

Jump-Diffusion Term Structure and Itô Conditional Moment Generator

Hao Zhou¹

Mail Stop 91
Federal Reserve Board
Washington, DC 20551

First Draft: January 1998
This Version: April 2001

¹I would like to thank George Tauchen for his valuable advice. I am also grateful to an associate editor, two anonymous referees, Ravi Bansal, Tim Bollerslev, Sanjiv Das, Peter Christoffersen, Lars Hansen, David Bates, Pete Kyle, Ronald Gallant, Mark Coppejans, and Michael Hemler for their helpful suggestions. Comments from the seminar participants at Duke University, Brown University, University of Virginia, the Society of Nonlinear Dynamics and Econometrics 1999 Annual Meeting, Financial Management Association International 1999 Annual Meeting, the Econometric Society 2000 Annual Meeting, and Duke Conference on Risk Neutral and Objective Probability Distributions in 2000 are greatly appreciated. The views presented here are solely those of the author and do not necessarily represent those of the Federal Reserve Board or its staff. For questions and comments, please contact Hao Zhou, Trading Risk Analysis Section, Division of Research and Statistics, Federal Reserve Board, Washington DC 20551 USA; Phone 1-202-452-3360; Fax 1-202-728-5887; e-mail hao.zhou@frb.gov.

Abstract

This paper implements a Multivariate Weighted Nonlinear Least Square estimator for a class of jump-diffusion interest rate processes (hereafter MWNLS-JD), which also admit closed-form solutions to bond prices under a no-arbitrage argument. The instantaneous interest rate is modeled as a mixture of a square-root diffusion process and a Poisson jump process. One can derive analytically the first four conditional moments, which form the basis of the MWNLS-JD estimator. A diagnostic conditional moment test can also be constructed from the fitted moment conditions. The market prices of diffusion and jump risks are calibrated by minimizing the pricing errors between a model-implied yield curve and a target yield curve. The time series estimation of the short-term interest rate suggests that the jump augmentation is highly significant and that the pure diffusion process is strongly rejected. The cross-sectional evidence indicates that the jump-diffusion yield curves are both more flexible in reducing pricing errors and are more consistent with the Martingale pricing principle.

Keywords: Jump-Diffusion, Term Structure of Interest Rates, Conditional Moment Generator, Multivariate Weighted Nonlinear Least Square, Market Price of Risk.

JEL classification: C51, C52, G12

1 Introduction

In modeling the short-term interest rate, researchers face the challenge of accommodating all relevant features in a single model specification. Those features, include but are not limited to, (1) strong mean persistence (short-term non-stationary), (2) high volatility and volatility persistence, (3) fat-tailed distribution (large skewness and kurtosis), (4) long-run mean reversion (stationary), and (5) level-dependence of changes or volatility in the short rate. The celebrated CIR model (Cox et al., 1985b) and its various extensions, although appealing in their general equilibrium nature and closed-form solution, have difficulty in fitting all these features simultaneously for the US interest rate data (Brown and Dybvig, 1986).¹ Rigorous specification tests tend to reject the square-root model, as they use the historical short-rate data (Aït-Sahalia, 1996b; Conley et al., 1997; Gallant and Tauchen, 1998). Although the CIR-type model has an inherent advantage in fitting features (1), (4), and (5), as indicated in the literature, the model fails to capture feature (2)—the high volatility and volatility persistence.² From a time-series perspective, the underfitting of the volatility parameter is the major cause for rejecting the short-rate dynamics. From a cross-sectional perspective, the pricing error of term structure is large when an understated volatility parameter reduces the flexibility of yield curves.

Consequently, efforts to modify the square-root model largely concentrate on more flexible specifications of the volatility function. It is clear that the CIR model is just one special case of so-called linear CEV (constant elasticity of volatility) specification, where elasticity equals 0.5. Recent comparative studies (Chan et al., 1992; Conley et al., 1997; Tauchen, 1997; Christoffersen and Diebold, 2000) found that an elasticity close to 1-1.5 is more acceptable. Alternatively, one can estimate the volatility nonparametrically (Aït-Sahalia, 1996a; Stanton, 1997; Bandi and Phillips, 2000). The empirical findings of this list suggest that the square-root model fits reasonably well for the medium range of interest rates, although the nonlinearity at both the high and low ends is inconclusive. A pertinent approach is to introduce an unobserved stochastic volatility factor into the diffusion function, which finds

¹The bivariate extensions of CIR specification (Gibbons and Ramaswamy, 1993; Chen and Scott, 1993; Pearson and Sun, 1994) also meet with poor empirical performance. Duffie and Singleton (1997) found favorable evidence for a two-factor CIR model with serially-correlated error structure. Dai and Singleton (2000) estimated more flexible three-factor affine specifications similar to Chen (1996) and Balduzzi et al. (1996), and these models passed the specification test.

