \frac{d\zeta_{j\,k}(Z_{j\,k}; \boldsymbol{\theta})}{dZ_{j\,k}}.$$

Furthermore, if $\boldsymbol{X}$ is "joint-Markovian", ie.

$$p_{\boldsymbol{X}}(\boldsymbol{x}_0, \boldsymbol{x}_1, \dots, \boldsymbol{x}_J; \boldsymbol{\theta}) = p_{\boldsymbol{X}}(\boldsymbol{x}_0, t_0; \boldsymbol{\theta}) \prod_{j=1}^{J} p_{\boldsymbol{X}}(\boldsymbol{x}_j, t_j | \boldsymbol{x}_{j-1}, t_{j-1}; \boldsymbol{\theta})$$

then upon substitution, the likelihood for Z can be compactly written as

$$p_{\boldsymbol{Z}}(\boldsymbol{z}_0, \boldsymbol{z}_1, \dots, \boldsymbol{z}_J; \boldsymbol{\theta}) = p_{\boldsymbol{Z}}(\boldsymbol{z}_0, t_0; \boldsymbol{\theta}) \prod_{j=1}^{J} p_{\boldsymbol{Z}}(\boldsymbol{z}_j, t_j | \boldsymbol{z}_{j-1}, t_{j-1}; \boldsymbol{\theta}),$$

where

$$\begin{aligned} p_{\boldsymbol{Z}}(\boldsymbol{z}_0, t_0; \boldsymbol{\theta}) &= p_{\boldsymbol{X}}\big(\boldsymbol{\zeta}_0(\boldsymbol{y_0}), t_0; \boldsymbol{\theta}\big) \times \Big| \text{J}\big(\boldsymbol{\zeta}_0(\boldsymbol{z}_0; \boldsymbol{\theta})\big) \Big|, \\ p_{\boldsymbol{Z}}\big(\boldsymbol{z}_j, t_j | \boldsymbol{z}_{j-1}, t_{j-1}; \boldsymbol{\theta}\big) &= p_{\boldsymbol{X}}\big(\boldsymbol{\zeta}_j(\boldsymbol{z}_j), t_j | \boldsymbol{\zeta}_{j-1}(\boldsymbol{z}_{j-1}), t_{j-1}; \boldsymbol{\theta}\big) \times \Big| \text{J}\big(\boldsymbol{\zeta}_j(\boldsymbol{z}_j; \boldsymbol{\theta})\big) \Big|, \end{aligned}$$

and

$$\mathrm{J}\big(\boldsymbol{\zeta}(\boldsymbol{z}_j;\boldsymbol{\theta})\big) = \prod_{k=1}^{K} \frac{d\zeta_{j\,k}(Z_{j\,k};\boldsymbol{\theta})}{dZ_{j\,k}} \quad (j = 0, 1, \ldots, J).$$

3.3. **Full information maximum likelihood function.** We now first focus on the state variable $X(t, T_{Fk}, T_{Bk})$ which is driven by the stochastic differential equation

$$\begin{aligned} dX(t, T_{Fk}, T_{Bk}) = &- \frac{1}{2}\left[\sum_i \left(\int_{T_{Fk}}^{T_{Bk}} \sigma_i(t,s)ds\right)^2 + \sigma_\varepsilon^2\right] dt \\ &- \sum_i \int_{T_{Fk}}^{T_{Bk}} \sigma_i(t,s)ds\, d\widetilde{W}(t) + \sigma_\varepsilon d\widetilde{\varepsilon}_k. \end{aligned} \tag{3.6}$$

Suppose that the process is sampled at $J+1$ discrete points in time $t_0, t_1, \ldots, t_n$ (not necessarily equally spaced apart). Due to the Markovian nature of the stochastic process for $X(t, T_{Fk}, T_{Bk})$, the likelihood function for $(X_{0\,k}, X_{1\,k}, \ldots, X_{J\,k})$ [6], for a given parameter vector of interest $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, is

$$p_{\boldsymbol{X}}(X_{0\,k}, X_{1\,k}, \ldots, X_{J\,k};\boldsymbol{\theta}) = p_{\boldsymbol{X}}(X_{0\,k}, t_0;\boldsymbol{\theta}) \prod_{j=1}^{J} p_{\boldsymbol{X}}(X_{j\,k}, t_j | X_{(j-1)\,k}, t_{j-1};\boldsymbol{\theta}).$$

With this discrete sample, it is proved in Lo (1988) [7] that the transitional likelihood function has a Gaussian form

$$p_{\boldsymbol{X}}(X_{j\,k}, t_j | X_{(j-1)\,k}, t_{j-1};\boldsymbol{\theta}) = \left[2\pi\beta_{j(kk)}\right]^{-\frac{1}{2}} \exp\left[-\frac{\left(X_{j\,k} - X_{(j-1)\,k} - \alpha_{j\,k}\right)^2}{2\beta_{j(kk)}}\right],$$

where

$$\text{Drift } \alpha_{j\,k} = -\frac{1}{2}\left[\sum_i \int_{t_{j-1}}^{t_j} \left(\int_{T_{Fk}}^{T_{Bk}} \sigma_i(u,s)ds\right)^2 du + \int_{t_{j-1}}^{t_j} \sigma_\varepsilon^2 du\right], \tag{3.7}$$

$$\text{Variance } \beta_{j(kk)} = \sum_i \int_{t_{j-1}}^{t_j} \left(\int_{T_{Fk}}^{T_{Bk}} \sigma_i(u,s)ds\right)^2 du + \int_{t_{j-1}}^{t_j} \sigma_\varepsilon^2 du. \tag{3.8}$$

[6] Recall that $X_{j\,k} \equiv X(t_j, T_{Fk}, T_{Bk})$

[7] Lo (1988) proves the case where there is only one noise term. By substitution, it is a straight forward extension to prove the result for the multiple-noise case. In any event, the result is merely a consequence of the fact that the process (3.6) for $X(t, T_{Fk}, T_{Bk})$ is Gaussian due to the assumption of time dependent volatility functions.

If we incorporate cross-sectional data into our study to exploit the full information content of the yield curve, the transitional likelihood function will have a multi-dimensional Gaussian form

$$p_{\boldsymbol{X}}(\boldsymbol{x}_j, t_j|\boldsymbol{x}_{j-1}, t_{j-1};\boldsymbol{\theta}) = (2\pi)^{-\frac{K}{2}} \left|\boldsymbol{\Omega}_j\right|^{-\frac{1}{2}} \times \exp\left(-\frac{1}{2}\left(\boldsymbol{x_j} - \boldsymbol{x}_{j-1} - \boldsymbol{\alpha}_j\right)' \boldsymbol{\Omega}_j^{-1}\left(\boldsymbol{x_j} - \boldsymbol{x}_{j-1} - \boldsymbol{\alpha}_j\right)\right), \tag{3.9}$$

where

$$\text{Drift } \boldsymbol{\alpha}_j = (\alpha_{j\,1}\,,\, \alpha_{j\,2}\,,\, \ldots, \alpha_{j\,k}\,,\, \ldots\,,\, \alpha_{j\,K})'\,,$$

$$\text{Covariance matrix } \boldsymbol{\Omega}_j = \begin{pmatrix} \beta_{j(11)} & \beta_{j(12)} & \cdots & \beta_{j(1K)} \\ \beta_{j(21)} & \beta_{j(22)} & \cdots & \beta_{j(2K)} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{j(K1)} & \beta_{j(K2)} & \cdots & \beta_{j(KK)} \end{pmatrix},$$

and for $k_1 \neq k_2$ [8]

$$\beta_{j(k_1k_2)} = \sum_i \int_{t_{j-1}}^{t_j} \left(\int_{T_{Fk_1}}^{T_{Bk_1}} \sigma_i(u,s)ds\right)\left(\int_{T_{Fk_2}}^{T_{Bk_2}} \sigma_i(u,s)ds\right) du. \tag{3.10}$$

The log likelihood function for the state variable $\boldsymbol{X}$ is

$$\mathcal{L}_{\boldsymbol{X}}(\boldsymbol{\theta}) = \sum_{j=1}^{J} \ln\Big(p_{\boldsymbol{X}}\big(\boldsymbol{x}_j, t_j|\boldsymbol{x}_{j-1}, t_{j-1};\boldsymbol{\theta}\big)\Big). \tag{3.11}$$

In the above formula we have ignored the unconditional probability of the first observation at time t_0. As argued in Aït-Sahalia (2001), this unconditional probability is dominated by the sum of all conditional density terms when the sample size becomes large.

