Stochastic volatility, mean drift, and jumps in the short-term interest rate

Stochastic Volatility, Mean Drift, and Jumps in the Short-Term Interest Rate

TORBEN G. ANDERSEN, LUCA BENZONI, and JESPER LUND*

First Draft: April 2003
This Draft: April 2004

Abstract

We find that an intuitively appealing and fairly manageable continuous-time model provides an excellent characterization of the U.S. short-term interest rate over the post Second World War period. Our three-factor jump-diffusion model consists of elements embodied in existing specifications, but our approach appears to be the first to successfully accommodate all such features jointly. Moreover, we conduct simultaneous and efficient inference regarding all model components which include a shock to the interest rate process itself, a time-varying mean reversion factor, a stochastic volatility factor and a jump process. Most intriguingly, we find that the restrictions implied by an affine representation of the jump-diffusion system are not rejected by the U.S. short rate data. This allows for a tractable setting for associated asset pricing applications.

---

*Torben G. Andersen is at the Kellogg School of Management, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208, 847-467-1285, t-andersen@kellogg.northwestern.edu and the NBER. Luca Benzoni is at the Carlson School of Management, University of Minnesota, 321 19th Ave S, Minneapolis, MN 55455, 612-624-1075, lbenzoni@umn.edu. Jesper Lund is at Nykredit Bank, Quantitative Research, DK-1780 Copenhagen V, Denmark, 45-3342-1378, mail@jesperlund.com. We are grateful to Bjørn Eraker; Lorenzo Garlappi; Bob Goldstein; Lars Hansen; Bob Kimmel; Mike Johannes; Koji Kusuda; James MacKinnon; Nour Meddahi; Ken Singleton; Cristian Tiu; and seminar participants at the New York Federal Reserve; the University of Minnesota; the University of Rochester; the University of Texas at Austin; the University of Wisconsin at Madison; the Washington University in St. Louis; the CIRANO-CIREQ-MITACS 2003 Financial Econometrics Conference; the 2004 Econometric Society Conference in San Diego for helpful comments and suggestions. We are indebted with Huyan Qiu, who provided outstanding research assistance. Of course, all errors remain our sole responsibility. The most recent version of the paper can be downloaded from `http://www.umn.edu/~lbenzoni`

1 Introduction

Understanding the dynamics of the U.S. "short rate," i.e., the risk-free instantaneous yield or spot rate is important for both practical and theoretical reasons. The short rate governs the price of riskless fund transfers and is thus a key determinant of the intertemporal consumption and investment decisions of economic agents. In addition, the short rate impacts the expected returns of primary assets whose excess returns (over the short rate) are functions of systematic risk exposures and associated risk premia. Finally, the short rate serves as a direct input to pricing and hedging in the huge fixed-income securities market and the associated trading of fixed-income derivatives. Consequently, it is not surprising that short rate modeling has been an active research area for decades.

In spite of the voluminous literature, the consensus view is that existing models fail to capture important features of the short-term interest rate dynamics. An impetus for much recent work is Chan, Karolyi, Lonstaff, and Sanders (1992) (hereafter CKLS) who stress the empirical difficulties of one-factor continuous-time specifications within the Vasicek (1977) and Cox, Ingersoll and Ross (1985) class of models. This helped inspire a large body of research on nonparametric continuous-time models for the short rate; see, e.g., Aït-Sahalia (1996a and b); Conley, Hansen, Luttmer, and Scheinkman (1997) (henceforth CHLS); Jiang and Knight (1997); and Stanton (1997). Among these contributions, Aït-Sahalia (1996b) considers several specifications of the seven-day Eurodollar rate and concludes that the principal source of rejection of existing models is the strong nonlinearity of the drift coefficient. He finds that around its mean, where the drift is essentially zero, the spot rate behaves like a random walk. The drift then mean-reverts strongly when far away from the mean. Similar results are reported in Jiang and Knight (1997), based on a sample of Canadian interest rates and CHLS for a sample of U.S. Federal funds rates. CHLS stress that this dynamic behavior does not necessarily imply non-linearity, or even mean-reversion, of the drift coefficient, but instead may result from so-called volatility-induced stationarity governed mainly through the specification of the diffusion coefficient.

More recently, Bandi (2002) also argues that the constant-elasticity-of-variance model with linear drift fails due to the martingale nature of the short rate over most of its empirical range, with the indications of nonlinearities in the drift term driven exclusively by a small number of observations at the edges of the sample domain. Durham (2002) and Jones (2003) both conclude that the apparent nonlinearity of the drift coefficient may be induced by the misspecification of standard one-factor models and they conjecture that more general drift and diffusion specifications are necessary. Specifically, Jones (2003) advocates a stochastic drift model, while Durham (2002) emphasizes a more general representation of the diffusion term, possibly allowing for the presence of stochastic volatility.

Multi-factor models have previously been used in the attempt of capturing the short rate dynamics;

see, e.g., Andersen and Lund (1997); Balduzzi, Das, Foresi and Sundaram (1996); Jegadeesh and Penacchi (1996); and Longstaff and Schwartz (1992). These contributions find that the presence of latent factors improves the fit for the short rate data significantly relative to the CKLS model. Nevertheless, the consensus remains that the spot rate eludes standard parametric specifications, including those within the so-called "affine class" commonly used in asset pricing applications.[1]

In contrast, we document that the main features of the U.S. short-term interest rate, i.e., complex conditional heteroskedasticity, fat-tailed innovations and pronounced autocorrelation patterns, may be captured within an intuitively appealing and manageable continuous-time jump-diffusion setting. Relative to earlier contributions, we explore a more general parametric class of continuous-time models, allowing for *both* multiple latent factors, entering separately in the drift and diffusion coefficients, *and* jumps in the interest rate level. Furthermore, we investigate alternative representations both within and outside of the affine class. This is reflected in our estimation approach that is not tailored to a specific functional form and readily handles latent factors. Thus, our analysis implicitly provides an assessment of the severity of the constraints that routinely are imposed by inference techniques explicitly designed for affine representations. We also seek to identify the role that the different model components play in capturing the dominant features of the data. Finally, we provide a qualitative illustration of some implications of our estimated system for the term structure of interest rates.

Our empirical analysis exploits a long sample of weekly observations on the 3-month U.S. T-Bill yield spanning 1954 to 2000. We estimate the system through a variant of the simulated method of moments (SMM) introduced by Duffie and Singleton (1993). Specifically, we rely on the efficient method of moment (EMM) technique of Gallant and Tauchen (1996) to guide the selection of moment conditions. This ensures that the moments summarize the key statistical features of the data, greatly enhancing the efficiency of the inference. Finally, EMM offers a variety of powerful specification tests that allow us to assess the quality of the model fit as well as likely sources of misspecification. Since different models are confronted with the identical set of (informative) moment conditions, we may readily assess the relative performance of non-nested specifications.

Both of our favored models contain three factors featuring stochastic volatility, mean drift and jumps. The inclusion of the stochastic volatility factor is critical for a good fit, whereas the stochastic mean offers a more modest, but still significant, improvement. Specifically, we find it important to allow for a relatively fast mean-reversion of the short rate around a highly persistent time-varying central tendency process. Economically, the mean drift may be indicative of slowly evolving inflationary expectations, time-variation in the required real interest rate, or both. Furthermore, we find jumps

[1]Other recent contributions that have investigated the short term interest rate include Das (2002); Eraker (2001); Elerian, Chib, and Shephard (2001); Hamilton (1996); Johannes (2004); Li, Pearson, and Poteshman (2002), and more.

to be integral to the quality of fit and to relieve the stochastic volatility factor from accommodating extreme outlier behavior. Without jumps, the estimated volatility specifications are unable to match the degree of persistence in the conditional variance process typically observed in short-term interest rate series. Interestingly, we find little evidence of the so-called level effect. Moreover, our analysis suggests that correlation between the shocks to the short rate process and the other two latent factors is not critical in modeling the short rate series. Overall, there is no indication of misspecification in our preferred models. Along some relevant dimensions, our affine three-factor jump specification provides the superior fit, which lends support to jump-diffusion representations of the form suggested by Duffie, Pan, and Singleton (2000) and Chacko and Das (2002).

In summary, we argue that all main characteristics of the short-term interest rate process can be captured with an intuitively appealing and fairly manageable continuous-time model. This finding is especially interesting since one, from prior evidence, may have concluded that the short rate dynamics eludes models within a standard parametric jump-diffusion framework. We document a need for both the type of stochastic volatility and mean drift specifications explored separately in, e.g., Andersen and Lund (1997), Durham (2002) and Jones (2003). However, even jointly, these do not suffice for a satisfactory representation. In order to achieve a reasonable fit to our long U.S. short rate series we must also introduce a jump component. Even with this added complexity, one of our specifications remains within the affine model class and therefore provides a tractable basis for a broad range of asset pricing questions.

The remainder of the paper is structured as follows. Section 2 introduces our candidate models for the short-term interest rate, while Section 3 compares our approach to that pursued in other recent contributions. Empirical results and the details of the EMM implementation are documented in Section 4. Section 5 illustrates potential implications for term structure modeling and compares our qualitative findings with stylized empirical evidence from the term structure literature. Concluding remarks are in Section 6.

2 Candidate Models

We have two main objectives in mind when formulating our short rate models. First, we strive to retain an intuitive interpretation of the various features of the models, which facilitates a comparison to the extant literature. Second, we aim to preserve the tractability necessary for asset pricing applications and allow for rather straightforward economic interpretation. These objectives lead us towards two distinct, but related types of models. In their most general form, they fall, respectively, outside and within the affine class.

The first family of models takes the form:

$$dr_t = \kappa_1(\mu_t - \overline{\kappa}\,\lambda\, r_t\,/\kappa_1 - r_t)\,dt + \sqrt{V_t}\,r_t^{\gamma}\,dW_{1,t} + (\mathrm{e}^{Z_t} - 1)r_t\,dq_t\,, \quad \gamma \geq 0\,, \tag{1}$$

$$d\ln V_t = \kappa_2(\alpha - \ln V_t\,)\,dt + \eta_1\,dW_{2,t}\,, \tag{2}$$

$$d\mu_t = \kappa_3(\vartheta - \mu_t\,)\,dt + \eta_2\,\sqrt{\mu_t}\,dW_{3,t}\,, \tag{3}$$

where $W_i, i = 1,2,3$, are Brownian motions with correlation coefficients $\mathrm{corr}(dW_{1,t}, dW_{i,t}) = \rho_{1,i}$, $i = 2,3$, while q is a Poisson process uncorrelated with W_i, $i = 1,2,3$, and governed by the jump intensity parameter λ, i.e., $\mathrm{Prob}(dq_t = 1) = \lambda$. In the event that $dq_t = 1$, the short rate is subject to a jump, which is expressed as a fraction $(\mathrm{e}^{Z_t} - 1)$ of the current rate r_t, where $Z_t \rightsquigarrow \mathrm{N}(\mu_J, \sigma_J)$ is independent of q and W_1-W_3. It then follows that, conditional on a jump or equivalently $dq_t = 1$, the average jump size is

$$\overline{\kappa} = \mathrm{E}(\mathrm{e}^{Z_t} - 1) = \mathrm{e}^{(\mu_J + \sigma_J^2/2)} - 1\,. \tag{4}$$

Within the family (1)-(3), we recognize several well-known special cases. If we assume that the volatility and mean are constant, $\sqrt{V_t} = \sigma$ and $\mu_t = \mu$, and exclude jumps, $\lambda = 0$, we obtain the one-factor CKLS model:

$$dr_t = \kappa_1(\mu - r_t)\,dt + \sigma\,r_t^{\gamma}\,dW_{1,t}\,, \quad \gamma \geq 0\,. \tag{5}$$

According to this representation, r_t reverts towards the mean level μ, with κ_1 measuring the speed of the mean reversion. Subject to appropriate parameter restrictions, the presence of the r_t^{γ} term in the diffusion coefficient rules out negative rates and induces conditional heteroskedasticity, as volatility depends on the level of the short rate, the so-called *level effect*. Further imposing $\gamma = 0.5$ results in the affine one-factor Cox, Ingersoll, and Ross (1985) (CIR) model, while $\gamma = 0$ produces the Vasicek (1977) specification.

Relaxing the restriction that μ is constant yields the "central tendency" model, which is a special case of (1) given by

$$dr_t = \kappa_1(\mu_t - r_t)\,dt + \sigma\,r_t^{\gamma}\,dW_{1,t}\,, \quad \gamma \geq 0\,, \tag{6}$$

in combination with (3). The specification (6) and (3) allows for a time-varying mean level of the short rate to dictate the short-run mean reversion in the drift. This extension is consistent with the indications in Figure 1 that, while there are extended periods of strong drift, mean reversion is not simply related to a pull towards the overall mean of around 5.6%. The series often appear to drift downwards, even when rates are considerably below the mean, and to drift upwards even when rates are well above the mean. It thus appears as if r_t converges rather quickly towards a time-varying mean level, μ_t, whereas μ_t itself reverts more slowly towards the unconditional mean, ϑ. The specification resembles the central-tendency representations proposed originally by Jacobs and Jones (1986), and later by Beaglehole and Tenney (1991), Balduzzi et al. (1995), Jegadeesh and Penacchi (1996), and Jones (2003). The model is also similar in spirit to that of Brennan and Schwartz (1979), although their second factor is the consol yield, which is a traded asset (a bond price). An obvious interpretation of the time-varying mean is that of an (expected) inflation factor that, in accordance with the Fisher effect, manifests itself in the required nominal interest rate. Further drift may be induced by a slow evolution in the equilibrium real interest rate.

The central tendency model is easily extended to the case of a time varying volatility process. The stochastic volatility and mean drift model continues to be a special case of our family (1)-(3). Specifically, it is given by

$$dr_t = \kappa_1(\mu_t - r_t)\, dt + \sqrt{V_t}\, r_t^{\gamma}\, dW_{1,t}\,, \qquad \gamma \geq 0\,, \tag{7}$$

combined with equations (2)-(3). Andersen and Lund (1997) and Durham (2002) estimate versions of this model in which $\mu_t = \mu$. The stochastic volatility specification of Andersen and Lund (1997) is unusual in the interest rate literature, but it is inspired by the success of similar formulations for discrete-time asset returns, and discretized versions of this short rate model have been estimated by Kearns (1992), Torous and Ball (1995) and Vetzal (1996). Moreover, it resembles the stochastic volatility models of Fong and Vasicek (1991), and Chen (1996). If log-volatility in equation (2) is serially correlated, then ARCH-type effects will interact with the level effect, creating rather complex volatility dynamics.

Finally, extending the stochastic volatility and mean drift model to allow for jumps produces the general non-affine specification (1)-(3). Theory predicts that an unexpected information arrival should induce a discontinuity in the return process. For example, jumps in interest rates could be induced by an announced shift in monetary policy. Moreover, from a statistical perspective, adding a jump component may improve the descriptive power of our models as jumps can accommodate outliers in the short rate distribution.

