Design and Estimation of Quadratic Term Structure Models*

Markus Leippold†

Swiss Banking Institute, University of Zurich

Liuren Wu‡

Graduate School of Business, Fordham University

June 14, 2001

*We thank Marco Avellaneda, David Backus, Peter Carr, Pierre Collin, Silverio Foresi, Michael Gallmeyer, Richard Green, Massoud Heidari, Burton Hollifield, Chris Telmer, and Stanley Zin for helpful comments. We welcome comments, including references to related papers we inadvertently overlooked.

†**Correspondence Information:** Plattenstr. 14, 8032 Zurich, Switzerland; tel: (+41) 1-634-2951; fax: (+41) 1-634-4903; `leippold@isb.unizh.ch`.

‡**Correspondence Information:** 113 West 60th Street, New York, NY 10023; tel: (212) 636-6117; fax: (212)765-5573; `wu@fordham.edu`.

Design and Estimation of Quadratic Term Structure Models

ABSTRACT

We consider the design and estimation of quadratic term structure models. We start with a list of stylized facts on interest rates and interest rate derivatives, classified into three layers: (1) general statistical properties, (2) forecasting relations, and (3) conditional dynamics. We then investigate the implications of each layer of property on model design and strive to establish a mapping between evidence and model structures. We calibrate a two-factor model that approximates these three layers of properties well, and illustrate how the model can be applied to pricing interest rate derivatives.

JEL Classification Codes: G12, G13, E43.

Keywords: quadratic model; term structure; positive interest rates; humps; expectation hypothesis; GMM; caps and floors.

Term structure modeling has enjoyed rapid growth during the last decade. Among the vast number of different models, the affine class stands out as the most popular class due to its analytical tractability. Duffie and Kan (1996)'s characterization of the complete class has spurred a stream of studies on its empirical applications and model design. Examples include general econometric estimations by Chen and Scott (1993), Duffie and Singleton (1997), Dai and Singleton (2000b), and Singleton (1999), applications to the predictability of interest rates by Frachot and Lesne (1994), Roberds and Whiteman (1999), Backus, Foresi, Mozumdar, and Wu (2001), Duffee (1999), and Dai and Singleton (2000a), and currency pricing by Backus, Foresi, and Telmer (2001). While these applications claim success in one or two dimensions, inherent tension exists when one tries to apply the model to a wider range of properties. Even more troublesome, however, is a seemingly irreconcilable tension between delivering a relatively good empirical performance and excluding negative interest rates. Indeed, all of the relatively "successful" model designs within the affine framework, in terms of empirical performance, imply positive probabilities of negative interest rates. Examples include Backus, Foresi, Mozumdar, and Wu (2001), Backus, Foresi, and Telmer (2001), Dai and Singleton (2000b), Dai and Singleton (2000a), Duffee (1999), Duffie and Singleton (1997), and Singleton (1999).

Leippold and Wu (1999a) identify and characterize an alternative class, the quadratic class of term structure models, where bond yields and forward rates are quadratic functions of the state vector. Their property analysis indicates that the quadratic class is comparable to the affine class for analytical tractability but is more flexible for model design. In particular, positive interest rates can be guaranteed with little loss of generality or flexibility. In this paper, we consider the model design and estimation problem within the quadratic framework.

Although examples of quadratic models have appeared in the literature since the late eighties,[1] empirical applications have at best been sporadic. The most systematic empirical study, and hence the most germane to our work, is by Ahn, Dittmar, and Gallant (2001).

[1]Early examples include Longstaff (1989), Beaglehole and Tenney (1991), Beaglehole and Tenney (1992), El Karoui, Myneni, and Viswanathan (1992), Constantinides (1992) Jamshidian (1996), Rogers (1997), Ahn (1998), Leippold and Wu (1999b), and Leippold and Wu (1999a).

They apply the efficient methods of moments (EMM) of Gallant and Tauchen (1996), calibrate the maximally flexible three-factor quadratic model and various restricted versions to the US Treasury data, and compare their performance with that of affine models. Our approach goes in the opposite direction. We start with a list of what we view as the salient features of interest rates and interest rate derivatives and attempt to find a parsimonious quadratic specification which accounts for them. What we gain are some helpful insights into the mapping between parameters and data. The approach also highlights the impact of different pieces of evidence on model structure. Furthermore, while Ahn, Dittmar, and Gallant (2001) focus on the time series property of the interest rate data, we also consider the model's application in option pricing. We apply the transform method proposed in Leippold and Wu (1999a) to price interest rate caps and investigate the implications of the salient features of the interest rate derivatives on model design.

We take a series of steps that we think serve to integrate evidence with theory. In the first step, we classify the properties of interest rates into three categories: (1) general statistical properties, (2) forecasting relations, and (3) conditional dynamics. The most significant statistical properties, in our view, include an upward sloping mean yield curve, high (but different) persistence in yields of different maturities, and positive skewness in interest rate distributions. For forecasting relations, we examine the violations of the various forms of the expectation hypotheses (EH). We find that underlying all interest rates is a common feature revealed by a single-period forward rate regression. That is, violations of the EH are mainly at the short end. However, even at the short end, the EH violation is not statistically significant for eurodollar interest rates. Overall, the evidence on EH violations is much weaker for eurocurrencies than for US Treasuries. On conditional dynamics, we find that the mean hump-shaped term structure for conditional variance not only shows up in the implied volatility quotes of interest rate derivatives, but also reveals itself vividly in the variance of multi-period changes. We document the evidence using interest rate and interest rate derivatives (caps and floors) data on eurocurrencies (US dollar and Deutsche mark), but similar stylized behaviors have also been observed in US Treasuries and in other currencies.

With facts in hand, we turn to the model. We analyze the properties of bond yields and forward rates under the quadratic class and examine the implications of the documented evidence on model design. The quadratic relation between interest rates and the state variables not only provides a convenient approach to guarantee positive interest rates, but also directly incorporate nonlinearity in interest rate dynamics. Furthermore, the affine specification on market price of risk plays an important role in accounting for the violations in the expectation hypotheses while flexible interactions between state variables are indispensable in generating the observed hump-shaped dynamics. Finally, the affine specification in market price of risk also governs the difference in conditional dynamics between the objective measure and the risk neutral measure. It links and distinguishes the hump dynamics observed in the time series and that observed in implied volatility quotes of interest rate derivatives.

In the third step, we transform the property analysis into moment conditions and calibrate a two-factor quadratic model by generalized methods of moments (GMM). While three or even more factors might be necessary to fully capture the interest rate dynamics [e.g. Litterman and Scheinkman (1991), Knez, Litterman, and Scheinkman (1994)], a two-factor model is the minimum required to capture the stylized evidence listed above. The calibration exercise confirms with the property analysis on the relative contribution of each component of the model to different features of interest rates. We find that fitting the large hump shape observed in the annualized variance of multiperiod interest rate differences asks for strong interaction between state variables; the weak EH violation for eurodollar interest rates, on the other hand, reduces the significance of the estimates for the affine part of the market price of risk.

In the final step, given the parameter estimates, we explore the model's implications on pricing interest rate derivatives. Under the quadratic model, the implied volatility is a biased estimator of the conditional variance rate of interest rates under the risk neutral measure. The bias is mainly a result of the violation of the Black model's assumption on conditional log-normality on simply compounded LIBOR rates. The magnitude of the bias differs across strikes, resulting in the famous implied volatility smile across strikes. Nevertheless, for at-the-money spot interest rate caps, the term structure of the implied volatility quotes approximates the term structure of the conditional variance of interest rates. In particular, a hump in the

term structure of conditional variance transfers to a hump in the term structure of implied volatility.

Recently, Filipović (2001) proves, under certain regularity conditions, that if one represents the forward rate as a polynomial function of the diffusion state vector, then the maximal order of the polynomial is two for the model to be consistent. Consistency in this context, as discussed as in Björk and Christensen (1999) and Filipović (2000), means that the interest rate model will produce forward rate curves belonging to the parameterized family. Thus, the affine class of Duffie and Kan (1996) and the quadratic class of Leippold and Wu (1999a) essentially complete the search for consistent polynomial term structure models. This is a remarkable result as many functional forms can be approximated by a polynomial series. It also stresses the potentially important role played by quadratic models.

The structure of the paper is as follows. The next section documents the evidence using more than 10 years of interest rates on eurodollar and euromark, as well as five years of data on interest rate caps and floors. Section II lays out the framework for the quadratic term structure models. Section III analyzes the implications of the quadratic model on the three layers of evidence. Section IV transforms the analysis into moment conditions, calibrates a two-factor quadratic model by generalized methods of moments, and tests various parameter restrictions. Section V investigates the implications of the model estimates on the cross-sectional and term structure behavior of cap implied volatilities. Section VI concludes.

I. Evidence

We document the salient features of interest rates and interest rate derivatives on two eurocurrencies: US dollar and Deutsche mark. We collect from Datastream monthly LIBOR (London Inter Bank Offer Rate) rates with one, three, six, and twelve months of maturity, and swap rates with two, three, four, five, seven and ten years of maturity. We then extract the term

structure of spot rates at each date by fitting the extended Nelson-Siegel function to the spot rate:

$$y_n = \beta_0 + \beta_1 \frac{1 - e^{-n/\tau_1}}{n/\tau_1} + \beta_2 \left(\frac{1 - e^{n/\tau_2}}{n/\tau_2} - e^{-n/\tau_2} \right),$$

where y_n denotes the spot rate with maturity n and $\Theta = [\beta_0, \beta_1, \beta_2, \tau_1, \tau_2]^\top$ is the parameter vector, which is obtained by minimizing the sum square errors between the observed and implied LIBOR and swap rates. The day counting conventions for LIBOR rates are actual over 360 for US dollar and Deutsche mark while that the swap rates is 30/360 with semi-annual compounding payments for US dollar and annual compounding for Deutsche mark. The data starts in April 1987 and ends in September 2000 (162 observations).

In discrete time notation, we denote the continuously compounded spot rate on an n-period bond at date t as y_t^n. It is defined as

$$y_t^n = -n^{-1} \ln P_t^n,$$

where P_t^n denotes the dollar price at date t of a claim to one dollar at $t + n$. Here the discrete period is in number of months, corresponding to the monthly data. Forward rates are defined by

$$f_t^n = \ln \left(P_t^n / P_t^{n+1} \right),$$

so that spot rates are averages of forward rates:

$$y_t^n = n^{-1} \sum_{i=0}^{n-1} f_t^i.$$

We use the one-month spot rate as a proxy for the instantaneous interest rate, or the short rate: $r_t = y_t^1 = f_t^0$.

We have also downloaded from Datastream the daily implied volatility quotes for at-the-money spot caps on the two eurocurrency LIBOR rates with maturities of one, two, three, four, five, seven, and ten years. The data are from February 1st, 1995 to October 17th, 2000

(1490 observations). A cap consists of a series of caplets. Let R denote the simply compounded LIBOR rate. The payoff of the ith caplet is given by

$$\Pi^i = \tau L \left(R_{t+i\tau} - K\right)^+$$

where τ is the tenor (payment interval) of the cap contract, it is six month ($\tau = 1/2$) for dollar LIBOR and one year ($\tau = 1$) for mark. L denotes the notional amount of the contract. K is the strike rate and is set equal to the current LIBOR rate R_t for at-the-money spot contracts. The payment is made is *in arrears*, i.e. the payment of the ith caplet is made at $t + (i+1)\tau$. The implied volatility quotes are obtained under the framework of the Black model, where the LIBOR rate is assumed to follow a geometric Brownian motion with constant diffusion v. The price of a cap contract with maturity N (periods) under such a model is given by

$$C_t^N = \sum_{i=1}^{N-1} c_i = \tau \sum_i P_t^{i+1} \left[F_t^i N\left(d_{1i}\right) - K N\left(d_{2i}\right)\right],$$

where F_t^i is the simply compounded forward rate between the ith and $(i+1)$th payment period

$$1 + \tau F_t^i = P_t^i / P_t^{i+1}$$

and

$$d_{1i} = \frac{\ln F_t^i / K + v^2 (i\tau)/2}{\sqrt{i\tau} v}, \quad d_{2i} = d_{1i} - \sqrt{i\tau} v.$$

The implied volatility is an estimate for v such that the cap price implied by the Black formula is equal to the market price.

A. Statistical Properties of Interest Rates

Table I provides the summary statistics of the spot rates on the two eurocurrencies. The most significant features include:

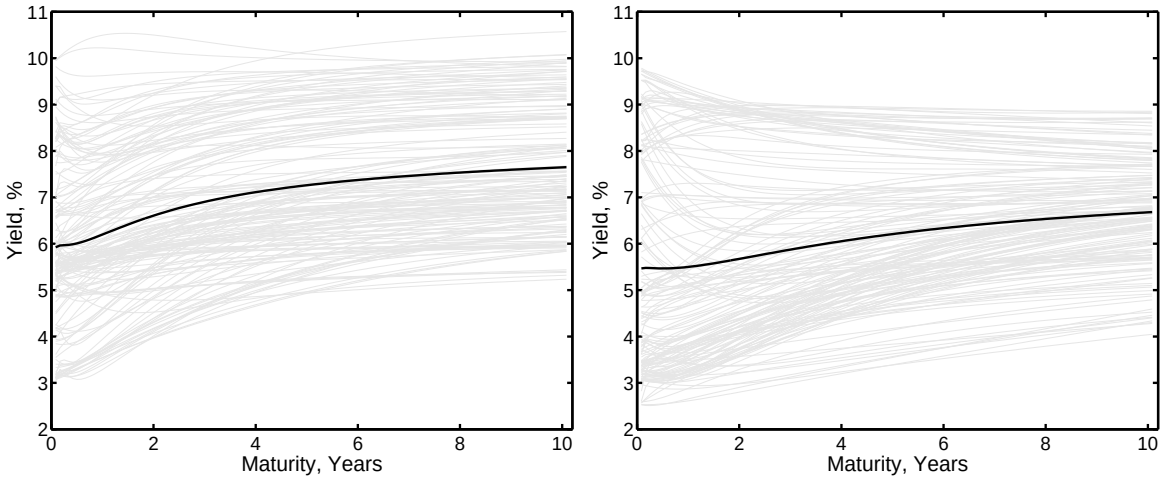

Figure 1. Term Structure of Interest Rates on Eurocurrencies
The gray lines in the background are term structures of interest rate on two eurocurrencies (US dollar in the left panel, Deutsche mark in the left panel) at each month from April 1987 to September 2000 (162 observations). The bold solid lines represent the mean yield curve.