Recall that there exists a transformation from $\boldsymbol{X}$ to $\boldsymbol{F}$ (see (3.4)) with inverse function $\boldsymbol{\zeta}$. It is clear that this transformation is on element-by-element basis. Applying the transformation formula, the likelihood function for $\boldsymbol{F}$ is

$$\mathcal{L}_{\boldsymbol{F}}(\boldsymbol{\theta}) = \sum_{j=1}^{J} \ln\Big(p_{\boldsymbol{X}}\big(\boldsymbol{\zeta}(\boldsymbol{F}_j), t_j|\boldsymbol{\zeta}(\boldsymbol{F}_{j-1}), t_{j-1};\boldsymbol{\theta}\big)\Big) + \sum_{j=1}^{J} \ln\left|\frac{\partial\boldsymbol{\zeta}_j(\boldsymbol{F}_j;\boldsymbol{\theta})}{\partial\boldsymbol{F}_j}\right|. \tag{3.12}$$

Applying the transformation the second time from $\boldsymbol{F}$ to $\boldsymbol{G}$, the quoted futures price in the market, with the inverse transformation function $\boldsymbol{\eta}$ (see (3.1)) results in the log

[8]Note that the drift $\alpha_{j\,k}$ and variance $\beta_{j(kk)}$ have been defined in (3.7) and (3.8) respectively.

likelihood function

$$\mathcal{L}_{\boldsymbol{G}}(\boldsymbol{\theta}) = \sum_{j=1}^{J} \ln \Big(p_{\boldsymbol{F}}\big(\boldsymbol{\eta}(\boldsymbol{G}_j), t_j | \boldsymbol{\eta}(\boldsymbol{G}_{j-1}), t_{j-1}; \boldsymbol{\theta}\big)\Big) + \sum_{j=1}^{J} \ln \left| \frac{\partial \boldsymbol{\eta}_j(\boldsymbol{G}_j; \boldsymbol{\theta})}{\partial \boldsymbol{G}_j} \right|. \tag{3.13}$$

4. MODELS AND DATA

4.1. **Models.** In this paper, we focus on single-factor HJM model, ie. there is only a single source of uncertainty. Since we are using futures data, which is usually actively traded for maturities less than 5 years, there is insufficient variation in the term structure across different maturities to separate the effect of different uncertainty sources. In addition, Dybvig (1990), as cited in Amin and Morton (1994), shows that almost all of the variation in forward rates with maturities less than five years can be explained by a dominant single factor.

The class of HJM model with which one is working is determined by the specification of the volatility function. We choose a fairly general "time-invariant" humped-volatility curve, ie. the volatility $\sigma(t,T)$ depends on $T-t$ only, not on the calendar date t, thus

$$\sigma(t,T) = [\sigma_0 + \sigma_1(T-t)]\exp(-\kappa(T-t)). \tag{4.1}$$

Despite the fact that the implied volatility functions obtained from caps and swaptions data often exhibit a humped volatility structure (Amin and Morton (1994), p. 160, and Hull and White (1996), p. 33), as far as we know, there has so far only the attempt of Ritchken and Chuang (1999) to estimate the humped-volatility model of the form (4.1) in the HJM framework[9]. The model nests many of the time-deterministic volatility forms considered in the literature so far:

- The exponential model (Hull and White (1990) Extended Vasicek Model): $\sigma(t,T) = \sigma_0 \exp(-\kappa(T-t))$
- The linear absolute model: $\sigma(t,T) = \sigma_0 + \sigma_1(T-t)$
- The absolute (or constant) model (Ho and Lee (1986) model): $\sigma(t,T) = \sigma_0$

The analytical expression for the log likelihood function of the quoted futures price under this volatility specification involves performing the integrations in (3.7), (3.8), (3.10), and details can be found in Appendix D.

4.2. **Data.** We apply the method outlined above to short term interest rate futures contracts traded on the Chicago Mercantile Exchange (CME). The CME contracts are

[9]Ritchken and Chuang (1999) also need to rely on the Markovianization of the interest rate system. Even though we estimate the same model, we do not rely on the property of Markovianization of the system

written on Eurodollar Time Deposits with a three-month maturity. The last trading day for each contract is the second London bank business day before the third Wednesday of the contract month, which rests in the March, June, September, December cycle. The data is taken from Datastream.

The CME Eurodollar futures contracts are chosen for their extreme liquidity. Table 1 reports the average daily trading volume of contracts used in our study.

Table 1: CME Eurodollar Futures Contracts

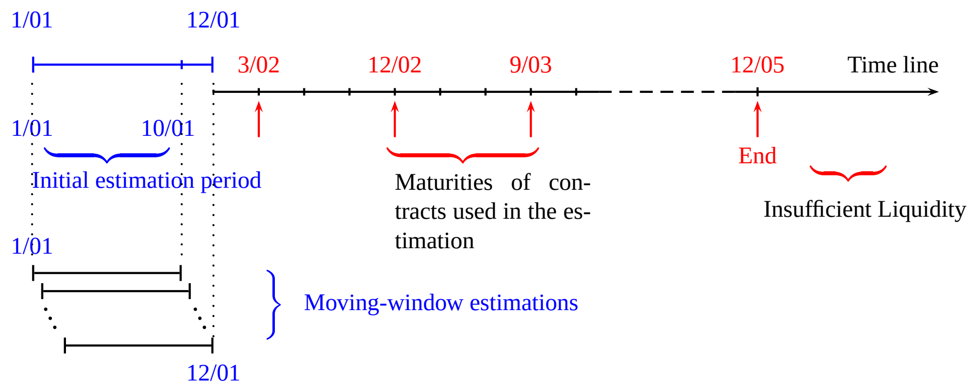
This table reports the contracts used in our estimation, as explained in the text, and their corresponding average daily trading volume in $US

Year	Begin Contract	End Contract	Total Number of Contracts	Observation per Series	Average Daily Trading Volume
1987	03/1988	09/1988[10]	2	211	5,653
1988	03/1989	12/1989	2	211	7,100
1989	03/1990	09/1991	3	211	7,119
1990	03/1991	09/1992	3	213	8,216
1991	03/1992	06/1994	4	212	8,238
1992	03/1993	06/1995	4	213	14,913
1993	03/1994	12/1997	6	210	11,840
1994	03/1995	12/1998	6	210	19,434
1995	03/1996	12/1999	6	210	15,397
1996	03/1997	12/2000	6	214	15,883
1997	03/1998	12/2001	6	210	16,990
1998	03/1999	12/2002	6	213	18,709
1999	03/2000	12/2003	6	209	16,497
2000	03/2001	12/2004	6	211	17,926
2001	03/2002	12/2005	6	210	30,762

The CME data covers the 15-year period from January 1, 1987 to December 31, 2001. The period is chosen so that the first 6-year period coincides with the data used in Amin and Morton (1994). We estimate our model for each year period separately, since the volatility parameters must reflect the current market condition, as also argued in Bühler et al. (1999). Each year we use trading data from January 1 to October 31 to form initial estimates. We then use trading data during November and December to check parameter stability, by the moving window approach. For each trading year, the futures series considered starts from the March contract maturing the following year, until the last actively traded contracts. To ensure a sufficient variation in futures prices,

[10] According to our design, we would have chosen the 12/1988 contract if it had been traded

FIGURE 3. Research Design



and so avoid possible singularity of the covariance matrix, the set of contracts used are spaced three quarter apart. For example, to estimate volatility parameters for 2001, we use March 2002, December 2002, September 2003, June 2004, March 2005 and December 2005 contracts (see Figure 3)[11]. Since the trading activities in each year are different, the number of contracts included in our analysis varies with time, as shown in table 1. From 1993 to 2001, 6 contracts are included in our analysis. On average, there are 211 observations for each series.