In recent years, the term structure literature has relied predominantly on models in the so-called affine class; see, e.g., Duffie and Kan (1996). Thus, we also consider representations of the form

$$dr_t = \kappa_1(\mu_t - \mu_J\,\lambda\,/\kappa_1 - r_t)\, dt + \sqrt{V_t}\, dW_{1,t} + Z_t\, dq_t\,, \tag{8}$$

$$dV_t = \kappa_2(\alpha - V_t\,)\, dt + \eta_1 \sqrt{V_t}\, dW_{2,t}\,, \tag{9}$$

$$d\mu_t = \kappa_3(\vartheta - \mu_t\,)\, dt + \eta_2 \sqrt{\mu_t}\, dW_{3,t}\,, \tag{10}$$

where W_1-W_3 and q satisfy the same assumptions as those stated for (1)-(3), except that here we assume $\mathrm{corr}(dW_{1,t}, dW_{i,t}) = 0$, $i = 2, 3$. Furthermore, we assume $Z_t \rightsquigarrow \mathrm{N}(\mu_J, \sigma_J)$. The system (8)-(10) extends the two- and three-factor models of Fong and Vasicek (1991), Chen (1996) and Balduzzi, Das, Foresi, and Sundaram (1996) by including a jump component and it extends the jump-diffusion specifications of Das and Foresi (1996) by incorporating stochastic volatility and mean factors. Still, it falls within the model class considered by Chacko and Das (2002), Duffie, Pan, and Singleton (2000) and Piazzesi (2003b), and thus provides a tractable setting for bond and option pricing applications.

The model (8)-(10) contains most of the features in (1)-(3). Thus, the interpretation of the specific components discussed for the prior system still applies here, although there also are some important differences. In order to preserve the affine structure of the model, we fix the γ coefficient at zero, thus suppressing the level effect. Further, in the system (8)-(10) the Brownian motion shocks are uncorrelated and jumps are normally distributed. This implies that the short rate can become negative. In practice, however, when the model parameters are fixed at reasonable values, the likelihood of observing negative rates is very small for most realizations of the state variables. Thus, this problem may be secondary compared to the advantage offered by the semi closed-form bond pricing formulas it delivers.[2] Ultimately, it is an empirical issue to assess the descriptive value of the affine model relative to non-affine specifications such as (1)-(3), which offer some additional desirable properties but are less practical for pricing applications.

3 Recent Developments in the Literature

The estimation of continuous-time diffusion models poses technical challenges. Early methods rely on the discretization of the underlying process and are subject to an estimation bias as documented by, e.g., Lo (1988). Further, it is often the case that the lack of analytical expressions for the models' transition densities as well as the presence of latent variables renders traditional estimation techniques infeasible. However, novel developments in continuous-time econometrics have spurred renewed interest in the estimation of interest rate diffusion models. Among the recent contributions, we have the nonparametric approaches of Aït-Sahalia (1996a and b), Conley, Hansen, Luttmer and Scheinkman (1997), Hansen and Scheinkman (1995), Jiang (1998), Jiang and Knight (1997), and Stanton (1997). One advantage of nonparametric techniques is that such methods lend themselves readily to the estimation of nonlinear models. This has produced an intense debate about the functional form of the drift and diffusion terms in one-factor interest rate models. Aït-Sahalia (1996a and b) and Stanton (1997) find that the degree of mean reversion is stronger in the tails of the interest rate process. On the other hand, other authors warn that such evidence should be interpreted with caution, as nonparametric methods may be subject to a small sample bias; see, e.g., Pritsker (1998) and Chapman

[2]See, e.g., Piazzesi (2003b) for a more extensive discussion on affine models and negative rates.

and Pearson (2000).[3]

Previous contributions have investigated the role of jumps in the short-term interest rate diffusion. Johannes (2004) estimates a one-factor jump-diffusion model nonparametrically using daily T-bill observations ranging from January 1965 through February 1999. He concludes that jumps are necessary to generate large enough moves in daily interest rates; also, he finds that jumps account for a significant portion of the interest rate volatility and serve an important economic purpose by providing a conduit for macroeconomic information to enter the term structure. Das (2002) estimates a model with normal jumps and ARCH conditional heteroskedasticity, relying on daily Fed funds rates from January 1988 through December 1997. He finds empirical support for the presence of jumps and provides evidence that their occurrence is linked to Fed activity as well as day-of-the-week effects. In contrast, we alleviate the impact of institutional arrangements such as the reserve maintenance period—discussed in detail by Hamilton (1996)—by only considering weekly observations. In addition, we assess the importance of jumps within a full-fledged continuous-time multi-factor setting, which provides a control for the influence of persistent stochastic volatility and mean drift factors on the inference. In some respects, our model is more in line with that of Piazzesi (2003a), but her work has a distinct macroeconomic regime focus. Her jumps correspond to the meeting dates of the Fed Open Market Committee meeting and release dates for macroeconomic data. Given the modeling choice, her sample is limited to five years of data on LIBOR, swap and Fed target rates. In contrast, to provide robust identification of the persistent and latent mean drift and volatility factors as well as the jump component in our full-fledged jump-diffusion model, we require a much longer sample of short-term interest rates.

Empirically, yields on zero-coupon bonds at close-by maturities are highly correlated. Litterman and Scheinkman (1991) demonstrate that virtually all variation in U.S. Treasury rates is captured by three factors, interpreted as changes in *level*, *steepness* and *curvature*. This evidence has motivated a large body of research on reduced-form term structure models, in which bond yields are expressed as an affine (or quadratic) function of a state vector. Estimation is often performed for a panel of bond yields through method of moment techniques, typically relying on simulation to deal with the presence of latent state variables. Dai and Singleton (2000), Ahn, Dittmar, and Gallant (2002), and Brandt and Chapman (2002) use EMM in their applications. Duffee and Stanton (2001) raise concerns about the finite-sample properties of EMM estimates obtained from a multi-dimensional score generator. To avoid this issue, Duffee (2002) and Collin-Dufresne, Goldstein and Jones (2003)

[3]Other recent papers that have studied the properties of the drift and diffusion terms in diffusion models for the short-term interest rate include, e.g., Bandi (2002), Das (2002), Jones (2003), Elerian, Chib, and Shephard (2001), Durham (2002), and Li, Pearson, and Poteshman (2002).

(henceforth CGJ) rely on quasi-maximum likelihood for model estimation. Likewise, Brandt and Chapman (2002) emphasize the trade-off between robustness and efficiency, and advocate an SMM approach using economic stylized facts to guide the choice of moment conditions, rather than relying on the score moments associated with a likelihood function. Our paper differs from the above along several dimensions. First, our factors have a clear statistical interpretation, which adds to the model transparency and facilitates the specification analysis. Second, the above models do not include jumps, which we deem crucial in obtaining a satisfactory fit. Third, their estimates rely on a panel of yields, while we focus on the short rate alone. Although our approach is, in theory, inefficient in this regard, it has several important advantages, which we now discuss.

A first advantage derives directly from the special position of the spot rate within the yield curve. Since the yields at longer maturities determine expected returns on fixed-income securities, their pricing will generally reflect the uncertain evolution of the underlying state variables and will thus depend upon the level and the dynamics of market factor risk premia. Thus, term structure modeling induces an additional layer of functional forms that need to be asserted and estimated. In contrast, short rate models avoid this complexity as performance is evaluated strictly under the objective probability measure. An important implication is that we can assess whether features like jumps as well as stochastic volatility and mean drift—with relevant term structure implications—are fundamental features inherent in the short rate. Furthermore, and related, we avoid making assumptions regarding the factor risk premia. Thus, we avoid any ambiguity as to whether rejection of models occurs because of a misspecification of the underlying dynamics, of the risk premia, or both. Second, our approach does not rely on a multi-dimensional score generator, and thus it is not subject to the concern of Duffee and Stanton (2001) about the finite sample properties of EMM estimates obtained using a multi-dimensional auxiliary model.

Third, and most importantly, forcing a multi-factor model to simultaneously fit the time-series and cross-sectional properties of bond yields will often produce outright counterfactual implications for the short rate dynamics. For example, CGJ argue that the model-implied spot-rate volatility obtained by inverting the yield curve using an affine multi-factor specification can be negatively correlated with the time series of volatility estimated with a standard GARCH approach. The problem occurs because the volatility state variable represents both a combination of yields (thus impacting the cross section of bond prices) and the quadratic variation of the short rate (thus impacting the time series of yields). In response, they advocate an "unspanned stochastic volatility" (USV) model, in which the cross section of bond prices is independent of volatility. Hence, the identification of the volatility factor now depends entirely on time-series data, as in our estimation procedure.[4] More

[4] Another reaction is exemplified by Diebold and Li (2002), who forecast the term structure from a standard three-

recently, Thomson (2003) has applied a new class of specification tests to term structure models of the LIBOR swap curve. Consistent with CGJ, he concludes that volatility is poorly identified by the cross section of bond yields. Even slight forms of model misspecification can induce large errors in imputed volatility. He documents three main problems for affine models at the short end of the yield spectrum: they ignore outliers in the data, they fail to capture the time-varying spot rate mean, and they err on the variance dynamics. And the forecasts of the conditional variance are especially inaccurate. Thus, he argues that time-series identification of volatility is more robust to model misspecification. We conclude that critical features of the short rate process appear incompatible with standard term structure models. Moreover, the misspecification is directly related to the need for a proper accommodation of the time-varying mean and variance as well as the outliers in the short rate data. These are, of course, the exact features that our empirically driven estimation procedure inherently is forced to focus on. Hence, until more robust implementations of term structure models are developed, there are good reasons to separate time-series estimation of the short rate dynamics from term structure modeling.

Several recent contributions extend the maximum likelihood approach for continuous-time model estimation. Aït-Sahalia (2002a, b) develops closed-form Hermite series expansions for the likelihood function of multivariate diffusions. Such approach is applied in Aït-Sahalia and Kimmel (2002) to obtain a closed-form expression for the likelihood function of the canonical affine models in Dai and Singleton (2000). Schaumburg (2002) extends Aït-Sahalia and Kimmel's approach for the purpose of conducting maximum likelihood estimation of diffusions driven by Levy-type processes. Of course, in the presence of dynamic latent (stochastic volatility) variables these likelihood-based techniques still face the non-trivial practical problem of integrating out the unobserved variables. In addition, Bates (2003) derives a rule to recursively update the joint characteristic function of latent variables and the data conditional upon past data. He applies the method to estimate an affine continuous-time jump-diffusion model using a sample of daily S&P 500 index returns. At the same time, considerable progress has been done with simulation-based Bayesian methods. Recent applications of the Monte Carlo Markov Chain (MCMC) estimation technique towards estimating continuous-time models for interest rates include Eraker (2001), Jones (2003), Elerian, Chib, and Shephard (2001). To our knowledge, however, none of these methods has been applied successfully to estimate a multi-factor continuous-time jump-diffusion model for interest rates.

dimensional VAR based on extracted level, slope and curvature factors without explicitly imposing any no-arbitrage constraints. They report a significant improvement in the out-of-sample forecast performance relative to currently popular term structure models as well as a random walk benchmarks.

4 Empirical Results

In this section, we report on our EMM implementation. Section 4.1 outlines the semi-nonparametric (SNP) model selection procedure used to obtain the moment conditions for the subsequent EMM estimation step. Section 4.2 details our EMM estimation and inference results.

4.1 The SNP Model

The finite-sample performance of the EMM procedure hinges on a sensible choice of the moment conditions used in the simulation-based EMM estimation step. The family of SNP densities introduced by Gallant and Nychka (1987) provides a natural starting point for this task. Loosely speaking, Gallant and Long (1997) show that the score function of an SNP density asymptotically spans the score of the true model, suggesting that the EMM methodology is asymptotically efficient when the order of the SNP model is expanded until an adequate statistical representation of the data is obtained.

In choosing our SNP model, we follow a careful model selection procedure. Andersen, Chung, and Sørensen (1997) find that the finite sample properties of the EMM estimates may deteriorate if an indiscriminate moment selection procedure is followed. To reduce the risk that random sample variation is encoded in the score vector by overfitting the auxiliary model, we use a Gaussian leading term, designed to capture the bulk of the dependency in the conditional mean and variance of the series. Next, we allow a squared Hermite polynomial expansion to accommodate any remaining non-normality and possible time-series dependency in the innovation process. An ARMA form is the natural candidate to capture the rich dynamics in the short rate conditional mean, while an ARCH-type representation generally provides a reasonable characterization of the conditional heteroskedasticity in interest rate data. Specifically, we adopt an EGARCH form. This choice is motivated not only by goodness-of-fit criteria, but also by the analysis in Andersen and Lund (1997), which shows that an EGARCH score generator has better "dynamic stability" properties than, e.g., a GARCH model, a consideration which proves important in the EMM implementation. Finally, we allow for an additional source of interaction via the interest level effect, i.e., we scale the EGARCH conditional variance term by $r_t^{2\delta}$. This last extension may be helpful in parsimoniously capturing the underlying dynamics of interest rate volatility and in identifying the corresponding level effect in the continuous-time models. In sum, this leads to the class of SNP densities:

$$f_K(r_t|x_t;\xi) = \left(\nu + (1-\nu\,) \times \frac{[P_K(z_t,x_t)]^2}{\int_{\mathrm{R}}[P_K(z_t,x_t)]^2\phi(u)du}\right)\frac{\phi(z_t)}{r_{t-1}^{\delta}\sqrt{h_t}}\,, \qquad \nu = 0.01\,, \tag{11}$$

where $\phi(.)$ is the standard normal density, $x_t = \{r_1,\ldots,r_{t-1}\}$ reflects the information set, ξ is the

SNP density parameter vector,

$$
\begin{aligned}
z_t &= \frac{r_t - \mu_t}{r_{t-1}^{\delta}\sqrt{h_t}}\,, \\
\mu_t &= \phi_0 + \sum_{i=1}^{s}\phi_i r_{t-i} + \sum_{i=1}^{u}\zeta_i(r_{t-i} - \mu_{t-i})\,, \\
\ln h_t &= \omega\,(1 - \sum_{i=1}^{p}\beta_i) + \sum_{i=1}^{p}\beta_i \ln h_{t-i} + (1 + \alpha_1 L + ... + \alpha_q L^q)\,[\,\theta_1 z_{t-1} + \theta_2\,(b(z_{t-1}) - \sqrt{2/\pi})\,]\,, \\
b(z) &= |z| \;\text{ for } |z| \geq \pi/2K, \quad b(z) = (\pi/2 - \cos(Kz))/K \;\text{ for } |z| < \pi/2K\,, \\
P_K(z,x) &= \sum_{i=0}^{K_z} a_i(x) z^i = \sum_{i=0}^{K_z}\left(\sum_{|j|=0}^{K_x} a_{ij}x^j\right) z^i\,, \qquad a_{00} = 1\,,
\end{aligned}
$$

where j is a multi-index vector, $x^j \equiv (x_1^{j_1}, \ldots, x_M^{j_M})$, and $|j\,| \equiv \sum_{m=1}^{M} j_m$. As in Andersen and Lund (1997), $b(z)$ is a smooth, twice-differentiable, function that closely approximates the absolute value operator in the EGARCH variance equation, with $K = 100$.