²Feature (3), large skewness and kurtosis, as an unconditional measure, is automatically satisfied if all the dynamic features are fitted. However, getting the unconditional moments correct does not guarantee the capturing of all the dynamic properties.

considerable support in empirical studies (Andersen and Lund, 1996, 1997). The stochastic volatility model retains the level dependence of short-rate changes, while greatly enhancing the long memory and high volatility features of the short-rate process. The jump-diffusion approach of interest-rate and no-arbitrage bond pricing is of more recent origin (Das, 1998), and its general equilibrium foundation was formulated by Ahn and Thompson (1988).³ The empirical approach (see Das, 1998, for instance) typically assumes a constant diffusion function as in Vasicek (1977). Under this setting, the unconditional volatility structure is enriched, while the conditional volatility persistence is still not fully captured.⁴

Conceptually, the jump-diffusion approach is more similar to the stochastic volatility strategy, but with better economic interpretation. The financial markets may receive information either through small, gradual perturbations or large, sudden shocks (Merton, 1976). The examples of microeconomic information flow include the temporary imbalances of supply and demand, changes in the economic outlook of a small number of market participants, and the publication of earning reports from several large companies within a week. At the same time, the monetary authority may adjust the discount rate by a quarter percent, OPEC oil agreements may produce a supply shock, or an Asian or Russian financial crises may affect asset demands across the world. These macroeconomic information shocks may completely alter the market perception of economic fundamentals. By their very nature, these shocks arrive only randomly, at certain points of time, and their impacts on market movements are in large, discrete sizes. Such discrete-size information shocks are reflected in the financial markets as data outliers. Many times the continuous sample path distributions (e.g., the CIR model) fail to explain the extraordinary volatility structure. However, the occasional volatility clustering is *the* feature of asset returns, and the asset pricing theory requires that the empirical methods are able to capture and explain both the smooth and the rough periods of the financial markets. A mixed model of a continuous Brownian motion and a discrete Poisson jump may be capable of capturing the real-time evolution of most financial asset prices (see Ingersoll, 1987, page 267-268).

³One can think of the jump-diffusion process as a special case of the stochastic volatility model (Merton, 1976; Heston, 1993). The jump-diffusion specifications considered in this paper have closed-form solutions to both dynamic moments and pricing functions, while the stochastic volatility model in general has to rely on numerical simulation in both estimation and pricing.

⁴Recently there is a growing literature on jump-diffusion interest rate modeling (see Baz and Das, 1996; Chacko and Das, 1999; Johannes, 1999; Piazzesi, 2000, among many others), which ranges from short-rate dynamics to fixed-income derivatives, from market-implied jumps to macroeconomic announcements, and from parametric to nonparametric specifications.

The innovation of this paper is to generate the parametric moment conditions and to construct a computationally efficient estimator that includes a diagnostics test. Maximum Likelihood Estimation (MLE) is available only for a very restricted class of jump-diffusion models (Lo, 1988). Our method differs with the infinitesimal generator of Hansen and Scheinkman (1995) (GMM) in that it fully exploits the conditional information, does not rely on simulations, as do Duffie and Singleton (1993) (SMM), uses model-dependent moments instead of data-dependent moments (Gallant and Tauchen, 1996) (EMM), generalizes to an arbitrary number of moments rather than only to conditional mean and variance (Fisher and Gilles, 1996) (QML), and has faster solutions for both estimation and pricing in comparison with the nonparametric approach (Aït-Sahalia, 1996a) (NP). As shown below, our method reduces a complicated task of solving a stochastic differential equation (SDE) to a simple matrix solution of an ordinary differential equation (ODE) system. The solution becomes a linear least-square problem with nonlinear parameter constraints, or, at most, a multivariate nonlinear least-square problem. The computational burden is reduced to only few minutes.⁵ In the literature, the most closely related method is to identify the stochastic differential equations with an orthogonal series representation (Hansen et al., 1998), which is attributed to the generalized eigenvalue-eigenfunction technique Wong (1964). The distinct feature of our paper is the maximum exploitation of the parametric information that is contained in the drift, diffusion, and jump specifications. The resulting estimation strategy retains a closed-form solution to bond prices, analytical expression for any order of moments, high computational efficiency, and the ability to be easily extended to multivariate scenario.