5. EMPIRICAL RESULTS

5.1. **First sample period: 1987-1992.** The estimates of parameters of the humped-volatility model for the first sample period can be found in table 2. This sample period is chosen to coincide with the sample used in Amin and Morton's (1994) implied volatility work. In their article, they conclude in favour of the absolute model, due to its ability to deliver stable parameter values, and to deliver profits when it trades on perceived mispricing. In line with their finding, it is not a surprise that we find a highly insignificant estimate of σ_1 and κ in most of the years, with the exception of 1988 and 1992, suggesting that the forward volatility is over-parameterized.

[11]For all of the years, we have repeated the estimation using different combination of futures price series (such as different starting contracts and different spacing between contracts), and the estimation results are not significantly different

Table 2: Humped-volatility model estimation for the first sample period 1987-1992

This table reports the estimation result for period 1987-1992 under forward rate humped-volatility specification, $\sigma(t,T) = (\sigma_0 + \sigma_1(T-t))\exp(-\kappa(T-t))$. Asymptotic standard errors of the estimate (White-consistent estimator) are inside parentheses. The symbol † indicates insignificant parameter values at 99% confidence level. All values of σ and their corresponding standard errors are reported in percentage.

Year	σ_0 (%)	σ_1 (%)	κ	σ_ε (%)	Log Likelihood
1987	2.299	0.3×10^{-5}	0.066	0.087	545.65
	(0.133)	$(0.414)^\dagger$	$(0.183)^\dagger$	(0.004)	
1988	0.848	1.622	0.706	0.072	712.12
	(0.296)	(0.629)	(0.128)	(0.003)	
1989	1.932	1.76×10^{-5}	0.246	0.119	969.3
	(0.123)	$(0.306)^\dagger$	$(0.160)^\dagger$	(0.004)	
1990	1.056	0.581	0.411	0.118	1034.4
	(0.217)	$(0.459)^\dagger$	(0.158)	(0.004)	
1991	1.115	3.02×10^{-5}	0.164	0.078	1760.1
	(0.065)	$(0.128)^\dagger$	$(0.116)^\dagger$	(0.002)	
1992	0.341	1.389	0.657	0.093	1565.0
	(0.163)	(0.275)	(0.032)	(0.003)	

We therefore re-estimate our models for 1987, 1989, 1990 and 1991 with the exponential, linear absolute and absolute models, and check the overall model fit by the likelihood ratio test. Among this class, the exponential model performs best, delivering significant parameter estimates without any significant loss of likelihood values (see table 3).

Similar to the implied value of Amin and Morton (1994), we find that the instantaneous volatility of the spot rate (which is σ_0 in the HJM model with time-deterministic instantaneous forward rate volatility) averages at about 1.3%. The standard error of our historical estimate is lower than the standard deviation of their implied values. Moreover, we do not find that our estimate is unstable with respect to initial parameter values nor to have large Hessian matrix (in order and magnitude) as reported in their study. The estimated volatilities of measurement error (σ_ε) are small in order and magnitude in all years.

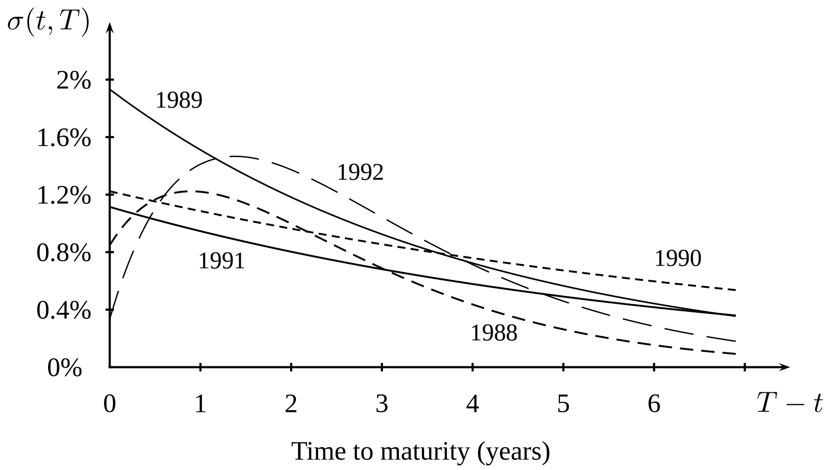
As can be seen from Figure 4, the humps in the volatility curves of 1988 and 1992 occur at about 1-1.5 years to maturity. The exponential volatility curves (in 1988, 1989, 1990) have positive decay factors κ, thus, the spot rate has higher instantaneous

TABLE 3. Estimation result for the first sample period 1987-1992

This table reports the best model for each year and the corresponding parameter values. Asymptotic standard errors of the estimate (White-consistent estimator) are inside parentheses. All values of σ and their corresponding standard errors are reported in percentage. The p-value for likelihood ratio test (between the humped model and the model reported here) are inside square brackets and under the corresponding likelihood value

Year	Model	σ_0 (%)	σ_1 (%)	κ	σ_ε (%)	Log Likelihood
1987	Exponential	2.299	-	0.066	0.087	545.65
		(0.133)	-	(0.030)	(0.004)	[1.000]
1988	Humped	0.848	1.622	0.706	0.072	712.12
		(0.296)	(0.629)	(0.128)	(0.003)	-
1989	Exponential	1.932	-	0.246	0.119	969.3
		(0.123)	-	(0.025)	(0.004)	[1.000]
1990	Exponential	1.2234	-	0.120	0.118	1033.9
		(0.088)	-	(0.031)	(0.004)	[0.480]
1991	Exponential	1.115	-	0.164	0.078	1760.1
		(0.065)	-	(0.016)	(0.002)	[1.000]
1992	Humped	0.341	1.389	0.657	0.093	1565.0
		(0.163)	(0.275)	(0.032)	(0.003)	-

FIGURE 4. Instantaneous forward volatility - First sample period



volatility than any forward rates, and short term forward rates have higher instantaneous volatility than longer term forward rates.

5.2. **Second sample period: 1993-2001.** In the second sample period, we find that the humped-volatility model is preferred to the exponential, linear absolute and the absolute model. The parameter estimates can be found in table 4.