We estimate the SNP densities in (11) by (quasi-)maximum likelihood (QML). In this application, our proxy for the riskless short rate is the three-month Treasury bill yield. Chapman, Long, and Pearson (1999) show that the bias induced by the three-month interest rate proxy is economically small for linear multi-factor models of interest rates. Shorter maturity T-bill rates are available but are adversely affected by idiosyncratic variation, as observed by Duffee (1996). The data series is weekly, spanning 1954 to 2000, and is obtained from the H.15 release of the Federal Reserve System.[5] These rates are quoted on a bank discount basis (see, e.g., Sinkey (1989) for a definition) and we convert them into continuously compounded yields prior to analysis. Our series is limited to the weekly frequency, although daily observations are available. This minimizes the impact of missing data points, possible holiday as well as day-of-the-week effects, and other institutionally driven features.[6] Wednesdays have the least number of missing observations, so we use the reported Wednesday rate. When this observation is missing, we use the Tuesday rate instead. This ensures a valid series from January 6, 1954 to June 28, 2000. A time-series plot of the data is provided in Figure 1, while summary statistics are given in Table 1.

Our model selection procedure within the family of SNP densities (11) is guided by the Bayesian (BIC) and Hannan-Quinn (H-Q) information criteria; we assign less importance to the commonly

[5]The H.15 data are available at several Web sites supported by the Federal Reserve System; see, e.g., the URL http://www.research.stlouisfed.org/fred/fredfile.html

[6]Hamilton (1996) documents the presence of institutional features in the daily Federal Funds rate, produced by, e.g., reserve maintenance requirements. These effects likely spill over other short maturity rates: e.g., Durham (2002) finds that the daily one-week maturity rates are noisy as well. To avoid such problems, some previous studies rely on monthly data. We follow the, by now, common practice of estimating the model at a weekly frequency.

used Akaike criterion (AIC) because of its tendency to overparameterize the models. Further, we pay close attention to the Ljung-Box tests for the autocorrelation of the (raw and squared) residuals, which provide guidance in detecting possible sources of misspecification in the conditional mean and variance leading terms of our densities.

Our analysis, summarized in Table 2, points towards an ARMA(4,1)-Level-EGARCH(2,1)-Kz(6)-Kx(0). More specifically, a relatively high order ARMA term is necessary to reproduce the rich conditional mean dynamics and to eliminate most of the dependencies in the residuals. Similarly, a combined analysis of the information criteria and Ljung-Box statistics indicates the need for a second-order EGARCH term, enriched with the presence of level effects. We experimented with higher-order polynomials, but found no support for that extension. Also, we rejected a Hermite(6,1) specification for the standardized residual in which, to conserve on the number of parameters, the non-homogeneous terms of order higher than two are fixed at zero.[7] The constrained Hermite(6,1) score generator was nonetheless used to check the robustness of our EMM estimates, as discussed in Section 4.2 below. As a final specification check on the score generator, we conducted extensive simulations from our SNP conditional density. Through Monte Carlo integration, we computed the moments of a long simulated series of short rate and confirmed that they converge to values close to those observed in the actual sample.

4.2 EMM Estimation

We denote the parameter vector of our continuous-time model ψ and let $\{r_t(\psi), x_t(\psi)\}_{t=1}^{\mathcal{T}(N)}$ represent a sample simulated from the model with $x_t(\psi) = \{r_1(\psi), \ldots, r_{t-1}(\psi)\}$ reflecting the information set. The EMM estimator of ψ is then defined by

$$\hat{\psi}_N = \arg\min_{\psi} \; m_{\mathcal{T}(N)}(\psi, \hat{\xi})' \, W_N \, m_{\mathcal{T}(N)}(\psi, \hat{\xi}) \,,$$

where $m_{\mathcal{T}(N)}(\psi, \hat{\xi})$ is the expectation of the score function, evaluated by Monte Carlo integration at the quasi-maximum likelihood estimate of the auxiliary model parameter $\hat{\xi}$,

$$m_{\mathcal{T}(N)}(\psi, \hat{\xi}) = \frac{1}{\mathcal{T}(N)} \sum_{t=1}^{\mathcal{T}(N)} \frac{\partial \ln f_K(r_t(\psi)|x_t(\psi); \hat{\xi})}{\partial \xi} \,,$$

and the weighting matrix W_N is a consistent estimate of the inverse asymptotic covariance matrix of the auxiliary score function. Following Gallant and Tauchen (1996), we estimate the covariance matrix of the auxiliary score from the outer product of the gradient. In simulating the short rate sequence $\{r_t(\psi), x_t(\psi)\}_{t=1}^{\mathcal{T}(N)}$, two antithetic samples of 75,000 $\times$ 25 + 5,000 rate are generated from

[7] A similar strategy has been used in Chernov, Gallant, Ghysels, and Tauchen (2002).

the continuous-time model at time intervals of $1/25$ of a week.[8] The first 5,000 observations are discarded to eliminate the effect of the initial conditions. Lastly, a sequence of $\mathcal{T}(N) = 75,000$ weekly rates is obtained by collecting the end-of-the-week observations from the simulated sample.

4.2.1 CKLS and CIR Models

Our representations nest a host of special cases that have been studied extensively in the literature, so it is worthwhile to ponder their empirical shortcomings in some detail. A natural starting point for our investigation is the one-factor representation of CKLS or CIR form in equation (5). One caveat is, however, that all such one-factor models fare so poorly that it may be treacherous to assign much meaning to the point estimates or even qualitative features of the fit. Nonetheless, one common finding is noteworthy. When we allow the level effect coefficient, γ, to be a free parameter, we invariably estimate it at a comparatively low value and with a great deal of imprecision. The culprit is that for the SNP density the corresponding level coefficient, δ, and the long-run volatility mean, ω, are highly correlated. The problem in identifying γ separately is also noted, for a shorter sample than ours in, e.g., Conley et al. (1997), Gallant and Tauchen (1997), and Tauchen (1997). We conclude that the fully specified CKLS model is poorly identified, and going forward we fix γ at the CIR value of $1/2$, a point estimate we could reject neither in this setting nor within the more general multi-factor representations explored subsequently. The discrepancy relative to some earlier empirical findings on the size of γ is likely driven by our longer sample that refutes the strictly monotone relation between the interest rate level and volatility. For example, at the very low interest rate levels in the 1950s and early 1960s the volatility was comparatively high. Likewise, the volatility in the 1970s was much higher than in the mid 1980s and early 1990s although the level of the short rate was about the same.

Parameter estimates, standard errors, and the overall goodness-of-fit test statistic for the CIR model are reported in Table 3, while score generator diagnostics are in Table 5. Note that, henceforth, $dt = 1/52$ and parameter estimates are given for weekly interest rate data expressed in decimal form on a yearly basis, as is common in the extant literature. The long-term mean μ of the interest rate is estimated at 5.1%, roughly in line with the 5.6% mean in our sample (Table 1). The κ_1 coefficient captures the persistence of the short-term interest rate. Our κ_1 estimate implies a first order autoregressive coefficient of $\exp(-\kappa_1/52) = 0.9967$ at the weekly level, and a half-life of shocks to

[8] Estimation conducted using considerably longer simulated samples produced nearly identical results. Furthermore, when the data generating process contains a jump component, the simulation step involves an additional layer of approximation as our procedure for generating jumps renders the EMM criterion function discontinuous in the parameter vector, and this creates problems for the numerical minimization of the EMM objective function, as in Andersen, Benzoni and Lund (2002). To avoid this problem jumps are smoothed using a close continuously differentiable approximation, as described in Appendix A.

r_t of 4 years, also in line with the stylized features of short-term interest rate series. However, although the parameter estimates are plausible, they are imprecisely determined and, more importantly, the model is overwhelmingly rejected at any reasonable significance level. This is consistent with, e.g., the findings of Aït-Sahalia (1996b), Bandi (2002), Durham (2002), and Tauchen (1997). The individual moment diagnostics in Table 5 suggest that one of the main problems is the model's inability to accommodate the tail behavior of our interest rate data. Indeed, the score components associated with the a_{20} and a_{60} terms in the polynomial expansion of the SNP density are highly significant.

4.2.2 Central Tendency

The central-tendency model, given by (6) and (3), extends CIR by adding a stochastic mean factor. Initial experimentation revealed that the correlation coefficient $\rho_{1,3}$ is insignificant and poorly identified by the SNP score moments. Thus, from here on we fix $\rho_{1,3} = 0$. For the same reason, we impose the identical constraint on the $\rho_{1,2}$ coefficient whenever volatility is stochastic.

The estimation results, reported in Tables 3 and 5, are not supportive of the central tendency model. Even though our point estimates are plausible, the parameters are poorly identified with standard errors that increase relative to CIR. Moreover, while the goodness-of-fit has improved slightly, the model is still overwhelmingly rejected. The reasons are readily identified. The moments associated with the higher order SNP coefficients continue to be highly significant, suggesting that the conditional innovations are no less troublesome than they were for the one-factor model. Hence, while the central-tendency factor may, indeed, improve upon the characterization of the drift coefficient it does not help in terms of accommodating the fat tails in the conditional error distribution.

4.2.3 Non-Affine Stochastic Volatility

The stochastic volatility model extends the CIR in a different direction, as the second factor now induces additional variation in the volatility dynamics. We focus initially on the specification given by (7) and (2) with, for the time being, a constant short-rate mean, i.e., $\mu_t = \mu$.

The estimation results, again in Table 3 and 5, are quite remarkable. The addition of the stochastic volatility factor has vastly improved the descriptive value of the model. It now attains a p-value of about 2%, and the standard errors suggest a considerable improvement in precision as well. The unconditional mean μ is 5.1%, and the estimate for the mean-reversion coefficient κ_1 is indicative of strong serial dependence in the mean dynamics, with an implied first order autoregressive coefficient of $\exp(-\kappa_1/52) = 0.9962$ at the weekly level, and a half-life of shocks to r_t of 3-4 years. Both estimates are in line with the characteristics of our interest rate data. However, our estimate for κ_2 appears at odds with the empirical properties of the short-term rate. The first order autoregressive coefficient for the volatility process implied by our estimate is $\exp(-\kappa_2/52) = 0.8877$ at the weekly level, considerably smaller than the value of the largest inverse root of the polynomial $1-\beta_1 L-\beta_2 L^2$, which is 0.9925 based on our discrete-time EGARCH estimates.

In this case, the EMM t-ratios do not pinpoint any particular source of misspecification—individually, each score component is insignificant (Table 5). Of course, these statistics provide only suggestive diagnostics and the 2% p-value for the overall goodness-of-fit test (Table 3) suggests that, despite the dramatic improvement relative to the CIR and central tendency specifications, the model may not be fully adequate.

4.2.4 Non-Affine Stochastic Volatility and Jumps

Overall, the empirical results from the stochastic volatility specification indicate potential model misspecification. A possible explanation is that, in this model, stochastic volatility serves the dual purpose of generating random shocks, and thus producing large outliers, while also accounting for the strong volatility persistence. Our estimates may suffer from the intrinsic tension between the two roles. On the one hand, because of the pronounced non-normality of the (conditional) short rate innovations, the volatility factor will tend to display an erratic pattern that accommodates outlier observations. On the other hand, the latent factor must also try to replicate the strong persistence in the conditional volatility of the interest rate. Our results suggest that the former feature interferes negatively with the latter function.

These observations motivate the extension of the model to incorporate a jump component, as in (1) and (2) with the constraint, for now, that the short rate mean is constant, i.e., $\mu_t = \mu$. Initial experimentation reveals that the parameter μ_J, associated with the average jump size, is insignificant and poorly identified. Thus, we impose the restriction $\mu_J = 0$ during estimation. Point estimates and EMM t-ratios are again reported in Table 3 and 5, respectively.

The estimates for μ and κ_1 are qualitatively consistent with the empirical features of the data and largely unchanged relative to the corresponding values in the pure stochastic volatility model. However, the presence of jumps has a remarkable impact on the volatility parameters. While the long-run volatility mean, α, estimate is unchanged, that of κ_2 is now considerably smaller and more precise. The first order autoregressive coefficient for the volatility process increases to $\exp(-\kappa_2/52) = 0.9723$ at the weekly level, which is very close to the result from our discrete-time EGARCH results and qualitatively consistent with findings in the extant literature. Interestingly, the estimate for the volatility-of-volatility coefficient η_1 is now also much smaller and more precise than that found in the pure stochastic volatility model. These findings are fully consistent with the hypothesis above: Jumps mitigate the task of volatility in accommodating the tail properties of the short rate distribution.

When jumps are introduced in the model, the volatility-of-volatility coefficient compensates for their presence and becomes smaller. Further, the κ_2 parameter is relieved from accommodating extreme outlier behavior and is estimated at a level consistent with a more persistent volatility process.

Turning to the jump coefficients, our estimates indicate that jumps are significant with an arrival rate of 5-6 jumps per year. Under the $\mu_J = 0$ restriction, the expected value $\mathrm{E}(\mathrm{e}^{Z_t}\text{-}1)$ of a jump, conditional on $dq_t = 1$, is very close to zero. Similarly, the standard deviation of the jump component implied by our estimate is $\sigma(\mathrm{e}^{Z_t}\text{-}1)$=0.0263, indicating that most discontinuities are in the order of $\pm$5-6% of the current interest rate level, corresponding to roughly 30-35 basis points.

4.2.5 Non-Affine Stochastic Volatility, Mean Drift, and Jumps

We finally explore the joint effect of simultaneously allowing stochastic volatility and a general time-varying mean drift specification.

We first consider the specification given by (7) and (2)-(3), with the jump component excluded. Our findings in Table 3 indicate that the model delivers a reasonable statistical fit with the p-value for the goodness-of-fit test achieving the respectable level of 14.08%. However, the interpretation of the volatility persistence coefficient κ_2 and the volatility-of-volatility term η_1 are troublesome for the reasons described earlier. The implied first order autoregressive coefficient is $\exp(-\kappa_2/52) = 0.8456$ at the weekly level, which seems unreasonably low given the apparent persistence in the volatility of interest rate data. Also, η_1 is quite large relative to the estimate from, e.g., the stochastic-volatility jump-diffusion model. Even if the EMM t-ratios do not clearly identify any particular types of misspecification, it still suggests that a jump extension may be successful in alleviating the dual role of volatility in accounting for both volatility persistence and outlier behavior.