The body of this paper is organized as follows: Section 2 applies the martingale pricing technique to the jump-diffusion short rate model and derives closed-form solutions to bond prices; Section 3 characterizes the first four conditional moments and constructs a nonlinear least-square estimator that includes diagnostics testing; Section 4 estimates the objective and risk-neutral term structure dynamics to provide empirical support for the jump-diffusion specifications; and Section 5 concludes.

⁵Alternatively, an equivalent spectral method of moments is developed by Chacko and Viceira (1999) and Singleton (2001). However, the selection of spectral moments remains as a difficult problem, whereas in the classical method of moments, a natural choice is the lower-order moments.

2 Jump-Diffusion Term Structure

A standard martingale pricing technique is applied to solve the bond price with the short rate following a jump-diffusion process. In this section, I simply lay out the basic structure and discuss some relevant issues, while technical details are delegated to the references mentioned in footnote (6).

2.1 Specifying the Short-Rate Process

Suppose that the evolution of short interest rates is governed by a square-root jump-diffusion process

$$dr_t = \kappa(\theta - r_t)dt + \sigma\sqrt{r_t}dW_t + JdN(\rho t), \quad (1)$$

where W_t is a standard Brownian motion, $N(\rho t)$ is a Poisson driving process with an intensity function ρ , and J is the jump size with distribution $\Pi(J)$. If there are no jumps ($\rho = 0$), equation (1) reduces to the benchmark square-root model (Cox et al., 1985a). In the following structural development and empirical application, I will focus on the case in which the jump-rate ρ is a constant and the jump-size J is a Uniform $[a, b]$ random variable. I then make extensive comparisons between this situation and the benchmark CIR case. Of course there are other variations. One is the affine jump-rate $\rho_0 + \rho_1 r_t$ with the jump-size J being an independent random variable. Another is the constant jump-rate ρ with the jump-size being an affine function $J_0 - r_t$.⁶ For notational simplicity, I assume that the Brownian motion, Poisson process, and random jump-size are all independent, although certain correlated structures can be accommodated without difficulties.

To guarantee that model (1) is well defined, certain regularity conditions need to be satisfied. The restrictions imposed on the square-root component is offered by Feller (1951)—(a) $\kappa > 0$, $\theta > 0$, and $\sigma^2 \leq 2\kappa\theta$, which ensures that the diffusion is in the domain $(0, \infty)$. Zero is not accessible except as a starting value, and the process never explodes to infinity.⁷ The next assumption serves to exclude the technical arbitrage from tampering with jump information—(b) $\rho \in \mathcal{F}_t^-$ and $\Pi(J) \in \mathcal{F}_t^-$, which says that the jump-intensity and jump-size distribution only depend on the left limit of r_t , to preserve the Markov property of the

⁶These jump-diffusion processes or their close variations have been studied in the literature, either from a no-arbitrage point of view (Baz and Das, 1996; Chacko and Das, 1999; Duffie et al., 2000) or within a general equilibrium framework (Ahn and Thompson, 1988). An earlier version of this paper provides closed-form solutions to bond prices for these jump-diffusion interest rate models. They are available upon request.

⁷Given mean reversion $\kappa > 0$ and positive long run mean $\theta > 0$, the process is also defined even if $\sigma^2 > 2\kappa\theta$ (Cox et al., 1985a). In this case zero is accessible but not absorbing.

interest-rate process. In other words, if one can predict the exact jump timing and jump-size at an instant before a jump occurs, one can make an arbitrarily large profit with certainty. The last assumption ensures that the diffusion and jump components are conforming to one another—(c) $0 \leq \rho < \infty$, $-r_{t-} < J < \infty$ at $N(\rho t) = 1$. This assumption ensures nonnegative jump intensity and a positive short rate.⁸ The second part of this assumption is not satisfied by some empirical results in Section 4, which motivates a “less” parametric jump-size specification only by its moments (see Section 4.3). Under assumptions (a)-(c), the jump-diffusion process (1) is well-defined and can be constructed properly.