Table 4: Humped-volatility model estimation for the second sample period: 1993-2001

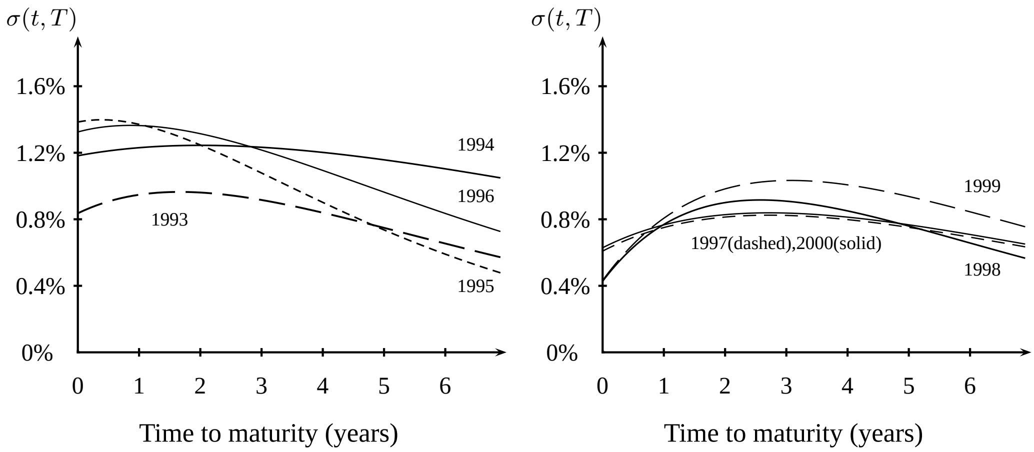
This table reports the estimation result for period 1993-2001 under forward rate humped-volatility specification, $\sigma(t,T) = (\sigma_0 + \sigma_1(T-t))\exp(-\kappa(T-t))$. Asymptotic standard errors of the estimate (White-consistent estimator) are inside parentheses. All values of σ and their corresponding standard errors are reported in percentage.

Year	σ_0(%)	σ_1(%)	κ	σ_ε(%)
1993	0.836	0.401	0.267	0.072
	(0.079)	(0.092)	(0.031)	(0.001)
1994	1.182	0.238	0.144	0.055
	(0.070)	(0.067)	(0.027)	(0.001)
1995	1.385	0.553	0.346	0.080
	(0.112)	(0.136)	(0.033)	(0.002)
1996	1.325	0.445	0.261	0.061
	(0.085)	(0.085)	(0.023)	(0.001)
1997	0.609	0.324	0.218	0.037
	(0.040)	(0.038)	(0.016)	(0.006)
1998	0.427	0.617	0.306	0.069
	(0.055)	(0.069)	(0.019)	(0.001)
1999	0.432	0.619	0.265	0.064
	(0.055)	(0.066)	(0.017)	(0.001)
2000	0.629	0.320	0.213	0.094
	(0.061)	(0.068)	(0.047)	(0.002)
2001	1.051	0.213	0.124	0.094
	(0.083)	(0.104)	(0.047)	(0.002)

A humped forward volatility curve implies that the instantaneous volatility of the spot rate is lower than short-term forward volatilities. However, forward volatility gradually decreases as time to maturity increases, and finally reaches a lower level than the spot rate volatility. Figure 5 shows that the humps usually occur at 1-3 years to maturity.

From Figure 5, there appears to have temporal clusters of spot rate volatilities (the instantaneous volatility of the spot rate is equal to the instantaneous volatility of a forward rate with zero time to maturity). The spot rate volatilities average at 1.1% during 1993-1996, then decrease to around 0.5% level during 1997-2000, and finally

FIGURE 5. Instantaneous forward volatility - Second sample period



bounce back to 1.1% in 2001 (see also table 4)[12]. At the end of the forward volatility curve, where there is still long time to maturity, the volatilities remain stable. The 7-year instantaneous forward volatility averages at about 0.7%.

5.3. **Stability of the estimates.** To check the stability of our estimates, we use a "moving window" approach. We use trading data from January to October each year to estimate our model. Then we move our window sample by 1 day, keeping sample size constant (ie. the drop-one/add-one method) to compute sequential estimates until the end of December each year. Figure 6 plots the series of instantaneous spot rate volatility σ_0 and the decay factor κ obtained in 1990. There are some fluctuations in the series, but they do not seem to be unstable. The results for other years are reported in table 5, which shows that the sequential estimates have low standard deviations.

[12]However, at this point of time, we do not have data to check whether the spot rate volatility will remain at this level for the next few years.

FIGURE 6. Moving Window Approach: Parameter Estimates for 1990

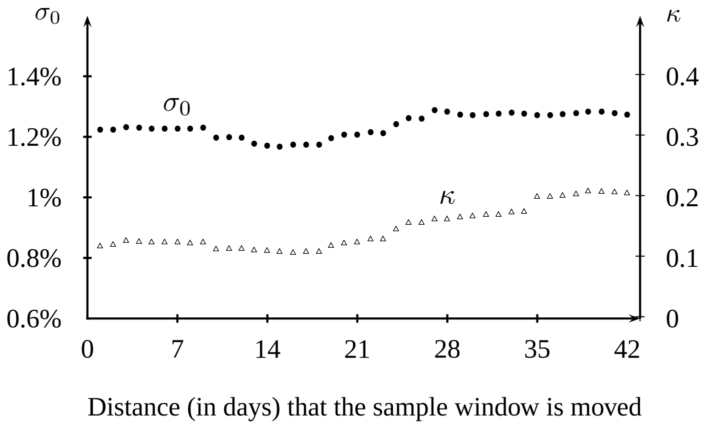


Table 5: Sequential estimates

The table reports the sequential estimate from moving window approach. Each year the first estimate window starts with trading data covering January to October. Then the sample is moved by 1 day (add-one/drop-one method), and the model is re-estimated. The process is repeated until all trading days in November and December are included in the samples. The value of σ_0 and their corresponding standard deviation are reported in percentage.

Year	σ_0 (%)		κ	
	Average	Standard Deviation	Average	Standard Deviation
1987	2.324	0.012	0.066	0.030
1988	0.895	0.093	0.788	0.073
1989	1.951	0.010	0.249	0.005
1990	1.236	0.039	0.148	0.035
1991	1.101	0.023	0.175	0.009
1992	0.372	0.055	0.661	0.008
1993	0.609	0.101	0.325	0.024
1994	1.121	0.031	0.179	0.019
1995	1.249	0.107	0.345	0.007
1996	1.153	0.141	0.313	0.032
1997	0.585	0.017	0.226	0.007
1998	0.423	0.014	0.303	0.010
1999	0.401	0.029	0.273	0.011
2000	0.572	0.025	0.242	0.013

FIGURE 7. Forward Volatility in 1990 (exponential shaped) and 2000 (humped shaped) - Moving Window Approach

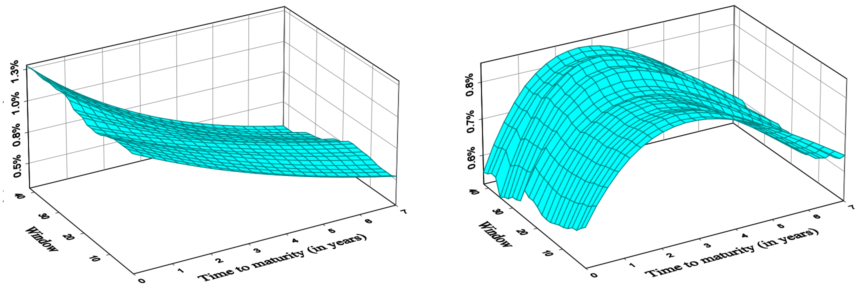


Table 5: (Continued)