Upon estimating the more general system (1)-(3), we find the descriptive value of the model to improve considerably relative to all of the previous specifications. The overall goodness-of-fit statistic now attains an impressive 25% p-value, and the score diagnostics suggest that all structural features of the mean and volatility dynamics are accommodated satisfactorily. And, reassuringly, the presence of jumps affects the volatility coefficient estimates substantially. The degree of persistence in the volatility process, controlled by κ_2, is now close to what we expect given the properties of interest rate data (Table 3). Similarly, the η_1 estimate decreases considerably compared to that from the no-jumps case. Further, both κ_2 and η_1 are much more precisely estimated, and all coefficients are highly significant. Moreover, the jump estimates are close to those found in the stochastic volatility model without the central tendency factor. Jumps occur at a rate of 5-6 per year and, given $dq_t = 1$, their mean is (near) zero. Furthermore, our σ_J estimate implies a standard deviation of $\sigma(\mathrm{e}^{Z_t}-1) = 0.0266$, indicating that most discontinuities are still in the order of $\pm$5-6% of the current interest rate level.

Finally, turning to the drift coefficients, the unconditional mean ϑ is 5.7%, virtually indistinguishable from the 5.6% sample mean. Interestingly, the role of κ_1 in the two-factor stochastic volatility model has been passed on to κ_3 in the three-factor model. The reversion of μ_t to the unconditional mean ϑ is slow, as indicated by the low value of κ_3, implying a first order autoregressive coefficient at the weekly frequency of $\exp(-\kappa_3/52) = 0.9937$, and a half-life of 2-3 years for shocks to μ_t. In contrast, the much stronger mean reversion indicated by κ_1 implies a weekly autoregressive coefficient of $\exp(-\kappa_1/52) = 0.9722$ and a half-life of shocks to r_t of about 6 months. The implication is that the short rate converges quite rapidly towards the time-varying mean μ_t, consistent with the periods of pronounced drift that are identifiable in Figure 1. At the same time, μ_t converges extremely slowly to the overall mean, thus inducing the seemingly borderline stationary mean dynamics of the short rate that are evident from the graphical display.

Our preferred specifications are remarkably different from the one-factor representations of, e.g., Aït-Sahalia (1996b), Conley et al. (1997), Eom (1998), and Stanton (1997). Although our three-factor model may appear elaborate, it retains a high degree of tractability and appears to provide a superior fit. Since the first EMM step consists of a semi-nonparametric approximation to the short rate process, the method can, in principle, detect any form of misspecification, including, of course, the omission of terms from the drift specification such as, e.g., r_t^{-1} and r_t^2, which are incorporated in Aït-Sahalia (1996b). However, because our chosen score generator has a homogeneous SNP density, $K_x = 0$, the mean dynamics of the conditional density are inherited solely from the ARMA(4,1) leading term. On the one hand, this restriction is arguably well motivated since the score generator model is selected on the basis of formal information criteria. It implies, on the other hand, that the omnibus specification test may have low power against a misspecified linear drift.[9] Hence, it may be reasonable to allow for a potentially overparameterized, non-homogeneous score generator in order to retain the ability to detect nonlinearity in the drift specification. Consequently, we estimate the three-factor model by EMM using the identical leading term, but with a Hermite(6,1) specification for the standardized residual with the non-homogeneous terms of order higher than 2 fixed at zero to retain relative parsimony. This does not produce any indications of misspecification in our continuous-time model. Moreover, all the score vector diagnostics associated with the non-homogeneous part of the conditional density (not reported here) remain insignificant. Finally, the associated changes in the point estimates of the continuous-time parameters are generally insignificant. In sum, the additional moments in the (constrained) Hermite(6,1) score generator do not appear to provide much useful information, nor do they suggest an inherent source of misspecification in our three-factor jump-diffusion model.

[9] We are grateful to George Tauchen for suggesting this possibility.

4.2.6 Affine Stochastic Volatility and Jumps

Motivated by the recent development of the affine modeling paradigm, we now investigate the affine version of the stochastic volatility model much in parallel with the account provided above. We initially explore the system (8)-(9) with a constant short rate mean, $\mu_t = \mu$, and no jumps. The results are reported in Tables 4 and 5. With the appropriate adjustments, they are very similar to those for the corresponding non-affine model. Although the estimates for the μ and κ_1 coefficients are plausible, those for κ_2 and η_1 have the same problems as noted before. The first order autoregressive coefficient for the volatility process is $\exp(-\kappa_2/52) = 0.8825$ at the weekly level, which is practically identical to the value for the non-affine model (0.8877). In sum, the estimated volatility process has fatter tails than what we would expect and it reverts too rapidly to its long-term mean. The EMM t-ratios are again not particularly helpful in identifying the sources of misspecification, but the model has an overall p-value of 2%, in line with the one obtained in the non-affine case. This suggests that extensions along the lines pursued in the log-volatility model will also be successful in this context.

Thus, we turn to the system (8)-(9), still with the $\mu_t = \mu$ restriction, but now enriched by the presence of jumps. The estimates for μ and κ_1 are largely unchanged. The unconditional mean is found to be 5.1%, roughly consistent with the 5.6% sample mean. The κ_1 estimate implies a first order autoregressive coefficient $\exp(-\kappa_2/52) = 0.9951$ at the weekly level, and half-life of shocks to r_t of 2-3 years. On the other hand, the estimates of the volatility coefficients change considerably and are now in line with prior expectations. Here, $\kappa_2 = 1.6645$, a sensible value and consistent with the corresponding non-affine estimate (1.4585). Further, the η_1 estimate is roughly one half of what was found in the prior case.

We again restrict the jumps to have a zero mean ($\mu_J = 0$). In the affine case jumps arrive at a slightly slower rate (approximately 3 per year). Their size is controlled by the σ_J term. For $dq_t = 1$, the discontinuities in the short rate typically fall within the $\pm 0.34\%$ range. Notice that jumps now are directly related to changes in the level of the interest rate, instead of representing a percentage change in the rate. With this in mind, the results in the affine case are qualitatively consistent with the prior ones, i.e., when $dq_t = 1$ jumps are typically in a range of $\pm 6\%$ of the average short rate. This is synonymous to relatively frequent but fairly small jumps. They are, however, consistent with jumps of the magnitude associated with Federal Reserve Bank changes in the federal funds rate of a quarter or half percentage point. The presence of jumps not only allows the stochastic volatility dynamics to accommodate the economically important longer-term persistence, but it also improves the quality of the fit considerably. The model p-value is now a very respectable 29% and the score diagnostics do not indicate any particular source of misspecification.

4.2.7 Affine Stochastic Volatility, Mean Drift, and Jumps

We consider now the affine three-factor model (8)-(10), in which the jump component is temporarily switched off. Our findings appear in Tables 4 and 5. The model provides a reasonable statistical fit for our short rate data (the p-value is 17.27%). However, the volatility coefficients suffer from the same problems of interpretation as encountered when estimating the prior non-jump specifications. Namely, the mean reversion coefficient κ_2 and the volatility-of-volatility parameter η_1 both appear unrealistically large.

Again, adding jumps to the model resolves the issue. When we estimate the system (8)-(10) with the only restriction that $\mu_J = 0$ the volatility process regains its natural interpretation. Our estimate for κ_2 is 1.6271, indicative of high volatility persistence and very close to the value in the non-affine model. Further, the η_1 estimate drops to roughly one-half of the value obtained in the no-jump case.

Turning to the stochastic mean coefficients, the unconditional mean ϑ is estimated at 5.3%, statistically indistinguishable from the sample mean. As in the non-affine setting, the role of κ_1 is passed on to κ_3 in the three-factor model. Although less precisely estimated, the magnitude of κ_3 is close to that of the corresponding non-affine model term. This finding is qualitatively consistent with the seemingly borderline stationary mean dynamics of the short rate, evident from Figure 1. In contrast, the much stronger mean reversion indicated by κ_1 implies that the short rate converges rapidly towards the time-varying mean, μ_t, consistent with the periods of pronounced drift identifiable in the display as well.

Jump estimates are similar to those obtained for the two-factor stochastic volatility model. Jumps arrive at a rate of 3-4 per year, and are relatively small in magnitude (typically within the $\pm 6\%$ range of the average rate). Nevertheless, they have a significant impact not only on the volatility estimates, as discussed above, but also on the quality of the model fit. The p-value associated to the goodness-of-fit test statistic increases from 17.27 to 37.79%. Finally, the score t-ratios do not indicate any model misspecification. These results are quite remarkable. The affine model class seems to perform as well as, or even better than, the non-affine version explored earlier. The improvement in overall goodness-of-fit may in part be due to the slightly weaker identification of the structural parameters and associated increase in the imputed variance of the basic noise components. Nonetheless, there is no compelling reason to favor one model class over the other from a statistical perspective.

5 Yield Curve Illustrations

So far, our objective has been to develop a continuous-time model that captures the dynamics of the short-term interest rate, r_t. The preferred specification contains jumps and three factors: the short rate itself, the conditional variance of the short rate, V_t, and a stochastic mean level, μ_t. However, our findings have implications that extend beyond the short rate dynamics. The short rate process, along with assumptions about the market prices of risk (risk premia), determines the entire yield curve. If no additional state variables appear in the specification of the risk premia, as is customarily assumed, movements in the yield curve are determined solely by changes in the three factors r_t, V_t and μ_t.[10] In this section, we provide qualitative illustrations of the yield curve implications of our model for a plausible specification of the risk premia. Specifically, Litterman and Scheinkman (1991) analyze a large number of yields using principal components. They identify three dominant factors, interpreted as changes in level, steepness, and curvature. If a candidate short rate model is to be taken seriously, then it should be able to induce similar changes in the yield curve, strictly by time-variation in the state variables. We stress that this does not constitute a formal specification test, but rather a qualitative screening. A full-fledged analysis would require a comprehensive panel of bond yields and goes well beyond the scope of this paper. Nonetheless, this line of reasoning points to a serious limitation of the two-factor stochastic volatility jump-diffusion model and confirms the need for the third factor, the stochastic mean μ_t process associated with the central tendency models.[11]

In Section 4.2, we found that the empirical properties of the affine three-factor jump-diffusion model were similar to those of its non-affine counterpart. Thus, in this section we take advantage of the analytical tractability of the affine specifications to illustrate the implications of our estimates for the yield curve. In its risk-adjusted form, the model (8)-(10) becomes:

$$dr_t = \{\kappa_1(\mu_t - \mu_J^*\,\lambda^*/\kappa_1 - r_t) - \sqrt{V_t}\,\xi_1(t)\}\,dt + \sqrt{V_t}\,dW_{1,t}^* + Z_t^*\,dq_t^*\,, \tag{12}$$

$$dV_t = \{\kappa_2(\alpha - V_t) - \eta_1\sqrt{V_t}\,\xi_2(t)\}\,dt + \eta_1\sqrt{V_t}\,dW_{2,t}^*\,, \tag{13}$$

$$d\mu_t = \{\kappa_3(\vartheta - \mu_t) - \eta_2\sqrt{\mu_t}\,\xi_3(t)\}\,dt + \eta_2\sqrt{\mu}_t\,dW_{3,t}^*\,, \tag{14}$$

where $\xi_1(t)$-$\xi_3(t)$ are the model risk premia, $dW_{i,t}^* = dW_{i,t} + \xi_i(t)\,dt$, $i = 1, \ldots, 3$, q^* is an independent Poisson process with parameter λ^*, while $Z_t^* \leadsto \mathrm{N}(\mu_J^*, \sigma_J^*)$. The coefficients λ^*, μ_J^* and σ_J^* incorporate an appropriate risk-adjustment compared to their counterparts λ, μ_J and σ_J, previously estimated in Section 4.2 under the physical probability measure.

Following Dai and Singleton (2000) and Duffie, Pan and Singleton (2000), we characterize $\xi_1(t)$-

[10] It is generally straightforward to construct a pure exchange general equilibrium model that supports the chosen specification, see Duffie and Kan (1996).

[11] An alternative approach is to introduce additional state variables that only affect the risk premia dynamics, and not the short rate. There is evidence that this may be a fruitful avenue to pursue, as argued by, e.g., CGJ.

$\xi_3(t)$ by:[12]

$$\xi_1(t) = \sqrt{V_t}\,\xi_1\,, \tag{15}$$

$$\xi_2(t) = \eta_1\sqrt{V_t}\,\xi_2\,, \tag{16}$$

$$\xi_3(t) = \eta_2\sqrt{\mu_t}\,\xi_3\,. \tag{17}$$

In the remainder of this section, we fix the coefficients in (12)-(14) that are not affected by model risk adjustments at the EMM estimates in Table 4, while we calibrate the risk premia in (15)-(17) and the risk-adjusted coefficients λ^*, μ_J^* and σ_J^* to match qualitatively the stylized empirical evidence for the term structure of interest rates.

A natural benchmark for the risk premia is the local expectations hypothesis (EH), i.e., $\xi_i(t) = 0$ for all i. Unfortunately, the local EH implies that the yield curve, on average, is downward sloping (inverted), due to the "convexity bias;" see, e.g., Black (1995). In contrast, inspection of historical U.S. interest rates reveals that the yield curve tends to be upward sloping and quite steep at the short end (0-5 years), while it is relatively flat for maturities in excess of 5 years. Our specification of the market prices of risk is calibrated to reflect these properties. Specifically, we set $\xi_1 = -250$, $\xi_2 = 0$, and $\xi_3 = -55$. Since the market prices of risk are negative, the effect is to increase the drift in the short rate which offsets the downward convexity bias.[13] Note further that the risk adjustment in the process for r_t is increasing in the (stochastic) volatility, $\sqrt{V_t}$. This feature is important for our interpretation of the effect of volatility changes. The market price of risk for V_t is set to zero, mainly because its effect on the term structure is minimal. Our yield curves are computed from continuous compounded yields, so the yields are given by $R(r_t, V_t, \mu_t, t, T) = -\log B(r_t, V_t, \mu_t, t, T)/(T-t)$, where $B(r_t, V_t, \mu_t, t, T)$ is the time-t price of a bond with maturity T, which is computed using the semi closed-form formula reproduced in Appendix B. The yield curve depends on the current state vector, i.e., the vector of state variables (r_t, V_t, μ_t). The following analysis is predicated on a baseline case for the state vector from which one or more state variables are perturbed. We summarize the investigation in three scenarios, each demonstrating a "typical" yield curve shift.