2.2 Justification for the Pricing Kernel

The notion of no-arbitrage is equivalent to the existence of a state price density or pricing kernel, such that all the risky factors in the economy are compensated appropriately. In our jump-diffusion model, the nonlinear jump-size uncertainty cannot be exactly hedged by a finite set of bonds. Since the market is incomplete, one can assume that the jump-size risk is not priced in order to invoke a candidate pricing kernel as consistent with no-arbitrage. Alternatively, a preference-based pricing kernel can arise from a general equilibrium setting (Ahn and Thompson, 1988), where all three sources of risks—Brownian motion, random jump timing, and random jump magnitude—are priced nonlinearly, if they are correlated with the fundamental shocks to the economy. Intuitively, jumps to the interest rate and yield curve are all systematic; therefore both jump-rate and jump-size risks should be priced accordingly.

The pricing kernel adopted here will not try to identify separately the risk premia between the jump-rate and the jump-size for following reasons. First, if a linear pricing kernel is used with additive premiums for diffusion, jump-rate, and jump-size risks, neither the no-arbitrage argument nor the equilibrium consideration may support that usage (Merton, 1976). The randomness of jump timing and magnitude affects the short rate multiplicatively, and their marginal effects on optimal wealth are not linearly separable. Also, if the risk premia for jump-rate and jump-size enter the pricing kernel nonlinearly, one can not preserve the closed-form solution to the term structure in a linear format. Furthermore, the recent history of monetary policy suggests that the timing of rate adjustments is largely subject to speculation,

⁸Ahn and Thompson (1988) explicitly solved the term structure for a case where the jump intensity is ρr_t and the jump size is a constant $J < 0$. They further assume that the jump size J is such that a positive interest rate will always result, which is equivalent to assuming that $-r_{t-} < J < 0$ at $N(\rho t) = 1$.

while the possible direction (up or down) and the most likely magnitude of the rate changes (e.g., 25 basis points) are widely agreed upon by the market participants (Piazzesi, 2000). Finally, a constant jump premium may effectively be capable of compensating the joint risk of random jump timing and magnitude, as both factors here are specified with constant variances, and the product of the two evolves more like a single mixture random variable.

These considerations taken into account, I adopt the following pricing kernel:

$$\frac{d\pi(t)}{\pi(t)} = -r_t dt - \frac{\lambda_W}{\sigma} \sqrt{r_t} dW_t - \lambda_J [dN(\rho t) - \rho dt], \quad (2)$$

where λ_W is the diffusion risk premium parameter and λ_J is the jump risk premium parameter. The expected return of the pricing kernel should equal to the risk-free rate $E_t \left[\frac{d\pi(t)}{\pi(t)} \right] = -r(t)dt$. The functional form of $\frac{\lambda_W}{\sigma} \sqrt{r_t}$ is in accordance with Cox et al. (1985a), and the choice of constant parameter λ_J is similar to the constant volatility case (Vasicek, 1977). The choice of the functional form of risk premia is not unique. In fact, any general functional form is allowed for the market price of risk, as long as the arbitrage opportunity is ruled out (Stanton, 1997). Of course, one may not be able to obtain the closed-form affine bond prices. Even within the affine term structure specifications, more flexible choices of risk premia are shown to improve the fitting of bond yields (Duarte, 1999; Duffee, 2000).

2.3 Solving for the Bond Price

Given the short-rate process, the bond return process can be spelled out. The price of a discount bond $P(t) = P(r_t, t, T)$, at time t , with $T-t$ maturity is conjectured to be log-linear $P(r_t, T-t) = A(T-t) \exp\{-B(T-t)r_t\}$, with the boundary condition $P(r_T, T, T) = 1$. Generalized Itô's Lemma (Merton, 1971; Lo, 1988) delivers the bond return process

$$\frac{dP(t)}{P(t)} = \mu_P(t)dt + \sigma_P(t)dW_t + J_P(t)dN(\rho t), \quad (3)$$