1. While at each date, the term structure exhibits various shapes, the mean term structure is upward sloping. Figure 1 depicts the term structure at each date in the background (gray lines) and the mean term structure with a bold solid line.
2. Bond yields are highly persistent and the persistence varies across maturities. The first-order (monthly) autocorrelations range from 0.972 to 0.985 for the dollar rates and from 0.980 to 0.992 for the mark rates.
3. The past decade has been relatively uneventful for both interest rates. While the interest rates are positively skewed, the tails are no thicker than that of a normal distribution.

B. Violations of Expectation Hypotheses

A long established fact about interest rates is that the current term structure contains information about future interest rate movements. While the expectation hypothesis (EH) has long been regarded as a poor approximation of the evidence, it presents useful forms for interest rate forecasting. An enormous body of research to this effect has been surveyed repeatedly, most recently by Bekaert, Hodrick, and Marshall (1997), Campbell (1995), Campbell and Shiller (1991), and Evans and Lewis (1994). Recently, Frachot and Lesne (1994), Roberds and

Whiteman (1999), Backus, Foresi, Mozumdar, and Wu (2001), Duffee (1999), and Dai and Singleton (2000a) have tried to explain the evidence within the affine framework of Duffie and Kan (1996).

Many forecasting relations have been formulated based on various forms of the expectation hypothesis. The most common ones include: (i) one-period forward rate regression, (ii) one-period yield regression, (iii) multiperiod forward rate regression, and (iv) multi-period yield regression. They can be denoted, respectively, as

$$\begin{aligned}
f_{t+1}^{n-1} - r_t &= a_n + c_n \left(f_t^n - r_t\right) + e_{t+1}^n, \\
y_{t+1}^n - y_t^n &= a_n + d_n \left(\frac{y_t^{n+1} - r_t}{n}\right) + e_{t+1}^n, \\
r_{t+n} - r_t &= a_n + e_n \left(f_t^n - r_t\right) + e_{t+n}, \\
\sum_{i=1}^{n} \left(1 - \frac{i}{n+1}\right) \left(r_{t+i} - r_{t+i-1}\right) &= a_n + g_n \left(y_t^{n+1} - r_t\right) + e_{t+n}.
\end{aligned}$$

We abuse the notation by using the same letter a_n for the intercepts and e for the residuals of all four regressions, as our focus is on the slopes. The one-period forward rate regression is proposed recently by Backus, Foresi, Mozumdar, and Wu (2001). The study of the one-period yield regression dates back at least to Roll (1970). Recent empirical studies include Bekaert, Hodrick, and Marshall (1997), Campbell and Shiller (1991), and Evans and Lewis (1994) for the US and Bekaert, Hodrick, and Marshall (1996) and Hardouvelis (1994) for other countries. The multi-period forward rate regression has been estimated by Fama (1984), Fama and Bliss (1987), and Mishkin (1988). The multi-period yield regression has also been studied by Campbell and Shiller (1991) with US data and Bekaert, Hodrick, and Marshall (1996) with data for the US, UK, and Germany.

Of all the regressions, the single period forward rate regression of Backus, Foresi, Mozumdar, and Wu (2001) possesses the simplest form, based on the martingale property of forward rates under the expectation hypothesis. Figure 2 reports the regression slopes on the two eurocurrencies, together with their 95% confidence band. Regressions on the two eurocurrencies exhibit a similar feature: while deviations from the EH are largest at short maturities, evidence converges to the hypothesis at long maturities.

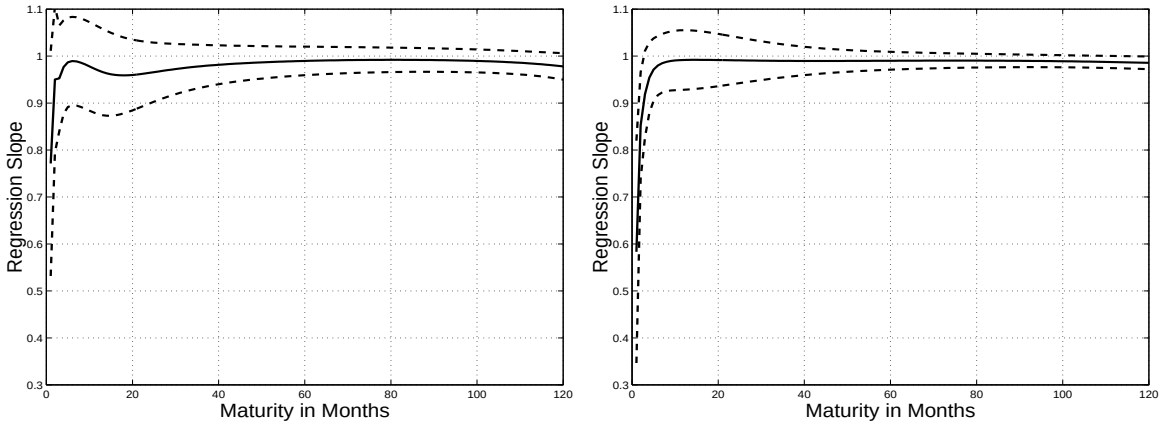

Figure 2. Single Period Forward Rate Regression Slopes
The solid lines depict the slopes of the single period forward rate regression:

$$f_{t+1}^{n-1} - r_t = a_n + c_n \left(f_t^n - r_t \right) + e_{t+1}^n,$$

on euro dollar (left) and mark (right). The dashed lines are the 95% confidence intervals constructed from the standard error estimate for each slope point estimate, with a normal distribution assumption. Standard errors are computed following Newey and West (1987) with six lags.

Backus, Foresi, Mozumdar, and Wu (2001) report similar qualitative properties for US Treasuries; however, the magnitude of the violations of the EH is much more significant in the US Treasuries data than implied by our estimates on eurocurrencies. For example, the one-month regression slope, c_1, is 0.5184 (0.1054) for US Treasuries (standard error in the parentheses),[2] significantly different from the null value of one, but our estimate for eurodollar is 0.7715 (0.1223), not significantly different from one under the 95% confidence level. The estimate for the mark $c_1 = 0.5832$ (0.1207) is significantly different from one; nevertheless, as maturity increases, the estimates for both eurocurrencies converge to one much faster than those for US Treasuries. For example, $c_{60} = 0.9897$ (0.0155) for eurodollar and 0.9901 (0.0097) for euro mark, neither of which is significantly different from unity; yet for US Treasuries, $c_{60} =$ 0.9635 (0.0124), which is significantly different from unity. The two series have mismatching subperiods. As we have the updated data set on US Treasuries, we re-estimate the regression

[2]Numbers for US Treasuries are from the last column of Table 1 in Backus, Foresi, Mozumdar, and Wu (2001).

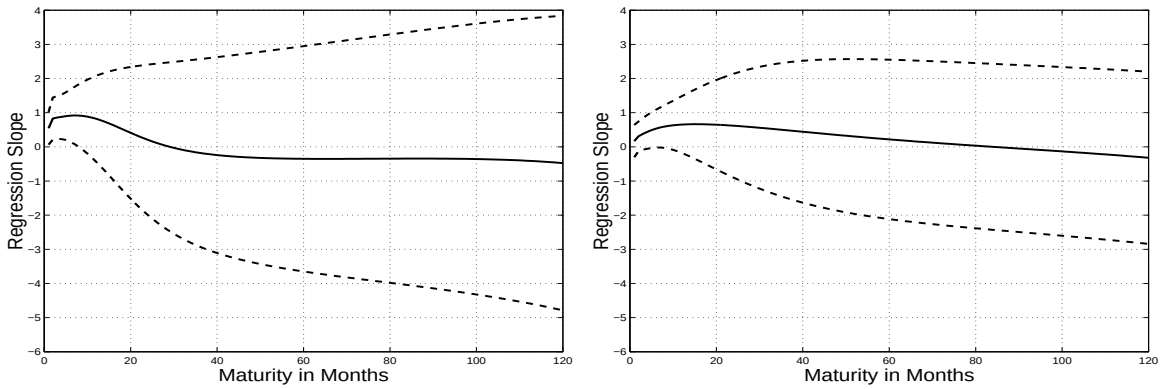

Figure 3. Single Period Yield Regression Slopes
The solid lines depict the slopes of the single period yield regression:

$$y_{t+1}^n - y_t^n \quad = \quad a_n + d_n \left(\frac{y_t^{n+1} - r_t}{n} \right) + e_{t+1}^n,$$

on euro dollar (left) and mark (right). The dashed lines are the 95% confidence intervals constructed from the standard error estimate for each slope point estimate, with a normal distribution assumption. Standard errors are computed following Newey and West (1987) with six lags.

slopes for the US Treasury data with the sample period approximately matching that of our eurocurrencies. The discrepancy observed above remains. In short, the EH violations observed in the eurocurrencies are not as significant as in the US Treasuries.

Longstaff (2000) discovers even more striking results at very short maturities. Using overnight, weekly, and monthly repo rates, he finds that the EH hypothesis cannot be rejected at both the unconditional level (i.e., the mean curve is flat) and the conditional level. He attributes the difference between his result and the US Treasuries to the idiosyncratic behavior of US Treasury bills. Our findings here is not as extreme. The mean yield curves are upward sloping for both eurocurrencies, indicating the existence of at least a constant risk premium. The single period regression slopes are significantly different from the null value of unity at least for euro mark at short maturities, implying that the risk premium may be time varying.

Figure 3 reports the slope estimates for the single period yield regression. In contrast to the forward regression, the deviations of the slope from one increase with maturity; however,

the standard errors also increase with maturity and hence make it hard to explain. The term structures of the slope estimates (not reported) of the multiperiod regressions take on more complicated forms. The accuracy of the estimates also deteriorates as maturity increases due to the significant overlapping. Backus, Foresi, Mozumdar, and Wu (2001) illustrate the inherent links between different forms of EH regressions and argue that they contain similar information content. In particular, they show that how a slight change in the forward regression slope can be transferred into a big swing in the yield regression. In our estimation, we follow Backus, Foresi, Mozumdar, and Wu (2001) and focus on the one-period forward regression.

C. Interest Rate Derivatives and The 'Hump' Dynamics

It has been widely noted that the conditional volatility of interest rates has a "hump-shaped" mean term structure. The conditional volatility first increases with horizon, reaches a plateau, and then decreases. While systematic documentation is rare, casual observations and modeling efforts abound, particularly among practitioners, illustrating its practical importance in risk management and option pricing. Examples include Amin and Morton (1994), Backus and Wu (1998), Bouchaud, Cont, El-Karoui, Potters, and Sagna (1998), Bühler, Uhrig-Homburg, Walter, and Weber (1997), Heath, Jarrow, Morton, and Spindel (1992), Hull and White (1996), Khan (1991), Mercurio and Moraleda (1996), Moraleda and Vorst (1997), and Ritchken and Chuang (1996).

The hump-shaped dynamics shows up in the data in a variety of ways. The most obvious is from the mean term structure of the Black implied volatilities of interest rate caps, floors, and swaptions. Figure 4 depicts the mean term structure of the Black implied volatility quotes for caps on dollar and mark LIBORs. While the volatility level differs across currencies, the mean hump shape is ubiquitous. Similar hump shapes are also observed in the implied volatility of interest rate floors and swaptions, as well as in other eurocurrencies.