We first consider the separate effect of changes in r_t. In Figure 2, the current benchmark is $(r_t, V_t, \mu_t) = (0.08, 0.007^2, 0.08)$, which generates an upward sloping yield curve. This is a common occurrence as the unconditional mean of the short rate is well below the risk neutral mean. The shocks in Figure 2 are changes in r_t of ± 0.01, or 100 basis points (bps). Maturities beyond 5 years

[12] The model risk-premia could be extended using the essentially affine specification advocated by Duffee (2002) and used by Brandt and Chapman (2002), or the even more general representation studied in Cheridito, Filipović and Kimmel (2003).

[13] The widely different magnitudes of the risk premia compensates for differences in scaling. Their relative impact in terms of generating an upwardly sloping yield curve (on average) is, in fact, of the same order of magnitude.

are barely affected by this change, so it is natural to refer to such shifts as changes in *steepness*. To rationalize this finding, note that the mean reversion parameter for r_t (towards the stochastic level μ_t) is $\kappa_1 = 1.7887$, corresponding to a half life of only 4-5 months. Thus, any shock to r_t wears off quickly. It is also worth noting the rich set of possible shapes for the yield curve. For example, in Figure 2, the yield curve for $r_t = 0.09$ has an inverse hump. This particular shape cannot be accommodated by the one-factor CIR model.

Litterman and Scheinkman (1991) find that the single most important change in the yield curve is the parallel shift, whereby yields at all maturities change by a similar amount. They refer to this as the level factor. To capture this feature we use a simultaneous change in r_t and μ_t. As a result, Figure 3 portrays an overall change in interest rate levels, which is qualitatively consistent with the evidence documented in Litterman and Scheinkman (1991).[14]

To summarize, different changes in r_t and μ_t induce shifts in the steepness and the level of the yield curve. In general, the shape of the yield curve reflects the expected movement in the short-term interest rates: r_t converges rather rapidly towards μ_t, which itself displays mean reversion. The influence of V_t on the yield curve is, arguably, less transparent. To develop some intuition, we first consider the Merton (1973) model with constant interest volatility, i.e., the one-factor arithmetic Brownian motion, $dr_t = \mu\, dt + \sigma\, dW_t$. The market price of risk is a constant, ξ, and the yield curve is given by

$$R(t,T) = r_t + \frac{1}{2}(\mu - \xi\,\sigma)(T-t) - \frac{1}{6}\sigma^2(T-t)^2. \tag{18}$$

If $\xi < 0$, an increase in σ causes short-term yields to rise while long term yields drop. To understand this result, note that a rise in volatility has two different effects on the yield curve: the risk neutral drift, $\mu - \xi\,\sigma$, becomes larger, which enhances yields. We call this the "risk premium" effect. However, at the same time the convexity bias rises, and this lowers yields. For short-term maturities the risk premium effect dominates, but for longer-term maturities the convexity bias is stronger.[15]

In order to investigate whether this intuition carries over to the stochastic volatility setting of our three-factor model, Figure 4 displays the yield curve with $(r_t, V_t, \mu_t) = (0.088, 0.007^2, 0.091)$ as the

[14] What we document is not a truly parallel shift as the 30-year rate varies only by a fraction of the 100 bps change at the short end of the curve. This, however, is a consequence of mean reversion in μ_t. Also, our findings are related to the result that, under weak restrictions on the economy, the yield on the long (asymptotic) zero-coupon bond cannot fall, see Dybvig et al. (1996). Reconciling the behavior of long-term interest rates with such theoretical implications is outside the scope of this paper.

[15] This result is not an artifact of the Gaussian model used here. Litterman et al. (1991) reach the identical conclusion for a short rate model which rules out negative rates, and where the volatility factor is related to the volatility of a long rate rather than the short rate. The relation between the short rate volatility and the yield curve is also studied in Christiansen and Lund (2002) and further investigated in Mele (2002).

baseline case. We consider changes in V_t of $\pm\, 0.003^2$. The shift in the yield curve is somewhat different from the constant volatility case, since yields at all (non-zero) maturities rise following an increase in V_t. Likewise, a decrease in V_t lowers all yields. This discrepancy is driven by the mean reversion in volatility. The risk premium effect implies that yields for medium-term maturities increase for positive shocks to V_t but—due to the mean reversion—the shock to V_t wears off before the convexity effect starts to dominate. Overall, changes in V_t only materially affect medium-term maturities (1 to 10 years), leaving the short rate and the 30-year yield unaffected. In fact, if the short rate risk is not priced, i.e. $\xi_1(t) = 0$, even large changes in V_t leave the yield curve virtually unaffected (with a slight decrease) which further bolsters the claim that the impact in Figure 4 is almost entirely due to a risk-premium effect. Since the yield curve is tilted mainly at intermediate maturities, we may associate changes in V_t with the third factor identified by Litterman and Scheinkman (1991), the *curvature* of the yield curve. Furthermore, Litterman et al. (1991) argue that a large part of the short rate volatility is explained by the difference between the 3 year yield and a weighted average of the 1 month and 10 year yields. The yield curves changes in Figure 4 are consistent with this claim since increased volatility elevates the 3 year rate relative to the 1 month and 10 year rates.[16]

The two-factor stochastic volatility model cannot accommodate the same range of dynamic yield curve features. To illustrate this, we recalibrate the market prices of risk, and attempt to replicate the typical yield curve shifts through perturbations to the factors r_t and V_t. The main difference, relative to the three-factor model, is the interpretation of changes in r_t. This is captured in Figure 5. Contrary to Figure 2, changes in r_t now have a significant impact on long-run maturities, and the overall picture is, in fact, much closer to Figure 3, which represents a change in interest rate levels (generated by a joint change in r_t and μ_t). Hence, the two-factor stochastic volatility model is unable to generate changes in the yield curve that are limited to the short-run maturities. The source of the problem is the overly simplistic mean dynamics. Specifically, the drift of the short rate is governed by a single mean reversion parameter, κ_1, which is almost zero. Thus, the model exclusively captures the strong persistence in the short rate. This precludes short-lived shocks to r_t, which are instrumental in generating the steepness changes, thus ruling out this empirically important feature of the yield curve dynamics.

In sum, our qualitative term structure illustrations suggest important roles for the type of stochastic volatility and mean factors, we have estimated from the short rate series, in the analysis of the full term structure. Obviously, these illustrations only serve as a preliminary screening of the properties

[16]Although volatility shocks may affect the inflection of the yield curve through the mechanisms detailed above, other sources of curvature may be required to capture the cross-sectional variation in bond yields. As mentioned, for example, CGJ advocate an unspanned stochastic volatility model and conjecture that a fourth latent factor is needed to generate the curvature effect.

of our model. We do, nonetheless, find it very telling that features such as complex mean drift, strong volatility persistence and sudden jumps are identifiable not only through a study of multiple yields but in fact also are necessary to explain the dynamics of short-term interest rates.

6 Conclusions

The objective of this paper is to identify a class of models that captures the salient features of the short-term interest rate and is sufficiently tractable to form the basis for asset pricing applications. To this end, we consider continuous-time specifications which lie within and outside the affine class. We extend classical specifications to a multi-factor setting, in which the latent variables may be readily interpreted as the conditional mean and volatility of the interest rate. Further, we enrich our models by incorporating a jump component. We conduct estimation via EMM using weekly U.S. 3-month T-Bill rates. We exploit the estimation procedure to generate powerful EMM tests and diagnostics which help us converge towards a couple of specifications that fit the short rate data satisfactorily. Along the way, we identify the features of the interest rate dynamics that account for the inadequate performance of widely used models nested within our general representations. Finally, we illustrate the qualitative implications of our estimated representations for the term-structure of interest rates.

Our analysis leads us from a simple one-factor model to three-factor specifications featuring stochastic volatility, mean drift and jumps. The inclusion of the stochastic volatility factor is critical in providing a good fit, whereas the stochastic mean factor offers a more minor, but still significant, improvement. Specifically, it is important to reproduce a relatively fast mean-reverting behavior of the short rate around a highly persistent time-varying central tendency process. Economically, the mean drift may be indicative of slowly evolving inflationary expectations, time-variation in the required real interest rate, or both. Finally, jumps are critical to the quality of the fit by directly accommodating outliers and thus relieving the stochastic volatility factor from this task. All our three-factor jump-diffusion models pass powerful specification tests, with the affine representation performing on par with the very best models. This result lends support to the affine class of jump-diffusion models, providing a convenient setting for asset pricing applications. Finally, we find that qualitative evidence gleaned from the term structure of interest rates is consistent with the inclusion of stochastic volatility and mean factors of the type that we have identified from the short-term interest rate series alone.

Arguably, our main contribution is to demonstrate that a standard continuous-time jump diffusion model can accommodate all main features of the U.S. short-term interest rate series. We accomplish this by exploiting an estimation technique that is sufficiently powerful to allow for joint inference regarding multiple latent factors and jump components. In contrast, the recent literature tends to conclude that the short rate process is incompatible with such models.[17] Of course, this does not imply that our reduced form modeling approach is *the* correct one—there is surely no such thing as a correct model—but it does point towards a role for this more traditional framework as a useful basis for many asset pricing applications.

Appendix A: Numerical Implementation of EMM

The algorithm used to simulate a sample $\{r_t(\psi), x_t(\psi)\}_{t=1}^{\mathcal{T}(N)}$ of short-term rates r_t and lagged observations x_t from the continuous-time jump-diffusion models (1)-(3) and (8)-(10) resembles closely that illustrated in Andersen, Benzoni, and Lund (2002) and Andersen and Lund (1997).

The short rate as well as the model's state variables is simulated using the Euler scheme—see, for example, Kloeden and Platen (1992). In our simulations, we set $dt = 1/52$ and generate weekly observations expressed in decimal form on a yearly basis. As the SNP model was fit on weekly short rate data expressed in percentage form on a yearly basis, our simulated rates are multiplied by a factor of 100 before being used in the EMM score generator.

Simulations from the stochastic volatility and mean drift terms are not problematic, thus we focus exclusively on the jump component. Poisson jumps are first approximated with a Binomial distribution; i.e., we replace dq_t with a random variable Y such that $\text{Prob}\{Y = 1\} = \lambda(t)\,dt$ and $\text{Prob}\{Y = 0\} = (1 - \lambda(t)\,dt)$. For this purpose, we generate a random variable U Uniform(0,1) and we smooth the discontinuity of Y over an interval centered around $1 - \lambda(t)$:

$$Y = \begin{cases} 0 & \text{if } 0 \leq U < 1 - \lambda(t)\,dt - h/2, \\ g(X) & \text{if } 1 - \lambda(t)\,dt - h/2 \leq U < 1 - \lambda(t)\,dt + h/2, \\ 1 & \text{if } 1 - \lambda(t)\,dt + h/2 \leq U \leq 1, \end{cases}$$

where $X = U - (1 - \lambda(t)\,dt - h/2)$ and $g(X) = -2/h^3 X^3 + 3/h^2 X^2$ for $0 \leq X \leq h$.

Notice that g is a C^∞ function and that it becomes steeper as the interpolation interval length h goes to zero. In our application, we fine-tune h by choosing the smallest possible interval size that eliminates the numerical problems in the EMM criterion function. This yields an accurate approximation to the jumps in the simulated short rate sequence.

Convergence conditions for the Euler approximations in a jump-diffusion setting are discussed in, for example, Kloeden and Platen (1989) and Protter and Talay (1997). These conditions are not explicitly verified for our specific approximation algorithm. As is often the case with high-level assumptions, such

[17] This is one reason that scholars have explored alternative short rate and term structure models incorporating regime shifts. This extensive literature includes, e.g., Hamilton (1988), Cai (1994), Gray (1996), Ang and Bekaert (2002a, 2002b), Bansal and Zhou (2002), Bansal, Tauchen and Zhou (2003), Dai, Singleton and Yang (2003), and Li and Yu (2003).

verification would be difficult. Nevertheless, it does not appear to constitute a problem for our application as extensive simulations verify that the moments of the simulated process converge.

At any iteration of the minimization, each jump $\kappa(t)$ is generated, in the event $dq_t = 1$, using the identical seed. Also, we obtain variance reduction through the use of antithetic variates in the simulation; see, for example, Geweke (1996) for a discussion of this technique.

As a final remark, it should be noted that short rates that are generated through the algorithm explained above may occasionally take on either negative or large values during the course of the EMM optimization. In order to avoid program crashes, we implement the following strategy. Given a simulated rate r_t, we compute the censored/smoothed rate $\hat{r}_t$ according to:

$$\hat{r}_t = \begin{cases} C_1 \exp(C_2\, r_t) + C_3 & \text{if } r_t \le r_{min}, \\ r_t & \text{if } r_{min} < r_t < r_{max}, \\ .5(r_t + r_{max} + \log(1 + r_t - r_{max})) & \text{if } r_{max} \le r_t, \end{cases}$$

where $r_{min} = 0.02\%$, $r_{max} = 75\%$, $C_1 = 0.5\, r_{min} \exp(-2)$, $C_2 = 2/r_{min}$, and $C_3 = 0.5\, r_{min}$.[18] Then, when evaluating the EMM criterion function, we use the *lagged* values of $\hat{r}$ to replace the corresponding simulated rates which are needed to compute (a) the level effect term, (b) the conditional mean, and (c) the non-homogeneous terms of the polynomial expansion in the SNP density.

Appendix B: The Affine Bond Pricing Model

Given the risk-adjusted model (12)-(14), the time-t price $B(r_t, V_t, \mu_t, t, T)$ of a bond maturing at T is

$$B(r_t, V_t, \mu_t, t, T) = \exp\left\{\gamma(t) + \beta_1(t)\, r_t + \beta_2(t)\, V_t + \beta_3(t)\, \mu_t\right\},$$

where $\beta_1(t) = (\mathrm{e}^{-\kappa_1(T-t)} - 1)/\kappa_1$, while $\beta_2(t)$-$\beta_3(t)$ and $\gamma(t)$ are the solution to a system of ordinary differential equations:

$$\begin{aligned} \dot{\beta}_2(t) &= \xi_1\beta_1(t) + (\kappa_2 + \eta_1^2\xi_2)\beta_2(t) - 0.5\beta_1(t)^2 - 0.5\eta_1^2\beta_2(t)^2\,, && (19) \\ \dot{\beta}_3(t) &= -\kappa_1\beta_1(t) + (\kappa_3 + \eta_2^2\xi_3)\beta_3(t) - 0.5\eta_2^2\beta_3(t)^2\,, && (20) \\ \dot{\gamma}(t) &= \mu_J\lambda\beta_1(t) - \kappa_2\alpha\beta_2(t) - \kappa_3\vartheta\beta_3(t) - \lambda g(\beta_1(t)) + \lambda\,, && (21) \end{aligned}$$

with $g(x) = \mathrm{e}^{x\mu_J + 0.5x^2\sigma_J^2}$, $\beta_2(T) = 0$, $\beta_3(T) = 0$, and $\gamma(T) = 0$. The solution to (19)-(21) can be computed in closed-form; see, e.g., Chen (1996). Nevertheless, the analytical expressions for $\beta_2(t)$, $\beta_3(t)$, and $\gamma(t)$ are quite cumbersome, thus it is more practical to solve the same system numerically using standard finite-differencing methods.