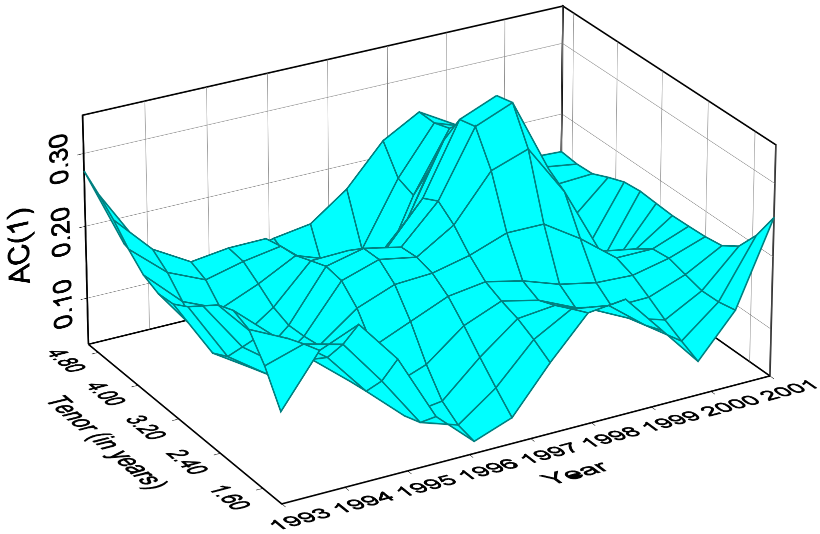
Year	σ_0 (%)		κ	
	Average	Standard Deviation	Average	Standard Deviation
2001	0.755	0.154	0.270	0.066

The very small changes of our parameter estimates implies that the resulting forward volatility curve experiences only slight and smooth movement over time. Figure 7 graphs two representative volatility curves, one having a simple exponential shape and the other having the humped shape. As time goes, the instantaneous forward volatility of long-to-maturity forward rates is more volatile than that of short-to-maturity forward rates. Overall, the smooth volatility surfaces indicate adequate stability in our estimation results.

5.4. **Model fit.** The model's goodness of fit are assessed by tests on residuals. Since the residuals of our estimates have different variances at each point of time by model construction, we carry out goodness of fit test by checking the estimated standardized residuals.

To test whether the standardized residuals come from a multivariate normal distribution, we employ the Omnibus test, which has been corrected for small sample bias and adapted to the multivariate case by Doornik and Hansen (1994). The test is derived from Shenton and Bowman (1977), who give the sample kurtosis a gamma distribution, and D'Agostino (1970), who approximates the distribution of sample skewness by the Johnson S_u system. Under this test, we can reject the null hypothesis of normal distribution for all sample periods, at 99% confidence level.

FIGURE 8. First order serial correlation in estimated standardized residuals
(The graph plots the absolute value of the correlation coefficients)



In addition, we calculate the serial correlation for the estimated standardized residuals. Even though our estimated correlation coefficients are much smaller than the level of 0.90 reported by De Jong and Santa-Clara (1999), they are still high. The absolute value of first order correlation coefficient averages at 0.16. The correlation reduces as the lag time lengthens, but is still at 0.05 level at lag order 30 (see table 6). Figure 8, which plots the absolute value of the first order residual serial correlation, shows that for most of the years, the residual serial correlation is higher in the mid-range maturities, whereas for the short rates and long rates, the serial correlation is weaker.

The existence of serial correlation in the estimated standardized residuals up to very long lags suggests that the model is misspecified. There are two ways to account for this striking autocorrelation feature. The first is to consider other HJM specifications where the instantaneous forward rate depends on the whole history of the term structure. This can be done by including either the instantaneous forward rate itself or the instantaneous spot rate into the specification of the instantaneous forward rate volatility. The second is to consider forward rate models with jump. The omission of a jump component when it exists will also result in autocorrelation. We leave these issues for future research.

Table 6: Residual serial correlation

This table reports the serial correlation coefficients for the standardized residuals from the HJM estimation, at lag order 1 and order 30.

Year	AC	Maturity (in years)					
		1.25	2	2.75	3.5	4.25	5
1987	$\rho(1)$	0.361	0.053				
	$\rho(30)$	0.044	-0.045				
1988	$\rho(1)$	-0.340	-0.090				
	$\rho(30)$	0.015	-0.159				
1989	$\rho(1)$	0.109	0.112				
	$\rho(30)$	0.127	-0.082				
1990	$\rho(1)$	0.044	-0.332	-0.062			
	$\rho(30)$	-0.091	0.007	-0.028			
1991	$\rho(1)$	0.208	0.286	-0.080	0.169		
	$\rho(30)$	0.052	0.009	0.054	0.018		
1992	$\rho(1)$	0.029	0.181	0.164	0.177		
	$\rho(30)$	-0.003	0.004	-0.011	0.002		
1993	$\rho(1)$	0.159	0.240	0.217	0.183	0.230	0.276
	$\rho(30)$	0.047	0.011	0.030	0.045	0.044	0.005
1994	$\rho(1)$	0.290	0.194	0.151	0.057	-0.143	0.099
	$\rho(30)$	-0.094	0.044	-0.024	-0.078	-0.003	0.031
1995	$\rho(1)$	0.105	0.160	0.127	0.197	0.235	-0.071
	$\rho(30)$	0.002	-0.041	-0.062	-0.061	-0.073	-0.140
1996	$\rho(1)$	0.079	0.097	0.214	0.148	0.129	0.122
	$\rho(30)$	-0.108	-0.023	-0.051	-0.043	-0.058	-0.033
1997	$\rho(1)$	-0.097	0.156	-0.185	-0.199	0.041	0.140
	$\rho(30)$	0.006	0.069	0.026	-0.049	0.013	0.020
1998	$\rho(1)$	0.241	0.148	0.176	0.365	0.259	0.259
	$\rho(30)$	0.057	-0.048	-0.102	0.019	-0.035	-0.046
1999	$\rho(1)$	0.064	0.155	0.281	0.301	0.181	0.173
	$\rho(30)$	0.012	-0.096	-0.072	0.007	0.014	0.014
2000	$\rho(1)$	0.104	0.108	0.079	-0.023	0.148	0.174
	$\rho(30)$	0.048	0.024	-0.024	0.135	0.084	0.074
2001	$\rho(1)$	0.254	0.172	0.161	0.172	0.143	0.146
	$\rho(30)$	0.087	0.003	-0.108	-0.047	-0.007	-0.003

6. Conclusion

This paper focuses on a method of estimation for a rich family of Heath-Jarrow-Morton term structure models where the instantaneous forward rate volatility is time deterministic, under which the process for the instantaneous spot rate may or may not be able to be Markovianized. It is important to choose a model that best describes the data. The resulting model then can be used by practitioners in their calibration procedure.

Among different methods of estimation, the Maximum Likelihood Estimator has favourable asymptotic properties. However, it cannot be applied directly in the HJM framework due to the need to assume and estimate a functional form for the market price of interest rate risk, and more importantly, due to the existence of the unobservable instantaneous forward rate volatility. The attempt to use futures rates as a proxy for forward rates leads to non-negligible estimation bias, which can be decomposed into a maturity bias component and a convexity bias component.

The major contribution of this paper rests on the realization that a futures contract can be viewed as a derivative instrument written on instantaneous forward rate, and therefore is driven by the same source of uncertainty as that driving the forward rate evolution. Using a likelihood transformation technique, and utilizing the result of Duan (1994) to simplify the likelihood function, we are able to derive the exact likelihood function for all model specifications that have deterministic volatility forms, albeit the likelihood function will be different in its degree of complexity.

To demonstrate our method, we focus on the humped-forward rate volatility specification suggested by the hump that is often revealed when an implied volatility function is backed out from caps and swaptions data. We use 15-year data (from 1987-2001) of CME Eurodollar futures data to estimate our model. We not only use time series, but also pool in cross-sectional data, ie. futures contracts that have different tenors at each point of time, in order to exploit the full information content of the yield curve.