A less direct approach comes from multiperiod differences of interest rates: changes $y_t - y_{t-k}$ over periods of length k. If the conditional volatility of y has hump-shaped dynamics, the

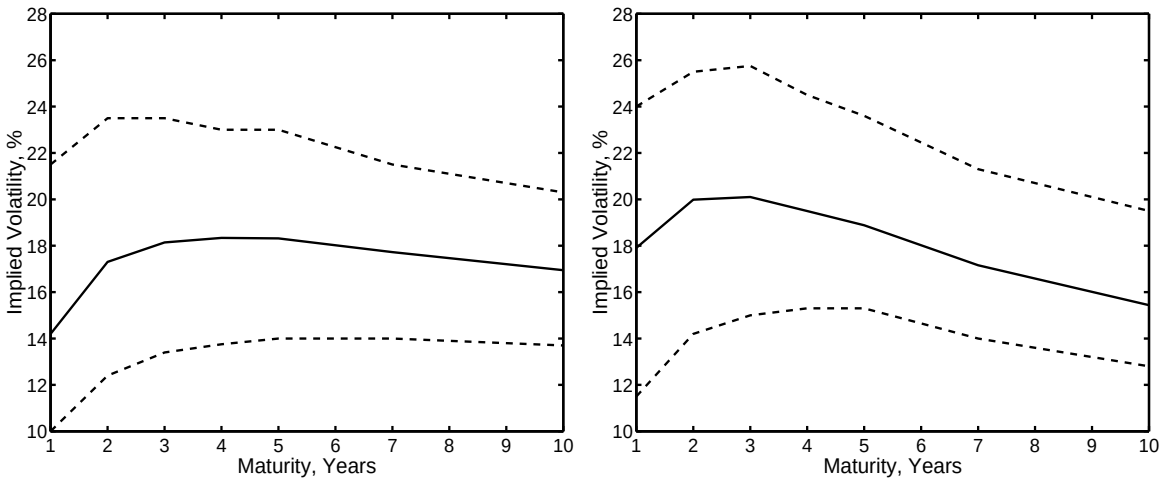

Figure 4. Hump-Shaped Conditional Dynamics In Interest Rate Caps
Solid lines are the mean term structure of at-the-money implied volatility quotes for interest rate caps for U.S. dollar (left) and Deutsche mark (right). Dashed lines depict the 95% and 5% quantiles. The data are from February 1st, 1995 to October 17th, 2000 (1490 observations).

variance of the multiperiod differences increases initially at a rate faster than k.[3] Equivalently, the annualized variance $Var(y_t - y_{t-k})/(kh)$ is hump shaped. Figure 5 depicts the annualized variance over k for the six-month spot rate on the two eurocurrencies, where $h = 1/12$ is the fraction of year per period. The hump shapes are very prominent for both currencies.

While both measures are approximations of the conditional dynamics,[4] the similarities are suggestive. The fact that hump dynamics are observed across different currencies and from different measures implies that it is a robust feature of the interest rate data.

[3]The unconditional variance of the difference captures the conditional variance of the level if y_t can be approximated as

$$y_t = \theta + \phi y_{t-1} + \sigma \varepsilon_t,$$

with $\phi = 1$. It can be used as an approximation for very persistent series such as interest rates when ϕ is less than but very close to one.

[4]The two measures differ, among other things, in both units and measures. The variance of multiperiod changes approximates the conditional variance of continuously compounded spot rates under the objective measure while the implied volatility approximates the conditional volatility *rate* of the simply compounded LIBOR rate under the risk neutral measure.

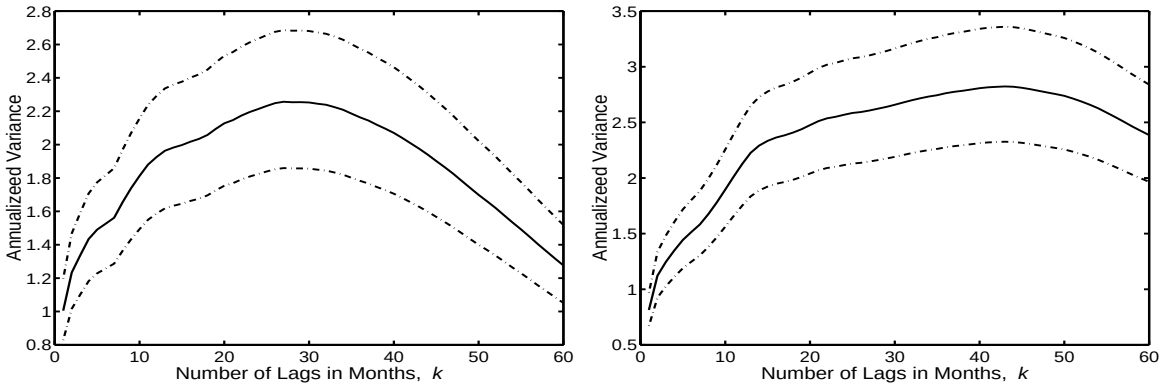

Figure 5. Hump-Shaped Dynamics In LIBOR Rates
Lines are annualized variance estimates of multiperiod changes in six-month spot rates (in annualized percentage), $Var\left(y_t - y_{t-k}\right)/(kh)$ for LIBOR rates on US dollar (left) and Deutsche mark (right). Data are monthly, from April 1987 to September 2000 (162 observations).

While the three dimensions do not exhaust the known properties of interest rates, they represent three layers of the data that a reasonable model should reproduce. The first layer summarizes the general statistical features of the time series and imposes minimal structure to the analysis. The second layer deals with predictability of interest rates and has far-reaching implications for interest rate forecasting. The third layer of property is even more subtle. It concerns with the conditional dynamics of the second moments, which has strong implications for model design and even more so for applications in risk management and derivatives pricing.

II. Quadratic Models

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leq t\leq \mathcal{T}}, \mathbb{P})$ be a stochastic basis. The filtration $\{\mathcal{F}_t\}_{0\leq t\leq \mathcal{T}}$ satisfies the usual conditions of right-continuity and completeness. We fix a strictly positive horizon date $\mathcal{T} > 0$. The process W is a d-dimensional Wiener process defined on $(\Omega, \mathcal{F}, \mathbb{P})$. We assume that the underlying filtration $\{\mathcal{F}_t\}_{0\leq t\leq \mathcal{T}}$ coincides with the usual $\mathbb{P}$-augmentation of the natural filtration of W. We introduce uncertainty in our economy by assuming that all assets are functions of an underlying Markov process X valued in some open subset $\mathcal{D}$ of $\mathbb{R}^d$.

A. Bond Pricing

Suppose that for any time $t \in [0, \mathcal{T}]$ and time-of-maturity $T \in [t, \mathcal{T}]$, the market value at time t of a zero-coupon bond with time-to-maturity $\tau = T - t$ is fully characterized by $P(X_t, \tau)$. The discrete-time notation P_t^n then corresponds to $P(X_t, nh)$ with $h = 1/12$ denoting the monthly time interval.

Definition 1 *In the* **quadratic class of term structure models***, the prices of zero-coupon bonds, $P(X_t, \tau)$, are exponential-quadratic functions of the Markov process X_t:*

$$P\left(X_t, \tau\right) = \exp\left[-X_t^\top A\left(\tau\right) X_t - \mathbf{b}\left(\tau\right)^\top X_t - c\left(\tau\right)\right], \tag{1}$$

where $A(\tau)$ is a nonsingular $d \times d$ matrix, $\mathbf{b}(\tau)$ is a $d \times 1$ vector, and $c(\tau)$ is a scalar.

Leippold and Wu (1999a) have identified the quadratic class in terms of the Markov process X_t, the instantaneous interest rate function $r(X_t)$, and the market price of risk $\gamma(X_t)$:

$$\begin{aligned} dX_t &= -\kappa X_t dt + dW_t; \\ r(X_t) &= X_t^\top A_r X_t + \mathbf{b}^\top X_t + c_r, \\ \gamma(X_t) &= A_\gamma X_t + \mathbf{b}_\gamma, \end{aligned} \tag{2}$$

where $\kappa, A_r, A_\gamma \in \mathbb{R}^{d\times d}$, $\mathbf{b}_r, \mathbf{b}_\gamma \in \mathbb{R}^d$, and $c_r \in \mathbb{R}$. Due to the symmetric nature of the quadratic models, we symmetrize A_r and $A(\tau)$ with no loss of generality. For the Markov process to be stationary, we also require that all the eigenvalues of κ be positive. We further normalize the diffusion of X_t to an identity matrix and its long run mean to zero.

As long as one is allowed to hold cash without cost, the instantaneous interest rate has to stay positive to guarantee no-arbitrage. A sufficient condition for the instantaneous interest rate to be positive is to restrict A_r to be positive definite and $c_r \geq \frac{1}{4}\mathbf{b}_r^\top A_r^{-1}\mathbf{b}_r$.

Straightforward application of Girsanov's theorem shows that, under the risk neutral measure $\mathbb{P}^*$, the drift of the Markov process remains affine with $\kappa^* = \kappa + A_\gamma$ and $\kappa^*\theta^* = -\mathbf{b}_\gamma$.

Stationarity for the Markov process under the risk neutral measure requires that all the eigenvalues of $\kappa^* = \kappa + A_\gamma$ be also positive.

Under the above specification, the coefficients for the bond price are determined by the following ordinary differential equations (ODE):

$$\begin{aligned} \frac{\partial A(\tau)}{\partial \tau} &= A_r - A(\tau)\kappa^* - (\kappa^*)^\top A(\tau) - 2A(\tau)^2; \\ \frac{\partial \mathbf{b}(\tau)}{\tau} &= \mathbf{b}_r - 2A(\tau)\mathbf{b}_\gamma - (\kappa^*)^\top \mathbf{b}(\tau) - 2A(\tau)\mathbf{b}(\tau); \\ \frac{\partial c(\tau)}{\tau} &= c_r - \mathbf{b}(\tau)^\top \mathbf{b}_\gamma + trA(\tau) - \mathbf{b}(\tau)^\top \mathbf{b}(\tau)/2, \end{aligned} \tag{3}$$

subject to the boundary conditions: $A(0) = 0$, $\mathbf{b}(0) = 0$, and $c(0) = 0$. In calibration, corresponding to the monthly data we use, we adopt a discrete-time version of the model. In particular, we choose monthly frequency, use the one-month rate as a proxy for the short rate, and solve the ODEs numerically by the Euler's method.[5] Given bond prices, bond yields are obtained straightforwardly:

$$y(X_t, \tau) = -\frac{1}{\tau} \ln P(X_t, \tau) = \frac{1}{\tau}\left(X_t^\top A(\tau) X_t + \mathbf{b}(\tau)^\top X_t + c(\tau)\right). \tag{4}$$

The discrete-time notation for the monthly yield, y_t^n, corresponds to $y(X_t, nh)$, with $h = 1/12$. The monthly forward rate, f_t^n, denotes the time t forward rate between $t+nh$ and $t+(n+1)h$:

$$f_t^n = \frac{1}{h} \ln P(X_t, nh)/P(X_t, (n+1)h) = X_t^\top A_n^f X_t + X_t^\top \mathbf{b}_n^f + c_n^f, \tag{5}$$

with

$$A_n^f = \frac{A((n+1)h) - A(nh)}{h}, \quad \mathbf{b}_n^f = \frac{\mathbf{b}((n+1)h) - \mathbf{b}(nh)}{h}, \quad c_n^f = \frac{c((n+1)h) - c(nh)}{h}.$$

The free parameters of the quadratic model include: $\Theta = (\kappa, A_\gamma, \mathbf{b}_\gamma, A_r, \mathbf{b}_r, c_r)$. Given these parameters and the current state vector X_t, bond prices, yields, and forward rates can be determined by (1), (4), and (5).

[5] Refer to Sewell (1988) for a standard reference.

B. Pricing State-Contingent Claims

Under the quadratic class, Leippold and Wu (1999a) prove that, for any asset with an exponential-quadratic terminal payoff structure at time T:

$$\exp\left(-q_1(X_T) - \int_t^T q_2(X_s)ds\right), \tag{6}$$

its time-t price is also exponential-quadratic in the current state X_t:

$$\begin{aligned}\psi(q_1 + \int_t^T q_2, \tau) &\equiv \mathbb{E}\left[\frac{\xi_T}{\xi_t}\exp\left(-q_1(X_T) - \int_t^T q_2(X_s)ds\right)\middle|\mathcal{F}_t\right] \\ &= \exp\left(-X_t^\top A(\tau)X_t - \mathbf{b}(\tau)^\top X_t - c(\tau)\right),\end{aligned} \tag{7}$$

where ξ_t denotes the pricing kernel at time t and $q_j(X)$ denotes a quadratic function of X, namely $q_j(X) = X^\top A_j X + \mathbf{b}_j^\top X + c_j$. The coefficients $A(\tau)$, $\mathbf{b}(\tau)$ and $c(\tau)$ also satisfy the ordinary differential equations in (3), but with different boundary conditions, $A(0) = A_1$, $\mathbf{b}(0) = \mathbf{b}_1$, and $c(0) = c_1$, to reflect the different terminal payoff structure. Furthermore, the short rate coefficients $\{A_r, \mathbf{b}_r, c_r\}$ need to be replaced by $\{A_r + A_2, \mathbf{b}_r + \mathbf{b}_2, c_r + c_2\}$ in the ordinary differential equations to reflect the integral on $q_2(X_s)$.

The quadratic form $q_j(X_t)$ can either be regarded as interest rates (bond yield or forward rate) or rates of return on other assets. The integral can be regarded either as an average rate in Asian style payoffs or as a cumulation of continuous payoffs.

This result can be applied to price state-contingent claims of the general type:

$$G_{q_i,q_j}(y,\tau) \equiv \mathbb{E}\left[\frac{\xi_T}{\xi_t}e^{-q_i(X_T)}\mathbf{I}_{q_j(X_T)\leq y}|\mathcal{F}_t\right], \tag{8}$$

where y can be regarded as some transform of a strike and $\mathbf{I}_x$ is an indicator function: it equals one when x is true and zero otherwise. As an example, when $y = \infty$, the claim reduces to the asset priced in (7): $G_{q_i,q_j}(\infty,\tau) = \psi(q_i,\tau)$. When we further assume $q_i = 0$, the claim is equivalent to a zero-coupon bond: $G_{0,q_j}(\infty,\tau) = P(X_t,\tau)$. On the other hand, for any fixed number y, if we set $q_i = 0$, $G_{0,q_j}(y,\tau)$ represents a state price: the price of an asset that pays

one dollar if and only if the state event $q_j(X_T) \leq y$ occurs. In what follows, we would refer to $G_{q_i,q_j}(y,\tau)$ as a *state price* in its broadest meaning. We also relax the notation on quadratic forms and let q_i and q_j denote any quadratic forms, or integral of quadratic forms, or any affine combinations of them.