[18] We borrowed the spline transformation which is used in the $r_t \ge r_{max}$ case from Gallant and Tauchen (2002).

References

Ahn, Dong-Hyun, Robert F. Dittmar, and A. Ronald Gallant, 2002, Quadratic Term Structure Models: Theory and Evidence, *Review of Financial Studies* 15, 243-288.

Aït-Sahalia, Yacine, 1996a, Nonparametric Pricing of Interest Rate Derivative Securities, *Econometrica* 64, 527-560.

Aït-Sahalia, Yacine, 1996b, Testing Continuous-Time Models of the Spot Interest Rate, *Review of Financial Studies* 9, 385-426.

Aït-Sahalia, Yacine, 2002a, Maximum-Likelihood Estimation of Discretely-Sampled Diffusions: A Closed-Form Approximation Approach, *Econometrica* 70, 223-262

Aït-Sahalia, Yacine, 2002b, Closed-Form Likelihood Expansions for Multivariate Diffusions, Working Paper, Princeton University.

Aït-Sahalia, Yacine, and Robert L. Kimmel, 2002, Estimating Affine Multifactor Term Structure Models Using Closed-Form Likelihood Expansions, Working Paper, Princeton University.

Andersen, Torben G., Luca Benzoni and Jesper Lund, 2002, An Empirical Investigation of Continuous-Time Equity Return Models, *Journal of Finance* 57, 1239-1284.

Andersen, Torben G., and Jesper Lund, 1997, Estimating Continuous-Time Stochastic Volatility Models of the Short-Term Interest Rate, *Journal of Econometrics* 77, 343-377.

Andersen, Torben G., Hyung-Jin Chung, and Bent E. Sørensen, 1999, Efficient Method of Moments Estimation of a Stochastic Volatility Model: A Monte Carlo Study, *Journal of Econometrics* 91, 61-87.

Ang, Andrew, and Geert Bekaert, 2002a, Short Rate Nonlinearities and Regime Switches, *Journal of Economic Dynamics and Control* 26, 1243-1274.

Ang, Andrew, and Geert Bekaert, 2002b, Regime Switches in Interest Rates, *Journal of Business and Economic Statistics* 20, 163-182.

Balduzzi, Pierluigi, Sanjiv R. Das, Silverio Foresi, and Rangarajan K. Sundaram, 1996, A Simple Approach to Three-Factor Models of Interest Rates, *Journal of Fixed Income* 6, 43-53

Balduzzi, Pierluigi, Sanjiv R. Das and Silverio Foresi, 1998, The Central Tendency: A Second Factor in Bond Yields, *Review of Economics and Statistics* 80, 60-72.

Bandi, Federico, 2002, Short-Term Interest Rate Dynamics: A Spatial Approach, *Journal of Financial Economics* 65, 73-110.

Bansal, Ravi, and Hao Zhou, 2002, Term Structure of Interest Rates with Regime Shifts, *Journal of Finance* 57, 1997-2043.

Bansal, Ravi, George Tauchen, and Hao Zhou, 2003, Regime-Shifts, Risk Premiums in the Term Structure, And the Business Cycle, Working Paper, Duke University.

Bates, David, 2003, Maximum Likelihood Estimation of Latent Affine Processes, Working Paper, University of Iowa.

Beaglehole, David R. and Mark S. Tenney, 1991, General Solutions of Some Interest Rate-Contingent Claim Pricing Equations, *Journal of Fixed Income* 1, 69-83.

Black, Fisher, 1995, Interest Rates as Options, *Journal of Finance* 50, 1371-1376.

Brandt, Michael, and David Chapman, 2002, Comparing Multifactor Models of the Term Structure, Working Paper, University of Pennsylvania and University of Texas at Austin.

Brennan, Michael J. and Eduardo S. Schwartz, 1979, A Continuous Time Approach to the Pricing of Bonds, *Journal of Banking and Finance* 3, 133-155.

Cai, Jun, 1994, A Markov Model of Switching-Regime ARCH, *Journal of Business and Economic Statistics* 12, 309-316.

Chacko, George and Sanjiv Das, 2002, Pricing Interest Rate Derivatives: A General Approach, *Review of Financial Studies* 15, 195-241.

Chan, K.C., G. Andrew Karolyi, Francis A. Longstaff and Anthony B. Sanders, 1992, An Empirical Comparison of Alternative Models of the Short-Term Interest Rate, *Journal of Finance* 47, 1209-1227.

Chapman, David A., John B. Long, Jr., and Neil D. Pearson, 1999, Using Proxies for the Short Rate: When Are Three Months Like an Instant? *Review of Financial Studies* 12, 763-806.

Chapman, David A., and Neil D. Pearson, 2000, Is the Short Rate Drift Actually Nonlinear?, *Journal of Finance* 55, 355-388.

Chapman, David A., and Neil D. Pearson, 2001, Recent Advances in Estimating Term-Structure Models, *Financial Analysts Journal* 57, 77-95.

Chen, L., 1996, *Interest Rate Dynamics, Derivatives Pricing, and Risk Managements*, Lecture Notes in Economics and Mathematical Systems, Vol. 435 (Springer Verlag, Berlin).

Cheridito, Patrick, Damir Filipović and Robert L. Kimmel, 2003, Market Price of Risk Specifications for Affine Models: Theory and Evidence, Working Paper, Princeton University.

Chernov, Mikhail, A. Ronald Gallant, Eric Ghysels, and George Tauchen, 2002, Alternative Models for Stock Price Dynamics, *Journal of Econometrics*, forthcoming.

Christiansen, Charlotte, and Jesper Lund, 2002, Revisiting the Shape of the Yield Curve: The Effect of Interest Rate Volatility, Working Paper, Aarhus School of Business.

Collin-Dufresne, Pierre, Robert Goldstein, and Christopher S. Jones, 2003, Identification and Estimation of Maximal Affine Term Structure Models: An Application to Stochastic Volatility, Working Paper, Carnegie-Mellon University, Washington University in St. Louis, and University of Southern California.

Conley, Timothy G., Lars P. Hansen, Erzo G. J. Luttmer, and José A. Scheinkman, 1997, Short-Term Interest Rates as Subordinated Diffusions, *Review of Financial Studies* 10, 525-577.

Cox, John C., Jonathan E. Ingersoll, Jr. and Stephen A. Ross, 1985, A Theory of the Term Structure of Interest Rates, *Econometrica* 53, 385-407.

Dai, Qiang, and Kenneth J. Singleton, 2000, Specification Analysis of Affine Term Structure Models, *Journal of Finance* 55, 1943-1978.

Dai, Qiang, and Kenneth J. Singleton, 2003a, Term Structure Modeling in Theory and Reality, *Review of Financial Studies* 16, 631-678.

Dai, Qiang, and Kenneth J. Singleton, 2003b, Fixed Income Pricing, in George M. Constantinides, Milton Harris, and Reneé M. Stulz, eds.: *Handbook of Economics and Finance*, Ch. 20, (North-Holland, Amsterdam).

Dai, Qiang, Kenneth J. Singleton, and Wei Yang, 2003, Regime Shifts in a Dynamic Term Structure Model of U.S. Bond Yields, Working Paper, New York University and Stanford University.

Das, Sanjiv, 2002, The Surprise Element: Jumps in Interest Rates, *Journal of Econometrics* 106, 27-65.

Das, Sanjiv and Silverio Foresi, 1995, Exact Solutions for Bond and Option Prices with Systematic Jump Risk, *Review of Derivative Research* 1, 7-24.

Diebold, Frank X., and Canlin Li, 2002, Forecasting the Term Structure of Government Bond Yields, Working Paper, University of Pennsylvania and U.C. Riverside.

Duffee, Greg R., 1996, Idiosyncratic Variation of Treasury Bill Yields, *Journal of Finance* 51, 527-551.

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

Duffee, Greg R., and Richard Stanton, 2001, Estimation of Dynamic Term Structure Models, Working Paper, University of California at Berkeley.

Duffie, Darrell, and Kenneth J. Singleton, 1993, Simulated Moments Estimation of Markov Models of Asset Prices,

Econometrica 61, 929-952.

Duffie, Darrell and Rui Kan, 1996, A Yield-Factor Model of Interest Rates, *Mathematical Finance* 6, 379-406.

Duffie, Darrell, Jun Pan, and Kenneth J. Singleton, 2000, Transform Analysis and Asset Pricing for Affine Jump Diffusions, *Econometrica* 68, 1343-1376.

Durham, Garland B., 2002, Likelihood-Based Specification Analysis of Continuous-Time Models of the Short-Term Interest Rate, *Journal of Financial Economics*, forthcoming.

Dybvig, Philip H., Jonathan E. Ingersoll, Jr. and Stephen A. Ross, 1996, Long Forward and Zero-Coupon Rates Can Never Fall, *Journal of Business* 69, 1-25.

Elerian, Ola, Siddhartha Chib, and Neil Shephard, 2001, Likelihood inference for discretely observed non-linear diffusions, *Econometrica* 69, 959-994.

Eom, Young H., 1998, An Efficient GMM Estimation of Continuous-Time Asset Dynamics: Implications for the Term Structure of Interest Rates, Working Paper, Yonsei University.

Eraker, Bjørn, 2001, MCMC Analysis of Diffusion Models with Applications to Finance, *Journal of Econometrics* 19, 177-191.

Fong, Gifford H. and Oldrich A. Vasicek, 1991, Fixed Income Volatility Management, *Journal of Portfolio Management* 17, 41-46.

Gallant, A. Ronald, and Jonathan R. Long, 1997, Estimating Stochastic Differential Equations Efficiently by Minimum Chi-Square, *Biometrika* 84, 125-142.

Gallant, A. Ronald, and Douglas W. Nychka, 1987, Semi-Nonparametric Maximum Likelihood Estimation, *Econometrica* 55, 363-390.

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

Gallant, A. Ronald, and George Tauchen, 1997, Estimation of Continuous Time Models for Stock Returns and Interest Rates, *Macroeconomic Dynamics* 1, 135-168.

Gallant, A. Ronald, and George Tauchen, 2002, SNP: A Program for Nonparametric Time Series Analysis, Working Paper, Duke University.

Geweke, John, 1996, Monte Carlo Simulation and Numerical Integration, in Amman, Hans M., David A. Kendrick and John Rust, eds.:

Handbooks in Economics, vol. 13, *Handbook of Computational Economics*, vol. 1, (North-Holland, Amsterdam).

Goldstein, Robert, 2000, The Term Structure of Interest Rates as a Random Field, *Review of Financial Studies* 13, 365-384.

Gray, Stephen F., 1996, Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process, *Journal of Financial Economics* 42, 27-62.

Hamilton, James D., 1988, Rational-Expectations Econometric Analysis of Changes in Regime: An Investigation of the Term Structure of Interest Rates, *Journal of Economic Dynamics and Control* 12, 385-423.

Hamilton, James D., 1996, The Daily Market for Federal Funds, *Journal of Political Economy* 104, 26-56.

Hansen, Lars P. and José A. Scheinkman, 1995, Back to the Future: Generating Moment Implication for Continuous-Time Markov Processes, *Econometrica* 63, 767-804.

Jacobs, R. L. and Robert A. Jones, 1986, A Two Factor Latent Variable Model of the Term Structure of Interest Rates, Working Paper, Simon Fraser University.

Jegadeesh, Narasimhan and George S. Pennachi, 1996, The Behavior of Interest Rates Implied by the Term Structure of Eurodollar Futures, *Journal of Money, Credit and Banking* 28, 426-451.

Jiang, George J., 1998, Nonparametric Modelling of U.S. Interest Rate Term Structure Dynamics and Implications on the Prices of Derivative Securities, *Journal of Financial and Quantitative Analysis* 33, 465-497.

Jiang, George J., and John L. Knight, 1997, A Nonparametric Approach to the Estimation of Diffusion Processes, with an Application to a Short-Term Interest Rate Model, *Econometric Theory* 13, 615-645.

Johannes, Michael, 2004, The Statistical and Economic Role of Jumps in Continuous-Time Interest Rate Models, *Journal of Finance* 59, 227-260.

Jones, Christopher S., 2003, Nonlinear Mean Reversion in the Short-Term Interest Rate, *Review of Financial Studies* 16, 793-843.

Kearns, Philip, 1992, Pricing Interest Rate Derivative Securities when Volatility is Stochastic, Working Paper, Rochester University.

Kimmel, Robert L., 2002, Modeling the Term Structure of Interest Rates: A New Approach *Journal of Financial Economics*, forthcoming.

Kloeden, Peter E., and Eckhard Platen, 1989, A Survey of Numerical Methods for Stochastic Differential Equations, *Stochastic Hydrology and Hydraulics* 3, 155-178.

Kloeden, Peter E., and Eckhard Platen, 1992, *Numerical Solutions of Stochastic Differential Equations* (Springer-Verlag, Berlin).

Li, Haitao, and Yuewu Yu, 2003, Short Rate Dynamics and Regime Shifts, Working Paper, Cornell University.

Li, Minqiang, Neil D. Pearson, and Allen M. Poteshman, 2002, Conditional Estimation of Diffusion Processes, Working Paper, University of Illinois at Urbana-Champaign.

Litterman, Robert and Jose A. Scheinkman, 1991,

Common Factors Affecting Bond Returns, *Journal of Fixed Income* 1, June, 54-61.

Litterman, Robert, Jose A. Scheinkman and Laurence Weiss, 1991, Volatility and the Yield Curve, *Journal of Fixed Income* 1, June, 49-53.

Lo, Andrew W., 1988, Maximum likelihood estimation of generalized Itô processes with discretely sampled data, *Econometric Theory* 4, 231-247.

Longstaff, Francis A. and Eduardo S. Schwartz, 1992, Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model, *Journal of Finance* 47, 1259-1282.

Mele, Antonio, 2002, Fundamental Properties of Bond Prices in Models of the Short-Term Rate, *Review of Financial Studies*, forthcoming.

Merton, Robert C., 1973, Theory of Rational Option Pricing, *Bell Journal of Economics and Management Science* 4, 141-183.

Piazzesi, Monika, 2003a, Bond Yields and the Federal Reserve, *Journal of Political Economy*, forthcoming.

Piazzesi, Monika, 2003b, Affine Term Structure Models, Manuscript prepared for the *Handbook of Financial Econometrics*, University of California at Los Angeles.

Pritsker, Matthew, 1998, Nonparametric Density Estimation and Tests of Continuous Time Interest Rate Models, *Review of Financial Studies* 11, 449-487.