For most of the years in our sample periods, we find that the humped-volatility model performs adequately among its class. The exponential model works better during the 3 years at the beginning of the sample period. There appears to be a temporal clustering of instantaneous spot rate volatility. At the longer end of the curve, where there is still long time to maturity, the instantaneous forward volatility stays at a fairly constant level. Our estimate remains stable with respect to initial parameter values and sample windows. However, the chosen volatility functional form does not fully capture all the features of the data.

We have set up a framework that allows estimation of all HJM model specifications which have time-deterministic instantaneous forward rate volatility, where the instantaneous spot rate process may or may not be able to be represented in a Markovianized system. Nevertheless, there still remains the empirical need and challenge to estimate and test other non-deterministic forms of forward rate volatilities, or forward rate volatility containing jump components. We intend to explore these issues in subsequent research.

APPENDIX A. THE EVOLUTION OF FUTURES PRICE UNDER HJM MODEL

Let $P(t, T_B)$ be the price at time t of a pure discount instrument that has a face value of \$1 and matures at time T_B, and let $B(t, T_B)$ be the corresponding log bond price, ie $B(t, T_B) = \ln P(t, T_B)$.

Denote by $F(t, T_F, T_B)$ the price at time t of a futures contract written on the pure discount instrument. The futures contract matures at time T_F.

Since futures contracts are marked-to-market, it is shown in Cox et al. (1981) that the futures prices are a Martingale under the equivalent measure $\widetilde{\mathcal{Q}}$:

$$\begin{aligned} F(t, T_F, T_B) &= \mathbb{E}_t^{\widetilde{\mathcal{Q}}}\big[F(T_F, T_F, T_B)|\mathcal{F}_t\big] \\ &= \mathbb{E}_t^{\widetilde{\mathcal{Q}}}\big[P(T_F, T_B)|\mathcal{F}_t\big] \\ &= \mathbb{E}_t^{\widetilde{\mathcal{Q}}}\big[\exp(B(T_F, T_B))|\mathcal{F}_t\big] \end{aligned}$$

We know that under $\widetilde{Q}$

$$\begin{aligned} B(T_F, T_B) &= -\int_{T_F}^{T_B} f(T_F, s)ds \\ &= -\int_{T_F}^{T_B} f(0, s)ds - \sum_i \int_{T_F}^{T_B}\int_0^{T_F} \sigma_i(u, s)\int_u^s \sigma_i(u, v)dv\,du\,ds \\ &\quad - \sum_i \int_{T_F}^{T_B}\left(\int_0^{T_F} \sigma_i(u, s)d\widetilde{W}_i(u)\right)ds \end{aligned}$$

By an application of the stochastic Fubini theorem

$$\begin{aligned} B(T_F, T_B) = &-\int_{T_F}^{T_B} f(0, s)ds - \sum_i \int_0^{T_F}\int_{T_F}^{T_B} \sigma_i(u, s)\int_u^s \sigma_i(u, v)dv\,ds\,du \\ &- \sum_i \int_0^{T_F}\left(\int_{T_F}^{T_B} \sigma_i(u, s)ds\right)d\widetilde{W}_i(u) \end{aligned}$$

Therefore[13]

$$F(t,T_F,T_B) = \exp\Bigg[-\int_{T_F}^{T_B} f(0,s)ds - \sum_i \int_0^{T_F}\int_{T_F}^{T_B} \sigma_i(u,s)\int_u^s \sigma_i(u,v)dv\,ds\,du$$
$$- \sum_i \int_0^t\int_{T_F}^{T_B} \sigma_i(u,s)ds\,d\widetilde{W}_i(u) + \frac{1}{2}\sum_i \int_t^{T_F}\left(\int_{T_F}^{T_B}\sigma_i(u,s)ds\right)^2 du\Bigg]$$

Using the expansion obtained as a result of substituting $t = 0$, the above formula can be reduced to

$$F(t,T_F,T_B) = F(0,T_F,T_B)\exp\Bigg[-\frac{1}{2}\sum_i \int_0^t \left(\int_{T_F}^{T_B}\sigma_i(u,s)ds\right)^2 du$$
$$- \sum_i \int_0^t\int_{T_F}^{T_B}\sigma_i(u,s)ds\,d\widetilde{W}_i(u)\Bigg]$$

Taking stochastic differentials gives the stochastic differential equation for $F(t,T_F,T_B)$ as (2.6) in the text:

$$\frac{dF(t,T_F,T_B)}{F(t,T_F,T_B)} = -\sum_i \int_{T_F}^{T_B}\sigma_i(t,s)ds\,d\widetilde{W}_i(t)$$

Appendix B. Variance of futures yield

To ease the notation set

$$\theta_i(t) = \int_{T_F}^{T_B}\sigma_i(t,s)ds.$$

It we set

$$z(t,T_F,T_B) = \frac{1}{T_B - T_F} + y(t,T_F,T_B)$$

then $\mathrm{var}(y(t,T_F,T_B)) = \mathrm{var}(z(t,T_F,T_B))$ and the stochastic differential equation (2.7) can be written as

$$dz(t,T_F,T_B) = \sum_i \theta_i^2(t)z(t,T_F,T_B)dt + z(t,T_F,T_B)\sum_i \theta_i(t)d\widetilde{W}_i(t).$$

With a view to calculating $\mathbb{E}_0[z(t,T_F,T_B)]$ and $\mathrm{var}_0[z(t,T_F,T_B)]$ we set

$$m(t) = \ln z(t,T_F,T_B) \tag{B.1}$$

and

$$n(t) = \ln\big(z(t,T_F,T_B)\big)^2 = 2m(t). \tag{B.2}$$

[13]We remind the reader that at time t the integral $\int_0^t \left(\int_{T_F}^{T_B}\sigma_i(u,s)ds\right)d\widetilde{W}_i(u)$ is a realized quantity

Application of Itô's lemma to (B.1) followed by an integration yields

$$m(t) = m(0) + \frac{1}{2}\sum_i \int_0^t \theta_i^2(u)du + \sum_i \int_0^t \theta_i(u)d\widetilde{W}_i(u) \tag{B.3}$$

and it follows from (B.2) that

$$n(t) = 2m(0) + \sum_i \int_0^t \theta_i^2(u)du + 2\sum_i \int_0^t \theta_i(u)d\widetilde{W}_i(u) \tag{B.4}$$

Since we assume that the volatility functions $\sigma_i(t,s)$ are deterministic functions of time, it follows that the $\theta_i(t)$ are deterministic functions of time. Hence (B.3) and (B.4) imply that both $m(t)$ and $n(t)$ are normally distributed and we readily calculate that

$$m(t) \sim \mathcal{N}\left(m(0) + \frac{1}{2}\sum_i \int_0^t \theta_i^2(u)du \ , \ \sum_i \int_o^t \theta_i^2(u)du\right), \tag{B.5}$$

and

$$n(t) \sim \mathcal{N}\left(2m(0) + \sum_i \int_0^t \theta_i^2(u)du \ , \ 4\sum_i \int_o^t \theta_i^2(u)du\right). \tag{B.6}$$

We recall that if a random variable $v(t)$ is distributed $\mathcal{N}\left(\mu(t), \sigma^2(t)\right)$ then

$$\mathbb{E}\left[\mathrm{e}^{v(t)}\right] = \mathrm{e}^{\mu(t)+\frac{1}{2}\sigma^2(t)}$$

Using this result we calculate from (B.5) and (B.6) that

$$\mathbb{E}_0\left[z(t,T_F,T_B)\right] = \mathbb{E}_0\left[\mathrm{e}^{m(t)}\right] = \exp\left(m(0) + \sum_i \int_0^t \theta_i^2(u)du\right) \tag{B.7}$$

and

$$\mathbb{E}_0\left[z(t,T_F,T_B)^2\right] = \mathbb{E}_0\left[\mathrm{e}^{n(t)}\right] = \exp\left(2m(0) + 3\sum_i \int_0^t \theta_i^2(u)du\right) \tag{B.8}$$

Using (B.7) and (B.8) and the relationship

$$\begin{aligned} \mathrm{var}\left[y(t,T_F,T_B)\right] &= \mathrm{var}\left[z(t,T_F,T_B)\right] \\ &= \mathbb{E}_0\left[z(t,T_F,T_B)^2\right] - \Big(\mathbb{E}_0\left[z(t,T_F,T_B)\right]\Big)^2 \end{aligned}$$

equation (2.7) is readily derived.