In general, we cannot directly solve the state price $G(y)$ in closed form, but we can do so for its Fourier transform $\chi(z)$, defined as

$$\chi_{q_i,q_j}(z) \equiv \int_{-\infty}^{+\infty} e^{izy} dG_{q_i,q_j}(y), \quad z \in \mathbb{R},$$

where we omit the second argument in τ in the state prices and their transforms in case no confusion occurs. Under the quadratic class, such a Fourier transform can be computed as the price of an asset with an exponential-quadratic terminal payoff:[6]

$$\chi_{q_i,q_j}(z) = \psi(q_i - izq_j). \tag{9}$$

Of course, the term "asset price" has to be used with caution since the "asset" has a complex-valued payoff function. But more importantly, the equality in (9) implies that the Fourier transform of the state-contingent claim retains the exponential quadratic form and hence the tractability of the quadratic class.

Given the Fourier transform $\chi_{q_i,q_j}(z)$, the state price $G_{q_i,q_j}(y)$ can be obtained by an extended version of the Lévy inversion formula,

$$G_{q_i,q_j}(y) = \frac{\chi_{q_i,q_j}(0)}{2} + \frac{1}{\pi}\int_0^{\infty} \frac{\mathcal{R}(z)\sin yz - \mathcal{I}(z)\cos yz}{z} dz, \tag{10}$$

where $\mathcal{R}(z)$ and $\mathcal{I}(z)$ denote the real and imaginary part of $\chi_{q_i,q_j}(z)$, respectively. The above inversion formula involves only *one* numerical integration, regardless of the dimension of the state space. The prices of many interest rate derivatives such as European options on zero-coupon bonds, interest rate caps and floors, exchange options, and even Asian style options can

[6] Refer to Leippold and Wu (1999a) for a proof.

all be expressed in terms of such a general state price. We can therefore price them efficiently via the inversion formula in (10). We apply the method to interest rate caps pricing.

III. Property Analysis

In this section, we analyze the implications of quadratic models on the three layers of properties of interest rates. The analysis lays a foundation for moment choices in our generalized method of moments (GMM) estimation in the subsequent section. More importantly, we ask what requirements the stylized evidence imposes on model design.

A. Statistical Properties

With the Markov process specified in (2), the state vector X is both conditionally and unconditionally normally distributed. In our discrete-time version with monthly intervals, let $\Phi = e^{-\kappa h}$ with $h = 1/12$ denote the monthly autocorrelation matrix of the state vector, we can then write the unconditional and conditional variance as

$$\text{vec}\ (V) \quad = \quad (I - \Phi \otimes \Phi)^{-1}\ \text{vec}\ (I)h; \quad V_k = \sum_{j=0}^{k-1} \Phi^j \left(\Phi^j\right)^\top,$$

where the subscript k denotes the conditional variance in k discrete periods (months). The k-period conditional mean is $\mu_{t,k} = \Phi^k X_t$ while the unconditional mean is zero.

Let q_t^n denote a generic quadratic form of normal variates:

$$q_t^n = X_t^\top A_n X_t + \mathbf{b}_n^\top X_t + c_n,$$

for any $(A_n, \mathbf{b}_n, c_n)$. q_t^n can be a bond yield, a forward rate, or the short rate. Fully utilizing the well-documented properties of quadratic forms of normal variates,[7] we can derive the following properties for q_t^n:

$$\begin{aligned}
\mathbb{E}\left[q_t^n\right] &= tr\left(A_n V\right) + c_n; \\
Var\left[q_t^n\right] &= 2tr\left(\left(A_n V\right)^2\right) + \mathbf{b}_n^\top V \mathbf{b}_n; \\
Cov\left(q_{t+k}^n, q_t^n\right) &= 2tr\left(\left(\Phi^k\right)^\top A_n \Phi^k V A_n V\right) + \mathbf{b}_n^\top \Phi^k V \mathbf{b}_n.
\end{aligned}$$

The monthly kth-order autocorrelation, $\rho(k)$, of a quadratic form q_t^n is then given by

$$\rho(k) = \frac{2tr\left(\left(\Phi^k\right)^\top A_n \Phi^k V A_n V\right) + \mathbf{b}_n^\top \Phi V \mathbf{b}_n}{2tr\left(\left(A_n V\right)^2\right) + \mathbf{b}_n^\top V \mathbf{b}_n}.$$

In the case of a one-factor model, the autocorrelation function, $\rho(k)$, can be written as a weighted average of the autocorrelation and its square of the Markov process:

$$a(n)\phi^{2k} + b(n)\phi^k = \rho(k), \tag{11}$$

with $\phi \equiv \exp(-\kappa h)$ being the monthly autocorrelation of the Markov process X, and

$$a(n) = \frac{2(A_n V)^2}{2(A_n V)^2 + b_n^2 V}; \quad b(n) = \frac{b_n^2 V}{2(A_n V)^2 + b_n^2 V}.$$

Therefore, in contrast to the AR(1) type behavior of one-factor affine models, bond yields under quadratic models enjoy a richer, nonlinear dynamics. In particular, the autocorrelation function of bond yields under a one-factor quadratic model can vary across maturities, in conformity with the data, while all one-factor affine models imply the *same* autocorrelation function across yields and forward rates of *all* maturities. Within the affine class, multiple factors are needed to generate the observed nonlinearities in the interest rate dynamics. In

[7] See, for example, Holmquist (1996), Johnson and Kotz (1970), Kathri (1980), and Mathai and Provost (1992).

contrast, nonlinearity is intrinsically built into the quadratic model through the quadratic term.

B. Forecasting Relations

To derive the EH regression slopes, we apply two important properties of quadratic forms of normal variates:

$$\begin{aligned} Cov(q_t^m, q_t^n) &= 2tr(A_m V A_n V) + \mathbf{b}_m^\top V \mathbf{b}_n; \\ Cov(q_{t+k}^m, q_t^n) &= 2tr\left(\left(\Phi^k\right)^\top A_m \Phi^k V A_n V\right) + \mathbf{b}_m^\top \Phi^k V \mathbf{b}_n. \end{aligned}$$

We start with the single period forward rate regression. The slope of such a regression, c_n, can be written as

$$\begin{aligned} c_n &= \frac{Cov\left(f_{t+1}^{n-1} - r_t, f_t^n - r_t\right)}{Var\left(f_t^n - r_t\right)} \\ &= \frac{2tr\left[\left(\Phi^\top A_{n-1}^f \Phi - A_r\right) V A_{frn} V\right] + \mathbf{b}_{frn}^\top V \left(\Phi^\top \mathbf{b}_{n-1}^f - \mathbf{b}_r\right)}{2tr(A_{frn}V)^2 + \mathbf{b}_{frn}^\top V \mathbf{b}_{frn}}, \end{aligned} \tag{12}$$

where $A_{frn} = A_n^f - A_r$ and $\mathbf{b}_{frn} = \mathbf{b}_n^f - \mathbf{b}_r$. The relation is relatively opaque, but its convergence to the stationary state is not. Suppose indeed that a stationary state exists, as $n \to \infty$, $A_n^f \to 0$ and $\mathbf{b}_n^f \to 0$, the regression slope converges to one:

$$\lim_{n\to\infty} c_n = \frac{2tr(A_r V)^2 + \mathbf{b}_r^\top V \mathbf{b}_r}{2tr(A_r V)^2 + \mathbf{b}_r^\top V \mathbf{b}_r} = 1.$$

This results confirms with the analysis of Dybvig, Ingersoll, and Ross (1996). As long as the interest rate processes are stationary, the variance of the forward rate falls with maturity. Therefore, for very long maturities, we are essentially regressing $-r$ on itself.

Backus, Foresi, Mozumdar, and Wu (2001) illustrate that intrinsic tension exists for a one-factor Cox, Ingersoll, and Ross (1985)(CIR) model to simultaneously fit the mean yield curve and the regression slope. To make $c_1 < 1$, the market price risk needs to be greater than zero while an upward sloping mean yield curve requires it to be negative. To release the tension,

they propose a negative CIR model, where the short rate is proportional to the negative of a CIR factor.

Duffee (1999) further illustrates that the inherent tension remains even for multi-factor CIR models when one tries to match the properties of the whole term structure of excess returns. To release the tension, he proposes the application of Gaussian state variables with a flexible affine market price of risk specification. Dai and Singleton (2000a) incorporates such a specification to explain the EH violations. In particular, Dai and Singleton (2000a) show that such a specification also releases the tension identified by Backus, Foresi, Mozumdar, and Wu (2001) as $\mathbf{b}_\gamma$ controls the shape of the mean yield curve (and hence should be negative) while the slope parameter A_γ controls the regression slope (and should be positive).

Note, however, that the one-factor affine example of Dai and Singleton (2000a) is merely a degenerating case of a one-factor quadratic model with $A_r = 0$. Under the quadratic class, we only use Gaussian state variables. Affine market price of risk is naturally incorporated into the framework. The incorporation of the quadratic term A_r further enriches the interest rate dynamics and can prevent the interest rate from being negative.

The slope coefficients of the other three types of regressions can also be derived analogously under the quadratic framework:

$$
\begin{aligned}
d_n &= \frac{2tr\left[\left(\Phi^\top A_n^y \Phi - A_n^y\right) V A_{yrn} V\right] + \mathbf{b}_{yrn}^\top V \left(\Phi^\top \mathbf{b}_n^y - \mathbf{b}_r\right)}{n\left[2tr\left(A_{yrn}V\right)^2 + \mathbf{b}_{yrn}^\top V \mathbf{b}_{yrn}\right]}, \\
e_n &= \frac{2tr\left[\left((\Phi^n)^\top A_r \Phi^n - A_r\right) V A_{frn} V\right] + \mathbf{b}_{yrn}^\top V \left((\Phi^n)^\top \mathbf{b}_n^f - \mathbf{b}_r\right)}{2tr\left(A_{frn}V\right)^2 + \mathbf{b}_{frn}^\top V \mathbf{b}_{frn}}, \\
g_n &= \frac{\sum_{i=1}^n \left(1 - \frac{i}{n+1}\right)\left[2tr\left[\left((\Phi^i)^\top A_r \Phi^i - (\Phi^{i-1})^\top A_r \Phi^{i-1}\right) V A_{yrn} V\right] + \mathbf{b}_r^\top \left(\Phi^i - \Phi^{i-1}\right) V \mathbf{b}_{yrn}\right]}{2tr\left(A_{yrn}V\right)^2 + \mathbf{b}_{yrn}^\top V \mathbf{b}_{yrn}},
\end{aligned}
$$

where $A_n^y, \mathbf{b}_n^y$ are coefficients on the bond yield y^n, $A_{yrn} = A_{n+1}^y - A_r$, and $\mathbf{b}_{yrn} = \mathbf{b}_{n+1}^y - \mathbf{b}_r$.

C. The Hump Dynamics

Conditional dynamics in general and conditional variance in particular have far-reaching implications in risk management and option pricing. A central feature of the conditional dynamics for bond yields, as we observed earlier, is that the conditional volatility or variance of bond yields has a hump-shaped mean term structure. Let $cv(k)_n = \mathbb{E}\left[Var_t\left[y^n_{t+k}\right]\right]$ denote the mean conditional variance of n-month yields with a conditional horizon of k periods. Let $av(k)_n = cv(k)_n/(kh)$ denotes the annualized mean conditional variance. The hump-shaped conditional dynamics implies that $av(k)$ increases with k at first, reaches a plateau, and then decreases as k further increases.

The conditional moments of quadratic forms of normal variates are given by

$$\begin{aligned}
\mathbb{E}_t\left[q^n_{t+k}\right] &= \mu^\top_{t,k} A_n \mu_{t,k} + tr\left(A_n V_k\right) + \mathbf{b}^\top_n \mu_{t,k} + c_n; \\
Var_t\left[q^n_{t+k}\right] &= 2\left[tr\left(\left(A_n V_k\right)^2\right) + 2\widetilde{\mu}^\top_{t,k} A_n V_k A_n \widetilde{\mu}_{t,k}\right],
\end{aligned}$$

where $\widetilde{\mu}_{t,k} = \mu_{t,k} + \frac{1}{2}A_n^{-1}\mathbf{b}_n$ and $\mu_{t,k} = \Phi^k X_t$. The mean term structure of the conditional variance under the quadratic class can be written as

$$cv\left(k\right)_n = 2tr\left[\left(A_n V_k\right)^2 + 2\left(\Phi^k\right)^\top A_n V_k A_n \Phi_k V\right] + \mathbf{b}^\top_n V_k \mathbf{b}_n. \tag{13}$$

Proposition 1 *Strong interdependence between elements of the state vector is essential in generating a hump-shaped conditional dynamics. Neither one-factor nor independent multifactor quadratic models are capable of generating the hump.*

Refer to appendix A for the proof. Similar necessary conditions are also required for affine models. However, the correlation structures between multi-factor CIR models are restricted. For example, Dai and Singleton (2000b) and Backus, Foresi, and Telmer (2001) both observe that the unconditional correlation between two square-root state variables can only be positive. Hence, while multi-factor CIR models in theory can generate a hump shape, the hump is often not large enough to match the evidence. In contrast, quadratic models only incorporate Ornstein-Uhlenbeck processes as the driving Markov process, the correlation structure between

state variables can be chosen freely. Fitting the hump shape hence becomes a relatively easy task.