where $\mu_P(t) = \frac{P_r \kappa(\theta - r) + \frac{1}{2} P_{rr} \sigma^2 r + P_t}{P}$, $\sigma_P(t) = \frac{P_r \sigma \sqrt{r}}{P}$, and $J_P(t) = \frac{P(r+J, T-t) - P(r, T-t)}{P}$ are the instantaneous drift, diffusion, and jump functions. Define a new process $M(t) = \pi(t)P(t)$, which is a martingale due to the fundamental no-arbitrage condition $P(t) = E_t \left[\frac{\pi(s)}{\pi(t)} P(s) \right]$, for any $s \geq t$ (or, in equilibrium, the first order condition); hence $M(t) = E_t[M(s)]$, for any $s \geq t$. Since $M(t)$ is defined as a martingale, it must have zero drift $\mu_M(t) = 0$, which can be spelled out as

$$\mu_P(t) = r_t + \frac{\lambda_W}{\sigma} \sqrt{r_t} \sigma_P(t) + \rho(\lambda_J - 1)E_J[J_P(t)]. \quad (4)$$

This formula simply says that the instantaneous bond return should be equal to the “risk free” rate plus the diffusion risk premium and the jump risk premium.⁹

Using the conjectured log-linear bond price formula, one thus reaches the fundamental evaluation equation,

$$\frac{1}{2}\sigma^2 r P_{rr} + \kappa(\theta - r)P_r + P_t - rP - \lambda_W r P_r + \rho(1 - \lambda_J)PE_J[e^{-BJ} - 1] = 0. \quad (5)$$

Note that if there are no jumps ($\rho = 0$), the pricing formula will reduce to the CIR model. When the jump risk is not priced ($\lambda_J = 1$), the pricing formula may look like the CIR model, but the diffusion parameters and the bond yields are different given the presence of jumps. A Similar situation occurs in option pricing when the jump risks are diversifiable (Merton, 1976), i.e., the model parameters and derivative prices are altered even though the jump risks are not priced. Finally if the jump size is nonrandom, one may delete the expectation operator; further to maintain an exact CIR-type bond pricing formula, one can either (a) not price the jump-rate risk ($\lambda_J = 1$) or (b) apply a Taylor series approximation to the term $[e^{-BJ} - 1]$ (Ahn and Thompson, 1988). If we restrict ourselves to the short-rate process (1), assuming that the composite jump risk is priced by a constant premium while also maintaining the log-linear pricing function, the solution will be characterized by an Ordinary Differential Equation (ODE) system,

$$B' = 1 - \frac{1}{2}\sigma^2 B^2 - (\lambda_W + \kappa)B, \quad (6)$$

$$\frac{A'}{A} = -\kappa\theta B + \rho(1 - \lambda_J)\left(\exp\left\{\frac{e^{-Bb} - e^{-Ba}}{-B(b-a)}\right\} - 1\right), \quad (7)$$

with boundary conditions $A(0) = 1$ and $B(0) = 0$.

2.4 The Issue of Negative Interest Rates

The issue of negative interest rates is intertwined with the choice of jump-size distribution and with its impact on solving both the expectation $E_J[e^{-BJ} - 1]$ and the pricing ODE system. The reality is, in United States the negative interest rate has not been an issue

⁹In general, calculating the drift $\mu_M(t) = 0$ by Itô's formula involves a joint conditional expectation (conditioning on \mathcal{F}_t^-) of the composite jump term with respect to both the Poisson driving process and the random jump magnitude. Since the jump timing and size are assumed to be independent, the jump rate ρ can be factored out first. Further, since the short-rate process (1) assumes that the jump size follows a Uniform distribution independent of the state variable, the conditional expectation can then be replaced by an unconditional expectation only with respect to $\Pi(J)$.

since World War II, although Japan recently experienced that event occurrence while in deep recession. A Negative nominal interest rate is not an arbitrage problem, if the deflation rate is more negative. Whether the real interest rate could be less than zero is subject to debate.

One advantage of the Poisson-Gaussian jump-diffusion model (Das, 1998) is that the Ornstein-Uhlenbeck process sustains negative interest rates. Consequently the choice of jump-size distribution is very general—almost any random variable with a moment generating function (MGF) is a candidate choice, provided that the MGF does not impose conflicting restrictions in solving the pricing ODE system. Nevertheless, volatility persistence and the level dependence of the short-rate changes, which are fairly stylized features of the interest rates, are not present in this setting. Note that the inclusion of jumps does allow the unconditional volatility to be more flexible, but since the jumps are rare events, the conditional volatility will remain quite restrictive under the Vasicek (1977) diffusion specification.