Protter, Philip, and Denis Talay, 1997, The Euler scheme for Lévy driven stochastic differential equations, *Annals of Probability* 25, 393-423.

Santa-Clara, Pedro, and Didier Sornette, 2001, The Dynamics of the Forward Interest Rate Curve with Stochastic String Shocks, *Review of Financial Studies* 14, 149-185.

Schaumburg, Ernst, 2002, Maximum Likelihood Estimation of SDEs Driven by Levy Type Processes, Working Paper, Northwestern University.

Sinkey, Joseph F., Jr., 1992, *Commercial Bank Financial Management in the Financial-Services Industry* (MacMillan, New York).

Stanton Richard H., 1997, A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk, *Journal of Finance* 52, 1973-2002.

Tauchen, George, 1997, New Minimum Chi-Square Methods in Empirical Finance, in David M. Kreps and Kenneth F. Wallis, eds.: *Advances in Economic Theory and Econometrics*, Seventh World Congress (Cambridge University Press, New York).

Thompson, Samuel, 2004, Identifying Term Structure Volatility from the LIBOR-Swap Curve, Working Paper, Harvard University.

Torous, Walter N. and Clifford A. Ball, 1995, Regime Shifts in Short-Term Riskless Interest Rates, Working Paper, University of California at Los Angeles and Vanderbilt University.

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

Vetzal, Kenneth R., 1997, Stochastic Volatility, Movements in Short-Term Interest Rates, and Bond Option Values, *Journal of Banking and Finance* 21, 169-196.

Tables and Figures

Table 1: Summary statistics. Data on weekly U.S. 3-month T-Bill Yields from 01/06/1954 to 06/28/2000 (N=2,426 observations). All figures are computed using weekly interest rate data expressed in percentage form on a yearly basis.

Mean	5.5614
Std. Dev.	2.8310
Skewness	1.1765
Kurtosis	4.9796

Autocorrelation of Yields:

Lag	*Autocorr.*	*Lag*	*Autocorr.*	*Lag*	*Autocorr.*
1	0.9960	8	0.9538	35	0.8382
2	0.9909	9	0.9479	40	0.8240
3	0.9855	10	0.9422	45	0.8032
4	0.9796	15	0.9161	50	0.7794
5	0.9732	20	0.8925	55	0.7589
6	0.9664	25	0.8683	60	0.7417
7	0.9599	30	0.8482	65	0.7165

Table 2: SNP Model Estimates. Data on weekly U.S. 3-month T-Bill yields, 01/06/1954 to 06/28/2000, (N=2,426 observations). Parameter estimates are for weekly interest rate data expressed in percentage form on a yearly basis and refer to the following model:

$$f_K(r_t|x_t;\xi) = \left(\nu + (1-\nu)\times\frac{[P_K(z_t,x_t)]^2}{\int_{\mathrm{R}}[P_K(z_t,x_t)]^2\phi(u)du}\right)\frac{\phi(z_t)}{r_{t-1}^{\delta}\sqrt{h_t}}\,,\qquad \nu = 0.01\,,$$

where $\phi(.)$ is the standard normal density,

$$\begin{aligned}
z_t &= \frac{r_t-\mu_t}{r_{t-1}^{\delta}\sqrt{h_t}}\,,\\
\mu_t &= \phi_0 + \sum_{i=1}^{s}\phi_i r_{t-i} + \sum_{i=1}^{u}\zeta_i(r_{t-i}-\mu_{t-i})\,,\\
\ln h_t &= \omega\,(1-\sum_{i=1}^{p}\beta_i) + \sum_{i=1}^{p}\beta_i\ln h_{t-i} + (1+\alpha_1 L + \ldots + \alpha_q L^q)\,[\,\theta_1 z_{t-1} + \theta_2\,(b(z_{t-1})-\sqrt{2/\pi})\,]\,,\\
b(z) &= |z| \ \text{ for } |z|\ge\pi/2K,\quad b(z) = (\pi/2-\cos(Kz))/K \ \text{ for } |z|<\pi/2K\,,\quad K=100\,,\\
P_K(z,x) &= \sum_{i=0}^{K_z} a_i(x)z^i = \sum_{i=0}^{K_z}\left(\sum_{|j|=0}^{K_x} a_{ij}x^j\right)z^i\,,\qquad a_{00}=1\,.
\end{aligned}$$

Parameter	*Estimate*	*(std. error)*
ϕ_0	0.0025	(0.0015)
ϕ_1	1.8607	(0.0974)
ϕ_2	-0.9024	(0.1160)
ϕ_3	0.0843	(0.0539)
ϕ_4	-0.0433	(0.0230)
ζ_1	-0.8319	(0.0937)
ω	-0.5053	(2.3824)
α_1	-0.8153	(0.1320)
β_1	1.6288	(0.2505)
β_2	-0.6315	(0.2484)
θ_1	-0.0828	(0.0340)
θ_2	0.4366	(0.0456)
δ	0.8171	(0.2166)
a_{10}	0.0305	(0.0187)
a_{20}	-0.1510	(0.0210)
a_{30}	-0.0322	(0.0122)
a_{40}	0.0702	(0.0107)
a_{50}	-0.0174	(0.0112)
a_{60}	-0.0996	(0.0115)

Table 3: EMM Estimates of the Continuous-Time Non-Affine (Jump-)Diffusion Models. Estimates are for the sample period 01/06/1954 to 06/28/2000. Standard errors are reported in brackets. Parameter estimates are for weekly interest rate data expressed in decimal form on a yearly basis and refer to the following model:

$$\begin{aligned} dr_t &= \kappa_1(\mu_t - \overline{\kappa}\,\lambda\, r_t\,/\kappa_1 - r_t)\,dt + \sqrt{V_t}\,r_t^{\gamma}\,dW_{1,t} + (\mathrm{e}^{Z_t}-1)r_t\,dq_t\,, \\ d\ln V_t &= \kappa_2(\alpha - \ln V_t\,)\,dt + \eta_1\,dW_{2,t}\,, \\ d\mu_t &= \kappa_3(\vartheta - \mu_t\,)\,dt + \eta_2\sqrt{\mu_t}\,dW_{3,t}\,, \\ Z_t &\rightsquigarrow \mathrm{N}(\mu_J,\sigma_J)\,,\ \ \overline{\kappa} = \mathrm{E}(\mathrm{e}^{Z_t}-1) = \mathrm{e}^{(\mu_J+\sigma_J^2/2)} - 1\,,\ \ \mu_J = 0\,,\ \ \mathrm{Prob}(dq_t = 1) = \lambda\,dt\,. \end{aligned}$$

In all models, $\rho_{1,i} = \mathrm{corr}(dW_{1,t}, dW_{i,t}) = 0$, $i = 2,3$, and $\gamma = 0.5$. In the CIR and the log-normal stochastic volatility (SV_1) models, the drift term is fixed at $\mu_t = \mu$. Similarly, in the CIR and central tendency (CT) models, the volatility term is fixed at $\sqrt{V_t} = \sigma$.

Parameter	*CIR*	*CT*	SV_1	SV_1J	SV_1-SD	SV_1J-SD
μ	0.0510 (0.0040)		0.0513 (0.0046)	0.0508 (0.0050)		
κ_1	0.1701 (0.1090)	1.8403 (5.0245)	0.1966 (0.1221)	0.2338 (0.1252)	1.3511 (0.9599)	1.4312 (0.3464)
σ	0.0249 (0.0027)	0.0255 (0.0028)				
α			-7.1535 (0.2289)	-7.1707 (0.3040)	-6.1453 (0.3386)	-7.0461 (0.2237)
κ_2			6.1956 (1.5021)	1.4585 (0.2184)	8.7248 (1.9792)	1.8506 (0.2640)
η_1			2.6774 (0.2854)	1.5327 (0.2034)	3.3885 (0.4095)	1.7709 (0.1539)
ϑ		0.0527 (0.0046)			0.0674 (0.0121)	0.0574 (0.0070)
κ_3		0.2790 (0.5002)			0.3005 (0.1967)	0.3319 (0.1591)
η_2		0.0362 (0.0349)			0.0880 (0.0315)	0.0560 (0.0105)
σ_J				0.0263 (0.0033)		0.0266 (0.0019)
λ				5.3967 (0.1974)		5.2980 (0.1188)
χ^2 [*d.f.*] (*P-Value*)	58.93 [16] ($< 10^{-5}$)	54.71 [14] ($< 10^{-5}$)	26.42 [14] (0.0229)	18.16 [12] (0.1109)	17.24 [12] (0.1408)	12.53 [10] (0.2510)

Table 4: EMM Estimates of the Continuous-Time Affine (Jump-)Diffusion Models. Estimates are for the sample period 01/06/1954 to 06/28/2000. Standard errors are reported in brackets. Parameter estimates are for weekly interest rate data expressed in decimal form on a yearly basis and refer to the following model:

$$\begin{aligned} dr_t &= \kappa_1(\mu_t - \mu_J\,\lambda\,/\kappa_1 - r_t)\,dt + \sqrt{V_t}\,dW_{1,t} + Z_t\,dq_t\,, \\ dV_t &= \kappa_2(\alpha - V_t\,)\,dt + \eta_1\,\sqrt{V_t}\,dW_{2,t}\,, \\ d\mu_t &= \kappa_3(\vartheta - \mu_t\,)\,dt + \eta_2\,\sqrt{\mu_t}\,dW_{3,t}\,, \\ Z_t \rightsquigarrow &\;\; \mathrm{N}(\mu_J, \sigma_J)\,,\;\; \mu_J = 0\,,\;\; \mathrm{Prob}(dq_t = 1) = \lambda\,dt\,. \end{aligned}$$

In the affine stochastic volatility model (SV_2), the drift term is fixed at $\mu_t = \mu$.

Parameter	SV_2	SV_2J	SV_2-SD	SV_2J-SD
μ	0.0512 (0.0055)	0.0526 (0.0018)		
κ_1	0.2593 (0.1449)	0.2555 (0.0350)	1.6313 (1.3334)	1.7887 (1.2393)
α	0.000054 (0.000016)	0.000059 (0.000014)	0.000069 (0.000022)	0.000052 (0.000019)
κ_2	6.4995 (1.2123)	1.6645 (0.5138)	6.2121 (0.9844)	1.7895 (0.4087)
η_1	0.0195 (0.0033)	0.0109 (0.0018)	0.0218 (0.0037)	0.0110 (0.0026)
ϑ			0.0533 (0.0069)	0.0525 (0.0061)
κ_3			0.2968 (0.2401)	0.2792 (0.2930)
η_2			0.0535 (0.0203)	0.0459 (0.0198)
σ_J		0.0017 (0.0002)		0.0016 (0.0002)
λ		2.9844 (0.2159)		3.2688 (0.1223)
χ^2 [*d.f.*] (*P-Value*)	25.90 [14] (0.0266)	14.19 [12] (0.2889)	16.42 [12] (0.1727)	10.74 [10] (0.3779)

Table 5: T-Ratios of the Average Score Components. Columns correspond to the following models: CIR, central tendency (CT), stochastic volatility without (non-affine, SV_1, and affine, SV_2) or with stochastic drift (SV_1-SD, SV_2-SD), extended for the presence of jumps (J). Estimates are for the sample period 01/06/1954 to 06/28/2000. The score vector components are relative to the parameters of the model:

$$
\begin{aligned}
f_K(r_t|x_t;\xi) &= \left(\nu + (1-\nu)\times \frac{[P_K(z_t,x_t)]^2}{\int_{\mathrm{R}}[P_K(z_t,x_t)]^2\phi(u)du}\right)\frac{\phi(z_t)}{\sqrt{h_t}}, \quad \nu = 0.01, \quad \phi(.) \text{ standard normal density},\\
z_t &= \frac{r_t-\mu_t}{r_{t-1}^{\delta}\sqrt{h_t}}, \; \mu_t = \phi_0 + \sum_{i=1}^{s}\phi_i r_{t-i} + \sum_{i=1}^{u}\zeta_i(r_{t-i}-\mu_{t-i}),\\
\ln h_t &= \omega\,(1-\sum_{i=1}^{p}\beta_i) + \sum_{i=1}^{p}\beta_i \ln h_{t-i} + (1+\alpha_1 L + \ldots + \alpha_q L^q)\,[\,\theta_1 z_{t-1} + \theta_2\,(b(z_{t-1}) - \sqrt{2/\pi})\,],\\
b(z) &= |z| \text{ for } |z| \geq \pi/2K, \quad b(z) = (\pi/2 - \cos(Kz))/K \text{ for } |z| < \pi/2K, \quad K = 100,\\
P_K(z,x) &= \sum_{i=0}^{K_z} a_i(x) z^i = \sum_{i=0}^{K_z}\left(\sum_{|j|=0}^{K_x} a_{ij}x^j\right) z^i, \qquad a_{00} = 1.
\end{aligned}
$$

Parameter	CIR	CT	SV_1	SV_1J	SV_1-SD	SV_1J-SD	SV_2	SV_2J	SV_2-SD	SV_2J-SD
ϕ_0	-0.4905	0.0626	0.2853	-0.1961	0.2504	0.7106	0.5100	0.4862	0.3671	0.2484
ϕ_1	-0.3369	0.1570	0.4943	-0.6309	0.4792	1.1362	0.5606	0.4783	0.9447	0.4651
ϕ_2	-0.3019	0.1675	0.5506	-0.6017	0.4776	1.1317	0.5908	0.4960	0.9405	0.4582
ϕ_3	-0.2760	0.1681	0.5980	-0.5789	0.4730	1.1158	0.6149	0.5105	0.9302	0.4474
ϕ_4	-0.2367	0.1843	0.6640	-0.5481	0.4755	1.1154	0.6485	0.5309	0.9353	0.4507
ζ_1	-1.7698	-1.2750	-1.8568	-2.1727*	0.6521	0.7551	-1.9792*	-1.7926	0.5529	0.6351
ω	-1.0896	-1.0353	-1.3611	-1.4101	1.4449	-0.9282	-0.9818	-0.9993	-0.6333	-1.4644
α_1	-0.1598	-0.2819	0.6136	1.4684	-2.0233*	0.9029	0.3821	1.0098	0.2191	1.4980
β_1	1.1763	1.1018	1.1794	1.2530	-1.5929	0.6899	0.7591	0.8170	0.3768	1.3791
β_2	1.1906	1.1172	1.1801	1.2494	-1.5892	0.6871	0.7592	0.8149	0.3759	1.3759
θ_1	0.4454	0.5962	0.5205	0.5124	0.6991	0.5729	-0.2099	-0.1262	-0.2080	-0.2484
θ_2	-1.8147	-1.9815	1.1725	0.6703	-1.6696	0.1786	0.7920	0.6421	0.4112	0.9361
δ	-0.6175	-0.5333	-0.6934	-0.6990	0.7475	-0.6481	-0.8229	-0.8306	-0.7746	-1.1678
a_{10}	0.8347	1.2693	0.1021	-0.2371	0.1575	0.7047	-0.0281	0.0579	-0.2020	0.0718
a_{20}	-4.0584*	-3.9628*	-1.3350	-1.0280	1.5429	-0.5320	-0.8312	-0.8980	-0.4135	-1.3584
a_{30}	1.0513	0.9302	1.2550	1.3539	1.4072	1.2760	1.4154	0.7821	1.3485	0.9029
a_{40}	-0.7527	-0.9557	0.6836	1.9016	-0.9432	1.2538	0.5861	1.2410	-0.0185	1.5950
a_{50}	1.1585	1.0787	1.2612	0.4621	0.5828	0.7153	1.5941	0.7500	1.6786	0.8497
a_{60}	4.4261*	4.4497*	1.6720	-0.1010	1.2775	0.0084	1.7907	0.1205	2.0996*	0.3004

* Average score components significantly different from 0.