From the time series data, the conditional dynamics can be approximately captured by the annualized variance of multiperiod changes. Under the quadratic model, the variance of k-period changes in yields or forward rates q_t^n is given by

$$\begin{aligned} v(k)_n &= Var(q_{t+k}^n - q_t^n) = 2\left[Var(q_t^n) - Cov(q_{t+kh}^n, q_t^n)\right] \\ &= 4tr\left(\left(A_n - \left(\Phi^k\right)^\top A_n \Phi^k\right) V A_n V\right) + 2\mathbf{b}_n^\top \left(I - \Phi^k\right) V \mathbf{b}_n. \end{aligned}$$

One can readily prove that similarly strong interactions between state variables are required for the annualized variance of multi-period changes to be hump-shaped.

Nevertheless, conditional dynamics implied from options prices and conditional dynamics inferred from the time series are dynamics under two different measures. The former is under the risk-neutral measure while the latter is under the objective measure. The correlation structure is hence captured by $\kappa^* = \kappa + A_\gamma$ in the former and by κ in the latter. Therefore, to simultaneously capture the observed conditional dynamics in both the time series and option prices, one also imposes constraints on the specification of the market price of risk.

In summary, the quadratic class of term structure models exhibits great potentials in simultaneously (1) matching the mean yield curve and the forecasting relations through the specification of the affine market price of risk $A_\gamma X_t + \mathbf{b}_\gamma$, and (2) generating the observed hump-shaped conditional dynamics by the flexible specification of the correlation structures between the state variables. Furthermore, the quadratic term enriches the dynamics of the interest rate by incorporating nonlinearity between state variables and interest rates and also provides a consistent approach to guarantee positive interest rates. In the next section, we further illustrate these properties by calibrating the quadratic model to the times series data on eurodollar interest rates.

IV. GMM Estimation

This section corroborates the property analysis in the previous section with empirical estimation. For this purpose, we choose the simplest model within the quadratic class which can approximate the three layers of properties of interest rates. As demonstrated in Dai and Singleton (2000a), a one-factor quadratic model suffices in capturing both the mean yield curve and the EH regression slopes. To capture the hump dynamics, however, we need at least a two-factor model to incorporate non-trivial correlation structures between state variables. Moment conditions are chosen based on the property analysis in the previous section. We calibrate the model to the time series of interest rates on euro dollars.

A. Moment Conditions and Inference

The parameter set of the quadratic model is: $\Theta \equiv (\kappa, A_r, \mathbf{b}_r, c_r, A_\gamma, \mathbf{b}_\gamma)$. We choose 40 moment conditions for the GMM estimation. These are taken from

1. **Mean yield curve**: Mean yields with maturities of 1, 3, 6, 12, 24, 60, and 120 months.
2. **Forecasting relations**: One period forward regression slopes, c_n, with maturities $n = 1, 3, 6, 12, 24$.
3. **Variance of multiperiod changes**: Variance of k-period (month) changes of 1-month and 6-month yields with k from 6 to 58 with an equal interval of 4.

Furthermore, given other parameters, we set c_r to match the mean short rate (μ_r), $c_r = \mu_r - tr(A_r V)$. We hence have 39 moment conditions left for the GMM estimation. The moment conditions and their values as functions of parameters are summarized in Table II.

Let $s_t^n = y_t^n - r_t$ denote the yield spread and $\Delta^{(k)} y_t^n$ the k-period change of the yield y_t^n. Let $\mu(x)$, c_n, and $v(k)_n$ denote, respectively, the mean of x, the slope of the forward regression with maturity n, and the variance of k-period changes in y_t^n, all implied from the model parameters

and computed based on the formulae in Table II. Given these notations, the three types of moment conditions can be written as:

$$
\begin{aligned}
h_t^1(n) &= s_t^n - \mu(s^n), && (n = 3, 6, 12, 24, 60, 120) \\
h_t^2(n) &= \left[\left(f_t^{n-1} - r_{t-1}\right) - c_n\left(f_{t-1}^n - r_{t-1}\right)\right]\left(f_{t-1}^n - r_{t-1}\right), && (n = 1, 3, 6, 12, 24) \\
h_t^3(n,k) &= v(k)_n - \left(\Delta^{(k)} y_t^n\right)^2, && (n = 1, 6;\ k = 6, 10, \cdots, 58)
\end{aligned}
$$

where k denotes the order and n denotes the maturity. The forward rate and the short rate in h_t^2 are demeaned: $f_t^n = f_t^n - \bar{f}_t^n$. The moment conditions are further scaled to have roughly unit variance. The first type of moment conditions captures the mean level of interest rates; the second type captures the cross-correlation between interest rates of different maturities; the third type captures the auto-correlation of interest rates. To reduce the persistence of the moment conditions, we construct the moment conditions on spreads (h^1 and h^2) or differences (h^3), instead of on levels.

We estimate one unrestricted and three restricted versions of the two-factor model:

- *Model A* (unrestricted two-factor quadratic model);
- *Model B* (independent two-factor quadratic model): $\kappa(1,2) = A_\gamma(1,2) = A_r(1,2) = 0$;
- *Model C* (two-factor quadratic model with constant market price of risk): $A_\gamma = 0$;
- *Model D* (two-factor Gaussian affine model): $A_r = 0$.

For identification reasons, we normalize κ and A_γ to be upper triangular and A_r symmetric. Model B disallows any interactions between the two state variables and hence serves as an over-identification test on how important such interactions are in capturing the properties of eurodollar interest rates. Model C, on the other hand, provides a test on the significance of the affine market price of risk A_γ. Finally, Model D tests the significance of the contribution of the quadratic term A_r.

Excluding c_r as a free parameter and the mean short rate as a moment condition, we have 39 orthogonality conditions and 13 free parameters in the unrestricted model (Model A). Hence, the model is over-identified and has 26 over-identification restrictions. Each of the

three restricted models (B, C, and D) has three more constraints on parameters. The number of over-identification restrictions increases to 29.

For inference, let T denote the number of observations and let $\mathbf{e}_T = [e^j(n,j)] \in \mathbb{R}^{22\times 1}$, with

$$e^j(n,k) \quad = \quad \frac{1}{T}\sum_{t=1}^{T} h_t^j(n,k), \quad j = 1,2,3,$$

denote the sample average of the orthogonality conditions. Let $\hat{\Theta}_T$ be the parameter estimates that minimize the objective function:

$$\mathbf{e}_T^\top W \mathbf{e}_T.$$

We normalize the orthogonality conditions to have unit variance and then set $W = I$ for the estimation. Asymptotically, under certain technical conditions (Hamilton, 1994), the estimator has a normal distribution:

$$\hat{\Theta}_T \sim N\left(\Theta, V/T\right), \text{ where } V = \left[D'WD\right]^{-1} D'WSWD\left[D'WD\right]^{-1}, \tag{14}$$

where S is the spectral density matrix, estimated following Newey and West (1987) with 12 lags, and D is the Jacobian matrix, defined as

$$D = \left.\frac{\partial \mathbf{e}_T^\top}{\partial \Theta}\right|_{\Theta=\hat{\Theta}_T}.$$

Under certain technical conditions, the moment conditions also have an asymptotic normal distribution,

$$\mathbf{e}_T \sim N\left(0, M/T\right), \text{ where } M = \left[I - D(D'WD)^{-1}D'W\right] S \left[I - D(D'WD)^{-1}D'W\right]^\top.$$

When the model is over-identified, as is the case in this paper, a χ^2 test can be constructed for the over-identifying restrictions,

$$T\mathbf{e}_T^\top (M)^{-1}\mathbf{e}_T \sim \chi^2(r-a),$$

where $r = 39$ is number of orthogonality conditions and a is number of free parameters (13 for Model A and 10 for Models B, C, and D).

B. Performance Analysis

Table III reports the GMM estimates and standard errors of the four two-factor quadratic models. Also included are the J-statistics and the probability of negative interest rates. Based on the J-statistics, all models are overwhelmingly rejected, indicating the need for more factors to fully capture the interest rate dynamics. Nevertheless, as illustrated by Figure 6, an unrestricted two-factor quadratic model (Model A) approximates very well the three layers of stylized evidence: the mean yield curve (top panel), the forward regression slopes (middle panel), and hump-shaped dynamics (bottom panel). The solid lines in the figure are implied by the model estimates in Table III. The dash-dotted lines correspond to the 10% and 90% quantiles. They are computed by the delta method, based on the covariance matrix of the parameter estimates and an asymptotic normal distribution assumption.

In Model B, we restrict the off-diagonal terms of κ, A_γ, and A_r all to zero. As a result, the two state variables do not interact with each other and do not interact in their impacts on the behavior of interest rates. In absence of such interactions, Model B fails to generate the hump-shaped dynamics observed in the interest rate data. Of the four models tested, the J-statistic is largest for Model B. Figure 7 depicts the fit of the hump-dynamics by the four models with parameter estimates from Table III. Of all the models tested, Model B stands out by its apparent failure in capturing the hump shape in the interest rate dynamics. For the other three models (A,C,D), the estimates of κ_{12} are all large and statistically significant, demonstrating the important role played by strong interactions between state variables in generating the hump dynamics.

Dai and Singleton (2000b) also find that the correlation structure between state variables plays an important role in improving the general empirical performance of affine models. Ahn, Dittmar, and Gallant (2001) have similar findings for the quadratic class. The analysis here

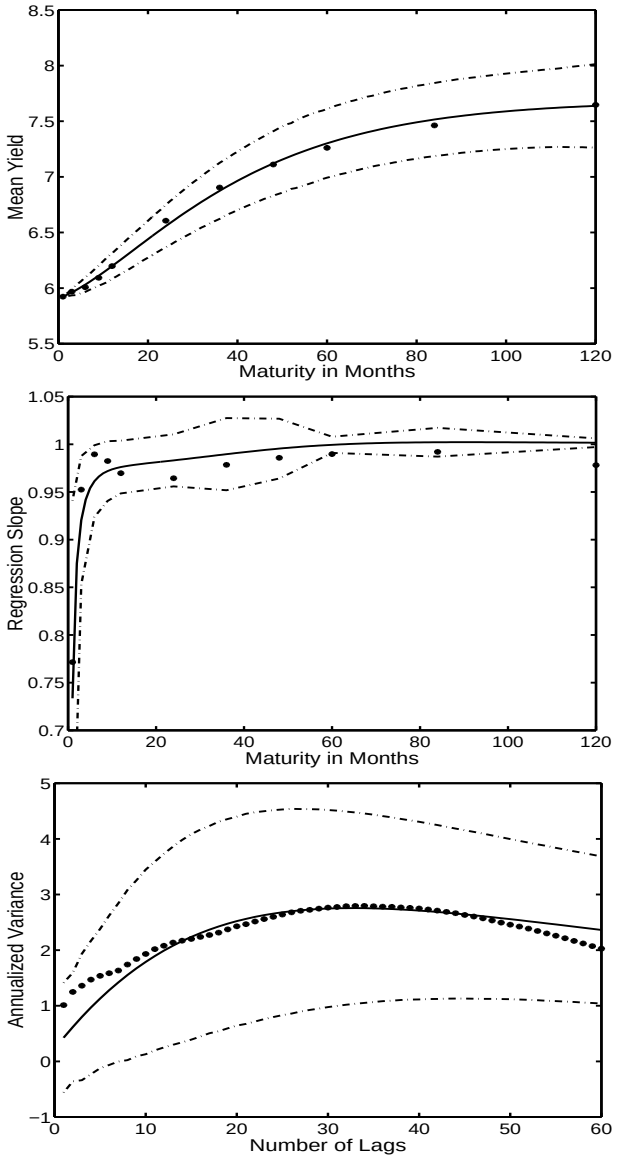

Figure 6. Performance of a Two-Factor Quadratic Model

Lines represent the performance of the unrestricted two-factor quadratic model (Model A) in matching the mean yield curve (top), the forward regression slope (middle), and the annualized variance of multiperiod changes of the six-month yield (bottom). Stars are data estimates. The solid lines correspond to the point estimates of Model A reported in Table III. The dashed lines correspond to the 10% and 90% quantiles, computed by the delta method based on the covariance matrix of the parameter estimates and an asymptotic normality assumption.