APPENDIX C. FIXED-MATURITY FORWARD RATE EVOLUTION

Consider an investor who holds a bond maturing at T_F and seek the return he or she would earn between T_F and $T_B(> T_F)$, if he or she contracted now at time t. The required rate of return is the discrete forward rate $f(t, T_F, T_B)$ defined by

$$P(t, T_F) = P(t, T_B) \exp\big(f(t, T_F, T_B)(T_B - T_F)\big)$$

ie.

$$\begin{aligned} f(t, T_F, T_B) &= \frac{1}{T_B - T_F} \ln\left[\frac{P(t, T_F)}{P(t, T_B)}\right] \\ &= \frac{1}{T_B - T_F} \int_{T_F}^{T_B} f(t, s) ds \end{aligned}$$

Recall that the evolution of the instantaneous forward rate is

$$f(t, T_B) = f(0, T_B) + \sum_i \left[\int_0^t \sigma_i(u, T_B) \int_u^{T_B} \sigma_i(u, v) dv\, du + \int_0^t \sigma_i(u, T_B) d\widetilde{W}_i(u)\right]$$

Therefore, the discrete forward rate $f(t, T_F, T_B)$ evolves according to

$$\begin{aligned} f(t, T_F, T_B) = {} & \frac{1}{T_B - T_F} \int_{T_F}^{T_B} f(0, s) ds + \frac{1}{T_B - T_F} \sum_i \int_{T_F}^{T_B} \int_0^t \sigma_i(u, s) \int_u^s \sigma_i(u, v) dv\, du\, ds \\ & + \frac{1}{T_B - T_F} \int_{T_F}^{T_B} \int_0^t \sigma_i(u, s) d\widetilde{W}_i(u)\, ds \\ = {} & \frac{1}{T_B - T_F} \int_{T_F}^{T_B} f(0, s) ds + \frac{1}{T_B - T_F} \sum_i \int_0^t \int_{T_F}^{T_B} \sigma_i(u, s) \int_u^s \sigma_i(u, v) dv\, ds\, du \\ & + \frac{1}{T_B - T_F} \int_0^t \int_{T_F}^{T_B} \sigma_i(u, s) ds\, d\widetilde{W}_i(u) \end{aligned}$$

The conditional variance of the fixed-maturity forward rate is thus readily calculated as

$$\operatorname{var}\big(f(t_j, T_F, T_B) | f(t_{j-1}, T_F, T_B)\big) = \frac{1}{(T_B - T_F)^2} \sum_i \int_{t_{j-1}}^{t_j} \left(\int_{T_F}^{T_B} \sigma_i(u, s) ds\right)^2 du$$

APPENDIX D. FULL INFORMATION LOG LIKELIHOOD FUNCTION FOR QUOTED FUTURES PRICES

The main task in deriving the log likelihood function is to calculate the Jacobian of the transformation and write out the drift vector and covariance matrix for each transition log likelihood function. These quantities then can be substituted directly to the formula in the text (equations 3.9, 3.11, 3.12 and 3.13) to write out the likelihood function for observable futures prices.

From 3.4

$$X_{jk} = \ln(F_{jk}) \equiv \zeta(F_{jk})$$

we have

$$\frac{\partial \zeta(F_{jk};\boldsymbol{\theta})}{\partial F_{jk}} = \frac{1}{F_{jk}}$$

From (3.2)

$$F_{jk} = \frac{1}{1+\left(1-\frac{G_{jk}}{100}\right)\tau} \equiv \eta\big(G_{jk}\big),$$

where $\tau = 90/360$ for CME Eurodollar futures, we find that

$$\frac{\partial \eta(G_{jk};\theta)}{\partial G_{jk}} = \frac{-\frac{\tau}{100}}{\left[1+\left(1-\frac{G_{jk}}{100}\right)\tau\right]^2}$$

The variance

$$\begin{aligned}\beta_{j(kk)} &= \int_{t_{j-1}}^{t_j}\left(\int_{T_{F_k}}^{T_{B_k}} \sigma(u,s)ds\right)^2 du + \int_{t_{j-1}}^{t_j} \sigma_\varepsilon^2 du \\ &= M^2 I_{00} + 2MNI_{01} + N^2 I_{02} + 2MRI_{11} \\ &\quad + 2NRI_{12} + R^2 I_{22} + \sigma_\varepsilon^2 (t_j - t_{j-1}),\end{aligned}$$

where

$$\begin{aligned}M &= \sigma_0(T_{Bk} - T_{Fk}) \\ N &= -\left(\frac{\sigma_0}{\kappa} + \frac{\sigma_1}{\kappa^2}\right)\left(\mathrm{e}^{-\kappa T_{Bk}} - \mathrm{e}^{-\kappa T_{Fk}}\right) - \frac{\sigma_1}{\kappa}\left(T_{Bk}\,\mathrm{e}^{-\kappa T_{Bk}} - T_{Fk}\,\mathrm{e}^{-\kappa T_{Fk}}\right) \\ R &= \frac{\sigma_1}{\kappa}\left(\mathrm{e}^{-\kappa T_{Bk}} - \mathrm{e}^{-\kappa T_{Fk}}\right) \\ I_{ab} &= \int_{t_{j-1}}^{t_j} \tau^a\,\mathrm{e}^{\kappa b\tau}\,d\tau \\ &= \Bigg(-\,\mathrm{e}^{\kappa b\tau}\Bigg[\frac{1}{(-\kappa b)}\tau^a + \frac{a}{(-\kappa b)^2}\tau^{a-1} + \frac{a(a-1)}{(-\kappa b)^3}\tau^{a-2} + \dots \\ &\qquad \dots + \frac{a(a-1)\dots 2}{(-\kappa b)^a}\tau + \frac{a(a-1)\dots 1}{(-\kappa b)^{a+1}}\tau^0\Bigg]\Bigg)\,\Bigg|_{t_{j-1}}^{t_j}\end{aligned}$$