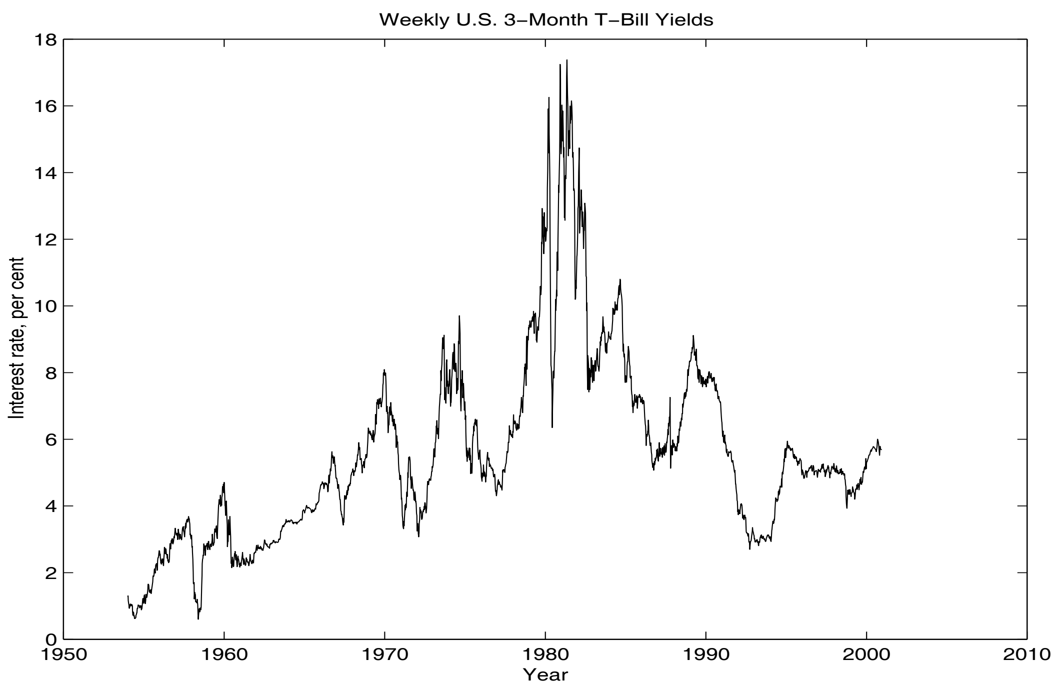

Figure 1: Weekly U.S. 3-month T-Bill Yields, 01/06/1954 to 06/28/2000.

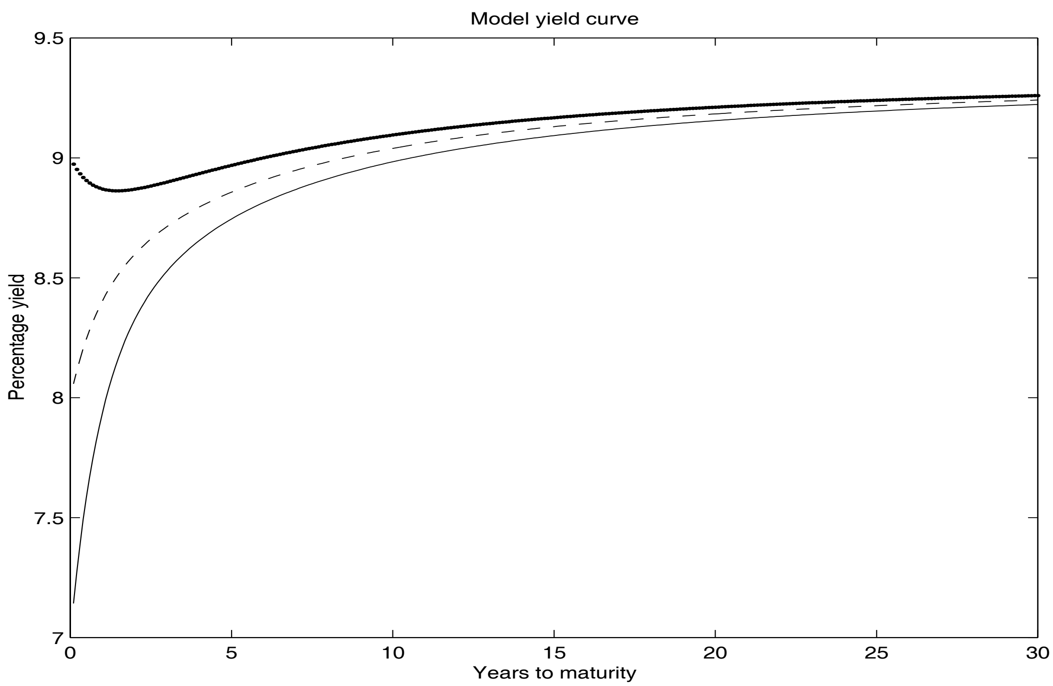

Figure 2: Yield curve from bond prices generated by the affine stochastic volatility and mean drift jump-diffusion model. The model coefficients as well the jump parameters μ_J^*, σ_J^*, and λ^* are fixed at the EMM estimates in Table 4. The risk premia coefficients (ξ_1, ξ_2, ξ_3) are set at $(-250, 0, -55)$. The state variables (V_t, μ_t) take values $(0.007^2, 0.08)$, while r_t is 0.09 $(\cdots)$, 0.08 (- -), and 0.07 (—).

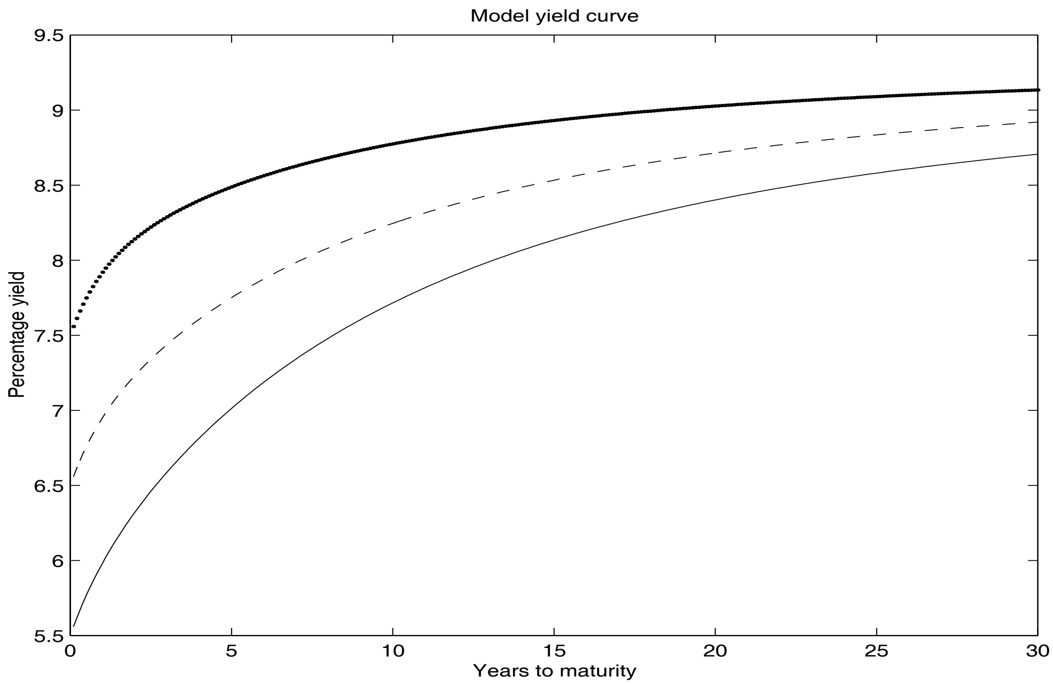

Figure 3: Yield curve from bond prices generated by the affine stochastic volatility and mean drift jump-diffusion model. The model coefficients as well the jump parameters μ_J^*, σ_J^*, and λ^* are fixed at the EMM estimates in Table 4. The risk premia coefficients (ξ_1, ξ_2, ξ_3) are set at $(-250, 0, -55)$. The state variables V_t takes value 0.007^2, while (r_t, μ_t) are (0.075, 0.075) $(\cdots)$, (0.065, 0.065) (- -), and (0.055, 0.055) (—).

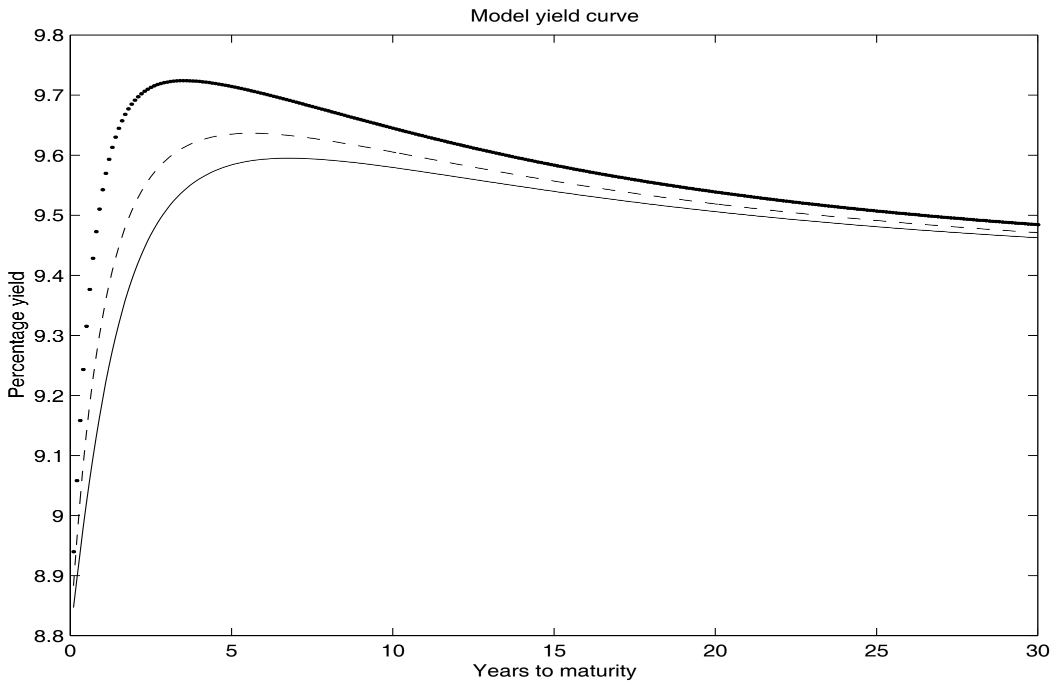

Figure 4: Yield curve from bond prices generated by the affine stochastic volatility and mean drift jump-diffusion model. The model coefficients as well the jump parameters μ_J^*, σ_J^*, and λ^* are fixed at the EMM estimates in Table 4. The risk premia coefficients (ξ_1, ξ_2, ξ_3) are set at $(-250, 0, -55)$. The state variables (r_t, μ_t) take values $(0.088, 0.091)$, while V_t is 0.01^2 $(\cdots)$, 0.007^2 (- -), and 0.004^2 (—).

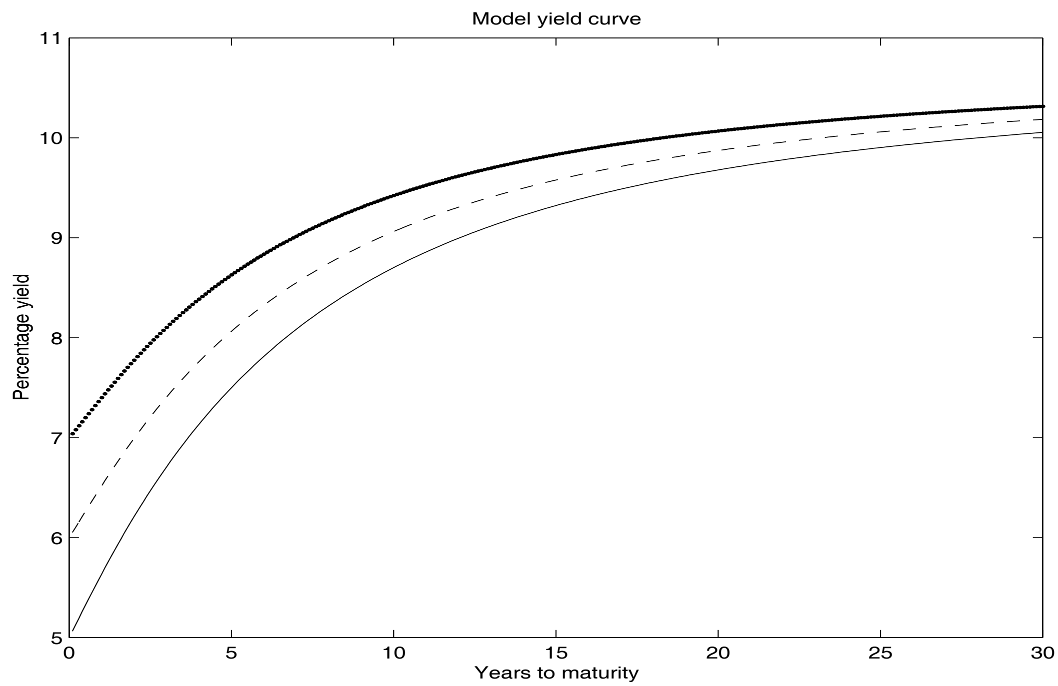

Figure 5: Yield curve from bond prices generated by the affine stochastic volatility jump-diffusion model (constant mean). The model coefficients as well the jump parameters μ_J^*, σ_J^*, and λ^* are fixed at the EMM estimates in Table 4. The risk premia coefficients (ξ_1, ξ_2) are set at $(-250, 0)$. The state variable V_t takes value 0.007^2, while r_t is 0.07 $(\cdots)$, 0.06 (- -), and 0.05 (—).