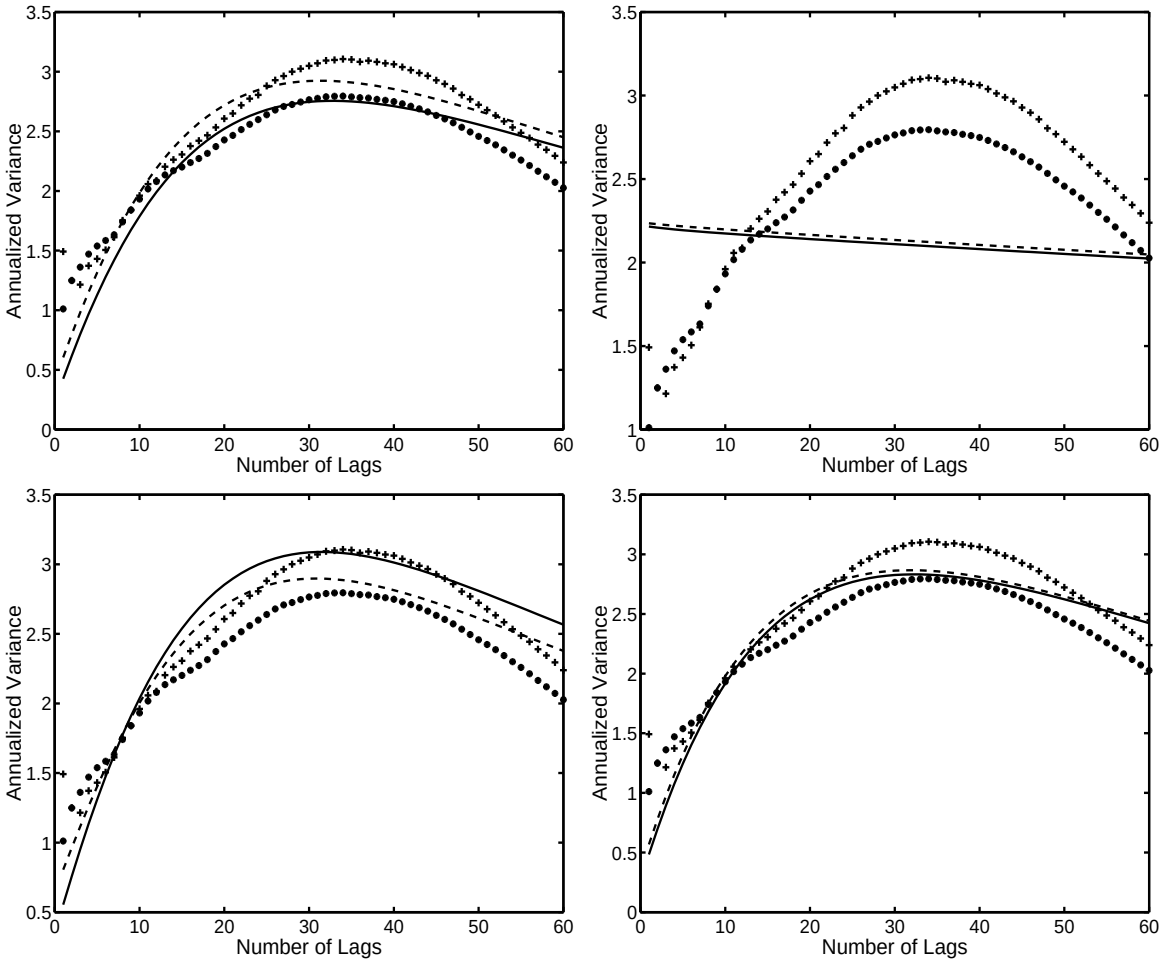

Figure 7. Fitting The Hump Dynamics for US LIBORs

The dashed and solid lines represent the performance of the four models: A (top left), B (top right), C (bottom left), and D (bottom right), in matching the annualized variance of multiperiod changes in one-month and six-month spot rates, respectively. Circles and stars are data estimates for the one-month and six-month spot rates. Parameter estimates of the models are reported in Table III.

pinpoints a dimension of the data which specifically requires such interactions and therefore provides an explanation as to why the general performance is improved.

In Model C, we set $A_\gamma = 0$. The market price of risk is therefore reduced to a constant vector. As illustrated in Figure 8 and the j-statistics in Table III, such a restriction does not deteriorate the model performance dramatically. The largest impact of such a restriction is on the confidence band of the regression slopes. As A_γ plays an important role in the EH regressions, setting $A_\gamma = 0$ significantly reduces the confidence band for the regression slopes. However, since the deviations from the EH are not statistically significant for eurodollars interest rates, a time-varying risk premium, as captured by a non-zero A_γ, is not called for. Indeed, the estimates for A_γ in the unrestricted model (Model A) are not significantly different from zero. Our experience, however, indicates that simultaneously fitting the mean yield curve and the regression slopes on the US Treasuries is extremely difficult with $A_\gamma = 0$. As we discussed in the data section, violations of EH is much more severe for US Treasuries than for eurocurrencies: the regression slope is much lower than one and the mean yield curve is also steeper. Simultaneously fitting them then often requires more time-variability in the risk premium and hence asks for a non-zero A_γ.

In Model D, we set $A_r = 0$. The model hence degenerates to a two-factor Gaussian affine model. One key weakness of such a model is the unrealistic implication that interest rates are normally distributed:

$$r \sim N\left(c_r, \mathbf{b}_r^\top V \mathbf{b}_r\right),$$

where V is the unconditional covariance matrix of the state vector X_t. Furthermore, the quadratic term also plays a role in keeping interest rate positive. For example, if we restrict the quadratic term A_r to be positive definite, we can rewrite the short rate function as

$$r(X_t) = \left(X_t + \frac{1}{2}A_r^{-1}\mathbf{b}_r\right)^\top A_r \left(X_t + \frac{1}{2}A_r^{-1}\mathbf{b}_r\right) + c_r - \frac{1}{4}\mathbf{b}_r^\top A_r^{-1}\mathbf{b}_r.$$

Obviously, as long as A_r is positive definite and $c_r \geq \frac{1}{4}\mathbf{b}_r^\top A_r^{-1}\mathbf{b}_r$, the short rate is guaranteed to be positive. Parameter estimates for Models A, B, and C in Table III all satisfy this positivity constraint for the short rate. In contrast, for Model D, as A_r is restricted to zero and interest

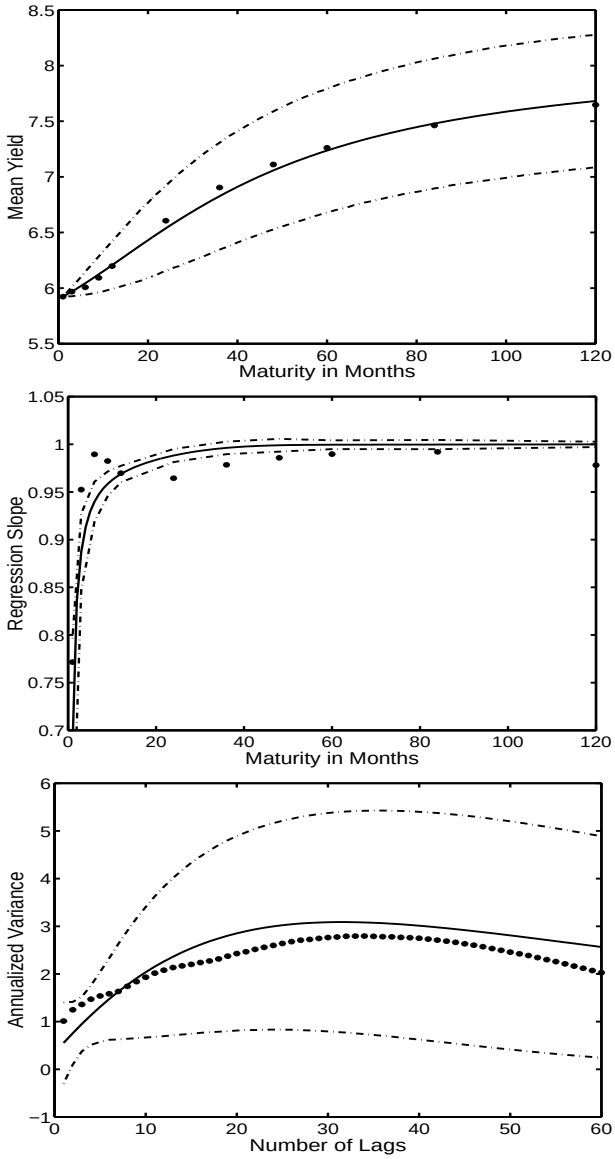

Figure 8. Performance of Model C with $A_\gamma = 0$

Lines represent the performance of Model C in matching the mean yield curve (top), the forward regression slope (middle), and the annualized variance of multiperiod changes of the six-month yield (bottom). Stars are data estimates. The solid lines correspond to the point estimates of Model A reported in Table III. The dashed lines correspond to the 10% and 90% quantiles, computed by the delta method based on the covariance matrix of the parameter estimates and an asymptotic normality assumption.

rates are normally distributed, positivity cannot possibly be satisfied. Under the estimates in Table III, Model D implies a probability of 23.21% at which the short rate can be negative.

C. Small Sample Bias

Due to the high persistence of interest rates and the small sample size on our interest rate data, concerns on small sample bias arise naturally. In particular, Bekaert, Hodrick, and Marshall (1997) and Bekaert and Hodrick (2000) find that the small sample bias can be significant for the EH regression slope estimates. Backus, Foresi, Mozumdar, and Wu (2001) confirm their results based on estimates on an affine model. In this section, we use Model A as a benchmark to investigate the small sample bias problem. In particular, we use the parameter estimates for Model A and generate 5,000 samples of 162 observations each for the state vector X. The state vector is simulated based on the following discrete approximation of the Markov process on a monthly interval:

$$X_{t+1} = \Phi X_t + \sqrt{h}\varepsilon_{t+1},$$

with $\varepsilon \sim NID(0, I)$. From the simulated series on the state vector, we calculate the spot rates and forward rates, and estimate the mean yield curve, hump dynamics, and the regression slopes from each sample. The small sample bias is then captured by the difference between the population values implied by Model A and the mean values from the simulation. In Figure 9, we plot together the data estimates (stars), the population values implied by Model A (solid lines), and the mean estimates from the 5,000 simulated series (dashed line). To demonstrate the sampling variability, we also plot the 10% and 90% percentiles (dash-dotted lines) for the estimates on the simulated series.

Consistent with the analysis in both Bekaert, Hodrick, and Marshall (1997) and Backus, Foresi, Mozumdar, and Wu (2001), the small sample bias for the forward rate regression is positive at the short end, implying that the violation of the EH might be more severe than it looks. However, the sampling variability is huge at short maturities, thus weakening the evidence on EH violations.

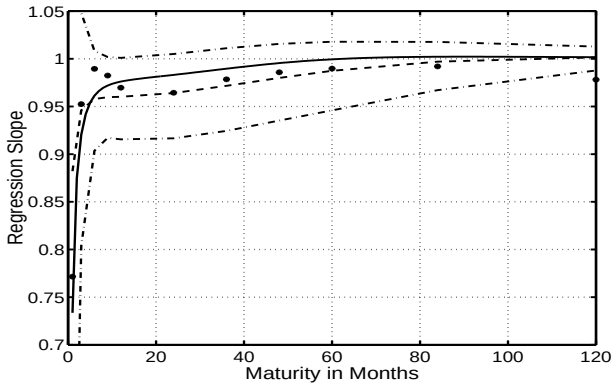

Figure 9. Sampling Variability and Small Sample Bias
Stars are estimates from the data. The solid lines are implied by Model A with estimates reported in Table III. The dashed lines are the mean values from a Monte Carlo simulation while the dash-dotted lines depict the 10% and 90% percentile.

V. Pricing Interest Rate Caps

To this point, we have focused exclusively on the time series properties of interest rates in guiding the deign and estimation of the quadratic model. Given our emphasis on the hump conditional dynamics, it is natural to exploit information from option markets as well. In this section, we investigate the option pricing performance of our estimated model by examining whether it can generate the average hump shape of the term structure of implied volatilities on interest rate caps.

The average shape of the term structure of implied volatilities will primarily be affected by two forces. The first is the conditional dynamics of the state vector under the risk neutral measure, captured mainly by $\kappa^* = \kappa + A_\gamma$. We expect the term structure of the implied volatility to be an approximate representation of the conditional volatility rate of the LIBOR rate. The difference between the conditional volatility and the implied volatility represents the second source: the pricing bias of the Black model due to conditional non-normality in the log of the simple rate. The simply compounded interest rate is log-normally distributed under the Black model but the deviation is obvious under the quadratic class. While we have analyzed the importance of a flexible correlation structure between the state variables in generating the

hump-shaped conditional dynamics, the role played by conditional non-normality is relatively opaque and calls for a numerical exercise.

Consider a cap contract written on the six-month euro dollar LIBOR rate, R^τ, with $\tau = 1/2$. The simple rate relates to the continuously compounded spot rate y^τ by

$$R^\tau = \frac{1}{\tau}\left(e^{\tau y^\tau} - 1\right).$$

Let t denote today, $t_1 = t + i\tau$ the settlement date, and $t_2 = t + (i+1)\tau$ the payment date of the ith caplet. The price of such a caplet, c_t, is

$$c_t^i \;=\; L\mathbb{E}\left[\frac{\xi_{t_2}}{\xi_t}\left(\exp(\tau y_{t_1}^\tau) - (1+\tau K)\right)^+ |\mathcal{F}_t\right] = L\mathbb{E}\left[\frac{\xi_{t_1}}{\xi_t}\left(1 - (1+\tau K)P(X_{t_1},\tau)\right)^+ \mathcal{F}_t\right].$$

The last equality is obtained by principle of conditional expectation. Straightforward application of the state price definition in (8) yields

$$c_t^i = L\left[G_{0,-q}(-\ln(1+\tau K), t_1) - (1+\tau K)G_{q,-q}(-\ln(1+\tau K), t_1)\right],$$

with $q = -\ln P(X_{t_1},\tau)$. We can therefore compute the caplet price from the transform method proposed in Section II. Furthermore, by virtue of linearity, we also have the Fourier transform on any affine combinations of the state price:

$$\chi(z) \equiv \int_{-\infty}^{\infty} e^{izy} d\left(b^\top G_{q_i,q_j}\right) = b^\top \psi(q_i - izq_j, \tau),$$

where q_i and q_j denote vectors of quadratic forms. The caplet price is an affine combination of two state prices. We can hence compute the Fourier transform of the caplet first and then obtain the caplet price by one numerical integration. The numerical integration is performed using a higher order adaptive quadrature method.