The covariance (where $k_1 \neq k_2$)

$$\begin{aligned}\beta_{j(k_1k_2)} &= \int_{t_{j-1}}^{t_j}\left(\int_{T_{Fk_1}}^{T_{Bk_1}} \sigma(u,s)ds\right)\left(\int_{T_{Fk_2}}^{T_{Bk_2}} \sigma(u,s)ds\right) du \\ &= M_1M_2I_{00} + (M_1N_2 + N_1M_2)I_{01} + N_1N_2I_{02} \\ &\quad + (M_1R_2 + R_1M_2)I_{11} + (N_1R_2 + R_1N_2)I_{12} + R_1R_2I_{22}\end{aligned}$$

where

$$M_1 = \sigma_0(T_{Bk_1} - T_{Fk_1})$$
$$M_2 = \sigma_0(T_{Bk_2} - T_{Fk_2})$$
$$N_1 = -\left(\frac{\sigma_0}{\kappa} + \frac{\sigma_1}{\kappa^2}\right)\left(\mathrm{e}^{-\kappa T_{Bk_1}} - \mathrm{e}^{-\kappa T_{Fk_1}}\right) - \frac{\sigma_1}{\kappa}\left(T_{Bk_1}\,\mathrm{e}^{-\kappa T_{Bk_1}} - T_{Fk_1}\,\mathrm{e}^{-\kappa T_{Fk_1}}\right)$$
$$N_2 = -\left(\frac{\sigma_0}{\kappa} + \frac{\sigma_1}{\kappa^2}\right)\left(\mathrm{e}^{-\kappa T_{Bk_2}} - \mathrm{e}^{-\kappa T_{Fk_2}}\right) - \frac{\sigma_1}{\kappa}\left(T_{Bk_2}\,\mathrm{e}^{-\kappa T_{Bk_2}} - T_{Fk_2}\,\mathrm{e}^{-\kappa T_{Fk_2}}\right)$$
$$R_1 = \frac{\sigma_1}{\kappa}\left(\mathrm{e}^{-\kappa T_{Bk_1}} - \mathrm{e}^{-\kappa T_{Fk_1}}\right)$$
$$R_2 = \frac{\sigma_1}{\kappa}\left(\mathrm{e}^{-\kappa T_{Bk_2}} - \mathrm{e}^{-\kappa T_{Fk_2}}\right)$$

and I_{ab} are defined as in the variance formulae.

The drift term is equal to minus a half of the corresponding variance term.

REFERENCES

Aït-Sahalia, Y. (2001), 'Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed-Form Approximation Approach', *Econometrica* . Forthcoming.

Amin, K. I. and Morton, A. J. (1994), 'Implied Volatility Functions in Arbitrage-free Term Structure Models', *Journal of Financial Economics* **35**, 141–180.

Amin, K. I. and Ng, V. K. (1997), 'Inferring Future Volatility from the Information in Implied Volatility in Eurodollar Options: A New Approach', *Review of Financial Studies* **10**(2), 333–367.

Bhar, R. and Chiarella, C. (1997*a*), 'Interest Rate Futures: Estimation of Volatility Parameters in an Arbitrage-Free Framework', *Applied Mathematical Finance* **4**(4), 181–199.

Bhar, R. and Chiarella, C. (1997*b*), 'Transformation of Heath-Jarrow-Morton models to Markovian Systems', *European Journal of Finance* **3**, 1–26.

Bjork, T. and Svensson, L. (2001), 'On the Existence of Finite Dimensional Realizations for Nonlinear Forward Rate Models', *Mathematical Finance* **11**, 205–243.

Bliss, R. R. and Ritchken, P. (1996), 'Empirical Tests of Two State-Variable Heath-Jarrow-Morton Models', *Journal of Money, Credit and Banking* **28**(3), 452–476.

Brennan, M. J. and Schwartz, E. S. (1979), 'A Continuous Time Approach to the Pricing of Bonds', *Journal of Banking and Finance* **3**, 133–155.

Brennan, M. J. and Schwartz, E. S. (1982), 'An Equilibrium Model of Bond Pricing and a test of market efficiency', *Journal of Financial and Quantitative Analysis* **17**, 301–329.

Bühler, W., Uhrig-Homburg, M., Walter, U. and Webber, T. (1999), 'An Empirical Comparison of Forward-Rate and Spot Rate Models for Valuing Interest-Rate Options', *Journal of Finance* **54**(1), 269–305.

Chapman, D. A., Long, J. B. J. and Pearson, N. D. (1999), 'Using Proxies for the Short Rate: When Are Three Months Like an Instant', *Review of Financial Studies* **12**(4), 763–806.

Chiarella, C. and Kwon, O.-K. (2001a), 'Formulation of Popular Interest Models under the HJM Framework', *Asia Pacific Financial Markets* **8**, 1–22.

Chiarella, C. and Kwon, O.-K. (2001b), 'Forward Rate Dependent Markovian Transformations of the Heath-Jarrow-Morton Term Structure Model', *Finance and Stochastics* **5**(2), 237–257.

Cox, J. C., Ingersoll, J. and Ross, S. A. (1981), 'The Relationship Between Forward Prices and Futures Prices', *Journal of Financial Economics* **9**, 321–346.

Cox, J. C., Ingersoll, J. and Ross, S. A. (1985), 'A Theory of the Term Structure of Interest Rates', *Econometrica* **53**(2), 385–407.

D'Agostino, R. B. (1970), 'Transformation to Normality of the null distribution of g_1', *Biometrika* **57**, 679–681.

De Jong, F. and Santa-Clara, P. (1999), 'The Dynamics of the Forward Interest Rate Curve: A Formulation with State Variables', *Journal of Financial and Quantitative Analysis* **34**(1), 131–157.

Doornik, J. A. and Hansen, H. (1994), 'An Omnibus Test for Univariate and Multivariate Normality'. Discussion pape, Nuffield College, Oxford.

Duan, J.-C. (1994), 'Maximum Likelihood Estimation Using Price Data of the Derivative Contract', *Mathematical Finance* **4**(2), 155–167. (2000) 'Correction: Maximum Likelihood Estimation Using Price Data of the Derivative Contract' *Mathematical Finance*, **10**(4), 461-462.

Dybvig, P. H. (1990), Bond and Bond Option pricing Based on the Current Term Structure, Technical report, Washington University in St.Louis, St.Louis, MO.

Flesaker, B. (1993), 'Testing the Heath-Jarrow-Morton/Ho-Lee Model of Interest Rate Contingent Claim Pricing', *Journal of Financial and Quantitative Analysis* **28**(4), 483–495.

Ho, L.-C., Cadle, J. and Theobald, M. (2001), 'Estimation and Hedging with a One-Factor Heath-Jarrow-Morton Model', *Journal of Derivatives* **8**(4), 49–61.

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**(4), 573–592.

Hull, J. and White, A. (1996), 'Using Hull-White interest-rate trees', *Journal of Derivatives* **3**(3), 26–36.

Inui, K. and Kijima, M. (1998), 'A Markovian Framework in Multi-Factor Heath-Jarrow-Morton Models', *Journal of Financial and Quantitative Analysis* **33**(3), 423–440.

Lo, A. W. (1988), 'Maximum Likelihood Estimation of Generalized Itô Processes with Discretely Sampled Data', *Econometric Theory* **4**(2), 231–247.

Musiela, M., Turnbull, S. M. and Wakeman, L. M. (1992), 'Interest Rate Risk Management', *Review of Futures Market* **12**(2), 221–261.

Pearson, N. D. and Sun, T.-S. (1994), 'Exploiting the Conditional Density in Estimating the Term Structure: An Application to the Cox, Ingersoll, and Ross Model', *Journal of Finance* **49**(4), 1279–1304.

Raj, M., Sim, A.-B. and Thurston, D. C. (1997), 'A Generalized Method of Moments Comparison of the Cox-Ingersoll-Ross and Heath-Jarrow-Morton Models', *Journal of Economics and Business* **49**(2), 169–192.

Ritchken, P. and Chuang, I. (1999), 'Interest Rate Option Pricing with Volatility Humps', *Review of Derivatives Research* **3**(3), 237–262.

Ritchken, P. and Sankarasubramanian, L. (1995), 'Volatility Structures of Forward Rates and the Dynamics of the Term Structure', *Mathematical Finance* **5**(1), 55–72.

Shenton, L. R. and Bowman, K. O. (1977), 'A Bivariate Model for the Distribution of $\sqrt{b_1}$ and b_2', *Journal of theAmerican Statistical Association* **72**, 206–211.

Vasicek, O. (1977), 'An Equilibrium Characterization of the Term Structure', *Journal of Financial Economics* **5**(2), 177–188.