The caplet price c_t depends on both the model parameters and the current state vector X_t. As we are interested in the average shape of the implied volatility, we set the state vector to its long run mean: $X_t = \mathbf{0}$, and price caps based on the term structure implied by Model A with parameter estimates reported in Table III.

A. Conditional Non-Normality and Volatility Smile

Under the Black model, the log of the LIBOR rate R_t^τ is assumed to be normally distributed. Therefore, conditional non-normality results in a pricing bias on the part of the Black model. Such a pricing bias is most vividly shown by the so-called "implied volatility smile" across moneyness (strike): when the implied volatility is plotted against moneyness, one often observes a smile (and sometimes a smirk) pattern while such a plot should generate a flat line under the conditional-normality assumption of the Black model. Thus, any non-flat pattern for such a plot becomes an indication of conditional non-normality for the underlying distribution under the risk neutral measure.

As an illustration, we compute caplet prices at different strikes based on the parameter estimates for Model A and $X = \mathbf{0}$. We then infer the Black implied volatility for each caplet and plot the implied volatility against a moneyness measure, which is defined as

$$Moneyness = \frac{\ln K/R_t}{\sigma\sqrt{i\tau}},$$

where $i\tau$ is the maturity of the i-th caplet, and $\sigma = 10\%$ is some fixed volatility level. While moneyness has been defined in many different ways in the literature, we use the above definition, and the $\sqrt{i\tau}$ scaling in particular, to facilitate comparison across maturities. The choice of $\sigma = 10\%$ is arbitrary, unless one intends to compare implied volatility patterns across different underlying interest rates.[8]

Figure 10 represents such a plot for Model A. We see that the implied volatility increases as the strike (moneyness) increases (further out of the money). Such a pattern implies that the Black model underprices out-of-the-money caplet contracts relative to at-the-money contracts. It indicates that Model A implies a thicker right tail (positive skewness) compared to the normal assumption of the Black model for the log LIBOR rate.

[8] One may then want to use σ to capture the average volatility level differences across different underlyings. Such practices have become industry standard.

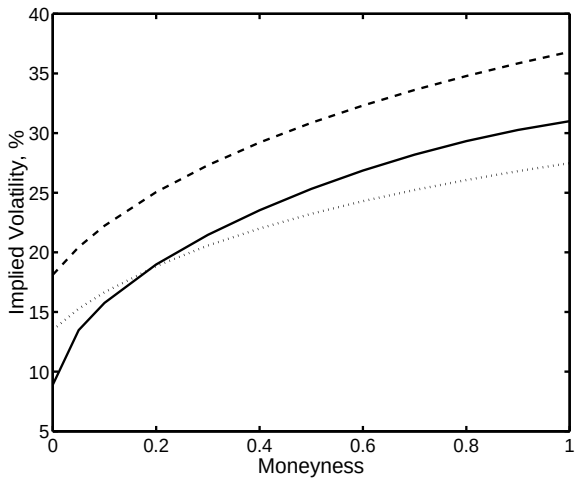

Figure 10. Caplet Volatility Smile
Implied volatility across moneyness are inferred from caplet prices, computed based on Model A, with parameter estimates reported in Table III and assuming $X = \mathbf{0}$. The maturities are one year (solid line), four years (dashed line), and ten years (dotted line). Moneyness is defined as $(\ln K/R)/\sigma\sqrt{i\tau}$, with $\sigma = 10\%$ and $i\tau$ being the maturity of the contract.

B. The Hump-Shaped Term Structure

To compare the model with the cap data, we also compute the Model A-implied prices of at-the-money interest rate caps at maturities from one year to ten years. We then infer the Black implied volatility from these cap prices. As we set $X = \mathbf{0}$ at its long run mean, we expect to obtain an average term structure for the implied volatility.[9] Figure 11 captures such a model-implied mean term structure. The figure illustrates a distinct hump shape that is reminiscent of those observed in the interest rate cap data in Figure 4.

A major source of this hump shape is from the conditional dynamics of the six-month LIBOR rate. In Figure 12, we plot the term structure of the annualized conditional volatility of the six-month spot rate under both the risk neutral measure (solid line) and the objective

[9]We recognize the potential differences incurred by averaging on different metric spaces: the mean term structure of implied volatility on the cap data in Figure 4 is obtained by averaging on implied volatility quotes while the mean term structure in this section computed from the quadratic model is obtained by averaging on the state vector X. Nevertheless, we do not expect the different averaging schemes to alter the qualitative feature of the hump shape.

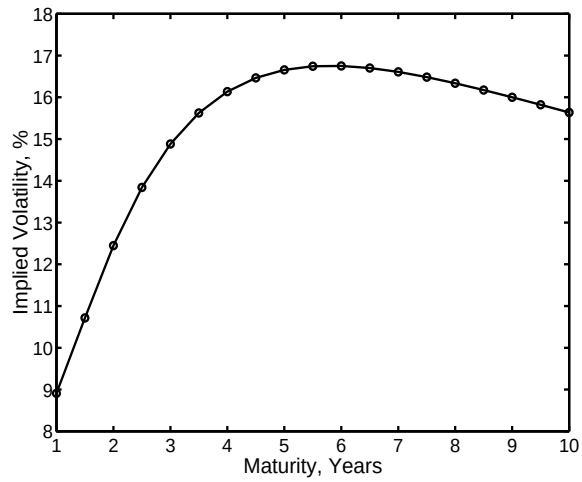

Figure 11. Hump Shape in Cap Implied Volatility
Implied volatilities at different maturities are inferred from at-the-money cap prices, computed based on Model A, with parameter estimates reported in Table III and with $X = \mathbf{0}$.

measure (dashed line). With $X = \mathbf{0}$, the conditional volatilities under the risk neutral measure are given by

$$\sigma^*(k) = \sqrt{\left[2tr\left(A_6 V_k^*\right)^2 + \mathbf{b}_6^\top V_k^* \mathbf{b}_6\right]/(kh)},$$

where $(A_6, \mathbf{b}_6)$ are the coefficients of the quadratic function for the six-month spot rate (continuously compounded) and $V_k^* = \sum_{j=0}^{k-1}(\Phi^*)^j((\Phi^*)^\top)^j$ is the k-period ahead conditional variance matrix of the state vector under the risk neutral measure. The conditional volatility under the objective measure is obtained analogously with V_k^* replaced by $V_k = \sum_{j=0}^{k-1}\Phi^j(\Phi^\top)^j$

Compare the solid line in Figure 12 with that in Figure 11, we see the close correspondence between hump shape in the conditional volatility under the risk neutral measure and the hump shape in the implied volatility. We recognize that, even in absence of the pricing bias, the implied volatility and σ^* measure different things: the implied volatility measures the conditional volatility of the log of the simply compounded interest rate while σ^* measures the conditional volatility of the continuously compounded interest rate. Nevertheless, the two conditional volatilities capture the same dynamics of the interest rate and exhibit very similar behaviors.

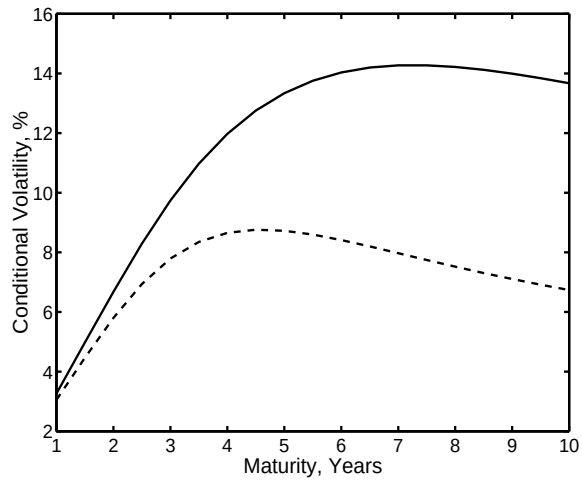

Figure 12. Annualized Conditional Volatility of 6-Month Spot Rate
Annualized conditional volatility under the risk neutral measure (solid line) and the objective measure (dashed line) are computed based on Model A, with estimates reported in Table III, assuming $X = \mathbf{0}$.

Note also that, in our estimation, we capture the hump shape dynamics in the interest rate by fitting the variance of multiperiod changes in interest rates. However, the variance of multiperiod changes only captures the conditional dynamics of the interest rate under the objective measure (the dashed line) but not that under the risk-neutral measure (the solid line). The moment conditions therefore only impose direct constraints on κ but not κ^*. The difference is captured by the affine market price of risk term $A_\gamma = \kappa^* - \kappa$. In our calibration, we intend to capture the market price of risk A_γ by matching the mean yield curve and the slopes of the single-period forward rate regressions. For future research, it would be interesting to investigate the implications of the options data on the market price of risk.

VI. Concluding Remarks

We document the stylized evidence on interest rates and interest rate derivatives, study its implications on model design, and calibrate a two-factor quadratic term structure model that approximates the stylized evidence relatively well. Along the way, however, we ignore a number of issues that deserve comment. First, a time-homogeneous two-factor model obviously cannot

match perfectly the whole term structure of interest rates at each day, let alone the term structure of interest rate caps. Factor analysis by, for example, Litterman and Scheinkman (1991) claims the existence of at least three factors. Furthermore, as a standard practice in the industry for derivatives pricing, one almost always takes the observed interest rate term structure as given, with minimal or no smoothing. To reconcile the observed term structure with a model such as ours, one often incorporates some time-inhomogeneous (and maybe also maturity-inhomogeneous) drift terms (fudge factors). Our choice of two factors represents the minimum number of factors required to capture the hump dynamics. Another issue is about the moment conditions for the GMM estimation, the selection of which is almost never unique. Ours is no exception. Admittedly, we could have used more "efficient" methods such as the maximum likelihood estimation, or the efficient methods of moments applied by, for example, Ahn, Dittmar, and Gallant (2001) and Dai and Singleton (2000b). Our choice of GMM and the particular moment conditions, on the other hand, highlights the mapping between the stylized evidence and the model structure, which is the focus of our paper.

Appendix A. Proofs

Proof of Proposition 1:

Recall from (13) that the k-period conditional variance of a quadratic form $q(X_t) = X_t^\top A X_t + \mathbf{b}^\top X_t + c$ is given by

$$cv(k) \equiv \mathbb{E}\left[q(X_{t+k})|\mathcal{F}_t\right] = 2tr\left[(AV_k)^2 + 2\left(\Phi^k\right)^\top AV_k A\Phi^k V\right] + \mathbf{b}^\top V_k \mathbf{b}. \tag{A1}$$

For the term structure to exhibit a hump shape, we need the annualized variance $cv(k)/(kh)$ to increase initially with k, which implies that

$$cv(2) - 2cv(1) > 0.$$

In a one-factor case, from (A1), we have

$$\begin{aligned} cv(2) - 2cv(1) &= \frac{2A^2(1+\phi^2)(1+\phi^4)}{1-\phi^2} + b^2(1+\phi^2) - \frac{4A^2(1+\phi^2)}{1-\phi^2} - 2b^2 \\ &= -2A^2(1+\phi^2)^2 - b^2(1-\phi^2) < 0. \end{aligned} \tag{A2}$$

as $\phi < 1$ for a stationary Markov process. Hence, the initial term structure is downward sloping. No hump can be generated from a one-factor model. The same is also true for completely independent multifactor models with A_r, κ, A_γ being diagonal matrices, in which case, the last line in (A2) represents the contribution from each diagonal element, which is negative as long as that element is stationary. $cv(2) - 2cv(1)$ can become positive only when the off-diagonal terms becomes dominating. Hence, strong interdependence between elements of the state vector is essential to generate hump dynamics.

The same argument also holds, almost in parallel, for the variance of multiperiod changes of $q(X_t)$,

$$v(k) \equiv Var\left(q(X_{t+k}) - q(X_t)\right) = 4tr\left(\left(A - \left(\Phi^k\right)^\top A\Phi^k\right) VAV\right) + 2\mathbf{b}^\top\left(I - \Phi^k\right) V\mathbf{b}. \tag{A3}$$

For the term structure to exhibit a hump shape, we need the annualized variance to increase initially with periods k, or $v(2) - 2v(1) > 0$. Yet, for a one-factor model, we have again

$$v(2) - 2v(1) = -4\left(1-\phi^2\right)^2 V^2 A_n^2 - 2\left(1-\phi\right)^2 b_n^2 V < 0.$$

Hence, the initial term structure is also downward sloping. No hump can be generated from a one-factor model. And the same conclusion holds for completely independent multifactor models.

References

Ahn, Dong-Hyun, 1998, Generalized squared-autoregressive-independent-variable term structure model, manuscript, University of North Carolina.

Ahn, Dong-Hyun, Robert F. Dittmar, and A. Ronald Gallant, 2001, Quadratic term structure models: Theory and evidence, *Review of Financial Studies* forthcoming.

Amin, Kaushik I., and Andrew Morton, 1994, Implied volatility functions in arbitrage-free term structure models, *Journal of Financial Economics* 35, 141–180.

Backus, David, Silverio Foresi, Abon Mozumdar, and Liuren Wu, 2001, Predictable changes in yields and forward rates, *Journal of Financial Economics* 59, 281–311.

Backus, David, Silverio Foresi, and Chris Telmer, 2001, Affine term structure models and the forward premium anomaly, *Journal of Finance* 56, 279–304.

Backus, David, and Liuren Wu, 1998, The "hump-shaped" mean term structure of interest rate derivative vols, manuscript, Fordham University.

Beaglehole, D. R., and M.S. Tenney, 1991, General solution of some interest rate-contingent claim pricing equations, *Journal of Fixed Income* 1, 69–83.

Beaglehole, D. R., and M.S. Tenney, 1992, A nonlinear equilibrium model of term strutcures of interest rates: Corrections and additions, *Journal of Financial Economics* 32, 345–454.

Bekaert, Geert, Robert Hodrick, and David Marshall, 1996, Peso problem explanations for term structure anomalies, NBER Working Paper 6147 Stanford University.

Bekaert, Geert, and Robert J. Hodrick, 2000, Exepctations hypothesis tests, *Journal of Finance* forthcoming.

Bekaert, Geert, Robert J. Hodrick, and David A. Marshall, 1997, On biases in tests of the exepctations hypothesis of the term structure of interest rates, *Journal of Financial Economics* 44, 309–348.

Bjórk, Tomas, and Bent Jesper Christensen, 1999, Interest rate dynamics and consistent forward rate curves, *Mathematical Finance* 9, 323–348.

Bouchaud, Jean Philippe, Rama Cont, Nicole El-Karoui, Marc Potters, and Nicolas Sagna, 1998, Phenomenology of the interest rate curve, *Applied Mathematical Finance* forthcoming.

Bühler, Wolfgang, Marliese Uhrig-Homburg, Ulrich Walter, and Thomas Weber, 1997, An empirical comparison of forward and spot-rate models for valuing interest-rate options, manuscript, University of Mannheim.

Campbell, John, 1995, Some lessons from the yield curve, *Journal of Economic Perspectives* 9, 129–152.

Campbell, John, and Robert Shiller, 1991, Yield spreads and interest rate movements: A bird's eye view, *Review of Economic Studies* 58, 495–514 Special Issue: The Econometrics of Financial markets.

Chen, Ren-Raw, and Louis Scott, 1993, Maximum likelihood estimation of a multifactor equilibrium model of the term structure of interest rates, *Journal of Fixed Income* 3, 14–31.

Constantinides, G. M., 1992, A theory of the nominal term structure of interest rates, *Review of Financial Studies* 5, 531–552.

Cox, John C., Jonathan E. Ingersoll, and Stephen R. Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385–408.

Dai, Qiang, and Kenneth Singleton, 2000a, Expectation puzzles, time-varying risk premia, and dynamic models of the term structure, manuscript, New York University and Stanford Univeristy.

Dai, Qiang, and Kenneth Singleton, 2000b, Specification analysis of affine term structure models, *Journal of Finance* 55, 1943–1978.

Duffee, Gregory R., 1999, Forecasting future interest rates: Are affine models failures, manuscript, University of California — Berkeley.

Duffie, Darrell, and Rui Kan, 1996, A yield-factor model of interest rates, *Mathematical Finance* 6, 379–406.

Duffie, Darrel, and Kenneth Singleton, 1997, An econometric model of the term structure of interest rate swap yields, *Journal of Finance* 52, 1287–1322.

Dybvig, Philip H., Jonathan E. Ingersoll, and Stephen A. Ross, 1996, Long forward and zero-coupon rates can never fall, *Journal of Business* 69, 1–25.

El Karoui, N., R. Myneni, and R. Viswanathan, 1992, Arbitrage pricing and hedging of interest rate claims with state variables: I theory, working paper, University of Paris.

Evans, Martin, and Karen Lewis, 1994, Do stationary risk premia explain it all? evidence from the term structure, *Journal of Monetary Economics* 33, 285–318.

Fama, Eugene F., 1984, The information in the term structure, *Journal of Financial Economics* 13, 509–528.

Fama, Eugene F., and Robert Bliss, 1987, The information in long-maturity forward rates, *American Economic Review* 77, 680–692.

Filipović, Damir, 2000, Consistency problems for HJM interest rate models, Phd dissertation Swiss Federal Institute of Technology Zurich.

Filipović, Damir, 2001, Separable term strutcures and the maximal degree problems, maniscript, ETH, Switzerland.

Frachot, Antoine, and Jean-Philippe Lesne, 1994, Expectation hypothesis and stochastic volatilities, manuscript, Banque de France.

Gallant, A. Ronald, and George E. Tauchen, 1996, Which moments to match?, *Econometric Theory* 12, 657–681.

Hamilton, James D., 1994, *Time Series Analysis.* (Princeton University Press Princeton, New Jersey).

Hardouvelis, Gikas, 1994, The term structure spread and future changes in long and short rates in the G-7 countries, *Journal of Monetray Economics* 33, 255–283.

Heath, David, Robert Jarrow, Andrew Morton, and Mark Spindel, 1992, Easier done than said, *Risk* 5, 77–105.

Holmquist, Björn, 1996, Expectations of products of quadratic forms in normal variates, *Stochastic Analysis and Applications* 14, 149–164.

Hull, John, and Alan White, 1996, Using Hull-White interest rate trees, *Journal of Derivatives* 3, 26–36.

Jamshidian, F., 1996, Bond, futures and option valuation in the quadratic interest rate model, *Applied Mathematical Finance* 3, 93–115.

Johnson, N. L., and S. Kotz, 1970, *Distributions in Statistics: Continuous Univariate Distributions 2.*Houghton Miffin Co. Boston chap. 29.

Kathri, C. G., 1980, Quadratic forms in normal variables, in P. R. Krishnaiah, eds.: *Handbook of Statistics* (North-Holland Publisher Co., Amsterdam).

Khan, Rui, 1991, Fixed income risk modeling, in *The Handbook of Fixed Income Securities* (McGraw-Hill, New York).

Knez, Peter J., Robert Litterman, and Jose Scheinkman, 1994, Explorations into factors explaining money market returns, *Journal of Finance* 49, 1861–1882.

Leippold, Markus, and Liuren Wu, 1999a, Asset pricing under the quadartic class, manuscript, University of St. Gallen/Fordham University.

Leippold, Markus, and Liuren Wu, 1999b, The potential approach to bond and currency pricing, manuscript, University of St. Gallen/Fordham University.

Litterman, Robert, and Jose Scheinkman, 1991, Common factors affecting bond returns, *Journal of Fixed Income* 1, 54–61.

Longstaff, Francis A., 1989, A nonlinear general equilibrium model of the term structure of interest rates, *Journal of Financial Economics* 23, 195–224.

Longstaff, Francis A., 2000, The term strutcure of very short-term rates: New evidence for the expectation hypothesis, *Journal of Financial Economics* forthcoming.

Mathai, A. M., and S. B. Provost, 1992, *Quadratic Forms in Random Variables.* (Marcel Dekker New York).

Mercurio, Fabio, and Juan M. Moraleda, 1996, A family of humped volatility structures, working paper, Erasmus University Rotterdam, The Netherlands.

Mishkin, Frederick, 1988, The information in the term structure: Some further results, *Journal of Applied Econometrics* 3, 307–314.

Moraleda, Juan M., and Ton C. F. Vorst, 1997, Pricing american interest rate claims with humped volatility models, *Journal of Banking and Finance* 21, 1131–1157.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Ritchken, Peter H., and I. Chuang, 1996, Option pricing and volatility humps, manuscript, Case Western Reserve University.

Roberds, William, and Charles Whiteman, 1999, Endogenous term premia and anomalies in the term structure of interest rates: Explaining the predictability smile, *Journal of Monetary Economics* 44, 555–580.

Rogers, L. C. G., 1997, The potential approach to the term structure of interest rates and foreign exchange rates, *Mathematical Finance* 7, 157–176.

Roll, Richard, 1970, *The Behavior of Interest Rates.* (Basic Books New York).

Sewell, Granville, 1988, *The Numerical Solution of Ordinary And Partial Differential Equations.* (Academic Press, Inc Boston).

Singleton, Kenneth, 1999, Estimation of affine asset pricing models using the empirical characteristic function, *Journal of Econometrics* forthcoming.

Table I
Summary Statistics of Interest Rates

Entries report summary statistics of annualized spot rates in euro US dollar and euro Deutsche Mark. The spot rates are estimated by the extended Nelsen-Siegel method from LIBOR rates and swap rates. Data are monthly, from April 1987 to September 2000 (162 observations). Our estimates replace population moments with sample moments.

Mat	Mean	Std	Skew	Kurt	Autocorrelation				
(Month)	(%)	(%)			1	2	3	4	5
				Panel A: US Dollar					
1	5.923	1.709	0.330	2.604	0.978	0.964	0.947	0.922	0.899
3	5.968	1.681	0.312	2.569	0.985	0.965	0.944	0.919	0.893
6	6.007	1.637	0.271	2.615	0.983	0.960	0.935	0.908	0.880
9	6.093	1.607	0.264	2.604	0.980	0.953	0.925	0.896	0.867
24	6.606	1.527	0.288	2.273	0.973	0.934	0.898	0.863	0.831
48	7.111	1.440	0.313	1.972	0.971	0.932	0.896	0.860	0.828
84	7.462	1.370	0.315	1.881	0.973	0.938	0.906	0.873	0.842
120	7.647	1.333	0.284	1.904	0.972	0.939	0.909	0.879	0.849
				Panel B: Deutsche Mark					
1	5.469	2.342	0.532	1.748	0.992	0.982	0.971	0.958	0.942
3	5.477	2.296	0.539	1.770	0.992	0.982	0.970	0.957	0.942
6	5.468	2.240	0.548	1.801	0.992	0.980	0.967	0.953	0.938
9	5.476	2.183	0.558	1.837	0.991	0.978	0.965	0.950	0.934
24	5.674	1.906	0.559	1.953	0.988	0.971	0.953	0.933	0.912
48	6.054	1.570	0.422	2.091	0.985	0.964	0.941	0.917	0.890
84	6.444	1.269	0.099	2.382	0.982	0.960	0.937	0.910	0.879
120	6.677	1.100	-0.165	2.622	0.980	0.960	0.937	0.908	0.878

Table II
Summary of Moment Conditions for GMM Estimation

The left hand side of the equations denotes the moment conditions while the right hand side denotes them as functions of parameters of the quadratic term structure models. μ_n denotes the mean, $\rho(k)_n$ the kth order autocorrelation, and $v(k)_n$ the variance of the k-period changes, of the n-month bond yield. g_n denotes the forward regression slope. $(A_n, \mathbf{b}_n, c_n)$ denote the coefficients of the n-month bond yield.

Mean yield curve	$\mu_n = tr(A_n V) + c_n$
Forward Regression	$c_n = \frac{2tr\left[\left(\Phi^\top A^f_{n-1}\Phi - A_r\right) V A_{frn} V\right] + \mathbf{b}^\top_{frn} V \left(\Phi^\top \mathbf{b}^f_{n-1} - \mathbf{b}_r\right)}{2tr\left(A_{frn} V\right)^2 + \mathbf{b}^\top_{frn} V \mathbf{b}_{frn}}$
Hump dynamics	$v(k)_n = 4tr\left(\left(A_n - \left(\Phi^k\right)^\top A_n \Phi^k\right) V A_n V\right) + 2\mathbf{b}_n^\top \left(I - \Phi^k\right) V \mathbf{b}_n$

Table III
Estimates of Quadratic Models on Euro US Dollars

Entries are GMM estimates (Newey-West standard errors in parentheses) of the parameters of models: (A) an unrestricted two-factor quadratic model; (B) an independent two-factor quadratic model with $\kappa(1,2) = A_\gamma(1,2) = A_r(1,2) = 0$; (C) a two-factor quadratic model with $A_\gamma = 0$; and (D) a two-factor Gaussian-affine model with $A_r = 0$. J-statistics, degree of freedom, p-value, and the probability that the short rate becomes negative are also reported.

Parameter	A		B		C		D	
$\kappa(1,1)$	0.527	(0.347)	4.443	(26.578)	0.511	(0.339)	0.620	(1.842)
$\kappa(1,2)$	−6.277	(2.797)	0		−6.309	(0.542)	−6.625	(2.910)
$\kappa(2,2)$	0.548	(0.349)	0.021	(0.009)	0.526	(0.321)	0.669	(1.781)
$A_r(1,1)$	0.023	(0.061)	0.026	(0.029)	0.025	(0.076)	0	
$A_r(1,2)$	−0.025	(0.131)	0		−0.041	(0.036)	0	
$A_r(2,2)$	0.210	(0.859)	0.122	(0.010)	0.391	(0.803)	0	
$\mathbf{b}_r(1)$	−0.266	(0.148)	−0.101	(0.059)	−0.178	(0.147)	−0.440	(0.274)
$\mathbf{b}_r(2)$	0.192	(0.465)	0.873	(0.759)	0.127	(0.135)	0.250	(0.696)
$A_\gamma(1,1)$	0.173	(0.348)	−3.923	(0.076)	0		0.148	(0.970)
$A_\gamma(1,2)$	0.998	(2.049)	0		0		−1.400	(4.241)
$A_\gamma(2,2)$	−0.379	(0.218)	−0.003	(0.003)	0		−0.619	(0.798)
$\mathbf{b}_\gamma(1)$	3.948	(0.417)	4.288	(0.474)	3.289	(0.520)	1.401	(1.539)
$\mathbf{b}_\gamma(2)$	−0.036	(0.157)	0.005	(0.046)	0.068	(0.093)	0.176	(0.096)
J-statistic	113.42		304.22		179.39		207.02	
Deg of Fr, v	26		29		29		29	
p-value	0		0		0		0	
$\Pr(r < 0)$	0		0		0		23.21%	