The disadvantage of the square-root jump-diffusion model adopted in this paper is that the short rate must be positive. Assuming that the jump-timing is independent of the Brownian motion, one has to require the jumps to be non-negative. This in turn implies a very narrow choice of jump-size distributions—not only to be bounded below from zero but also to have a MGF that does not restrict the ODE solution. As a result, only Uniform, Exponential, and Bernoulli distributions—including some very close variations—are qualified jump-size choices.¹⁰ Even with these drawbacks, a square-root diffusion function still exhibits some inherent time-varying volatility persistence and the level-dependence of short-rate changes. Therefore, adding jumps to the CIR model may further enrich the volatility pattern and better capture the short-rate dynamics.

3 Moment Generator and Estimation Strategy

In this section I characterize the conditional moments of a square-root jump-diffusion process in a convenient matrix form, and I implement a Multivariate Weighted Nonlinear Least Square for Jump-Diffusion (MWNLS-JD) estimator. Regularity conditions for consistent estimation and correct inference are laid out, and a conditional moment-based specification

¹⁰In empirical application, both Bernoulli and Exponential distributions are also tested, with very poor parameter estimates, and hence they are discarded. More general classes of models—some are susceptible to the problem of negative interest rates and others are not—are covered by Chacko and Das (1999) and Duffie et al. (2000) in the context of derivative pricing and characteristic function.

test is constructed appropriately.

3.1 A Characterization of Conditional Moment

Now let us focus on the first four conditional moments of the proposed jump-diffusion processes (1) $dr_t = \kappa(\theta - r_t)dt + \sigma\sqrt{r_t}dW_t + JdN(\rho t)$. Let $R_s = [r_s, r_s^2, r_s^3, r_s^4]'$ be the column vector of the first four powers of r_s for some $s \geq t$. An extended version of Itô's formula (Merton, 1971; Lo, 1988) delivers $E_t(R_s) = [E_t(r_s), E_t(r_s^2), E_t(r_s^3), E_t(r_s^4)]'$ in a matrix form:¹¹

$$\begin{bmatrix} E_t(r_s) \\ E_t(r_s^2) \\ E_t(r_s^3) \\ E_t(r_s^4) \end{bmatrix} = \begin{bmatrix} r_t + E_t \int_t^s \{\kappa(\theta - r_u) + \rho E_J[(r_u + J) - r_u]\} du \\ r_t^2 + E_t \int_t^s \{2r_u \kappa(\theta - r_u) + \sigma^2 r_u + \rho E_J[(r_u + J)^2 - r_u^2]\} du \\ r_t^3 + E_t \int_t^s \{3r_u^2 \kappa(\theta - r_u) + 3r_u \sigma^2 r_u + \rho E_J[(r_u + J)^3 - r_u^3]\} du \\ r_t^4 + E_t \int_t^s \{4r_u^3 \kappa(\theta - r_u) + 6r_u^2 \sigma^2 r_u + \rho E_J[(r_u + J)^4 - r_u^4]\} du \end{bmatrix}.$$

So the conditional moment is simply the realization of the four powers of r_t at the initial date plus the expected Riemann integral of the stochastic differential generated by Itô's Lemma. If one observes the continuous time record, the empirical counterparts of these conditional moments can be measured directly by numerical integration. Since the data is only available in discrete samples or since the continuous time record is contaminated by microstructure noises, the main challenge remains to tackle the integration without relying on the actual sample path. For instance, Stanton (1997) applied a stochastic Taylor series approximation to the integral and estimated by a nonparametric kernel regression approach. Hansen et al. (1998) adopted an orthogonal series approximation in a general eigenvalue-eigenfunction framework. Fisher and Gilles (1996) proposed a quasi-maximum likelihood estimator based on closed-form conditional mean and variance functions of the affine diffusion process. The approach developed below is more general in terms of the concerned jump-diffusion processes and more straightforward in terms of extending to any number of moments.

Taking derivatives of $E_t(R_s)$ with respect to the future time s and writing the results in a matrix form, we arrive at a system of ordinary differential equations

$$\frac{dE_t(R_s)}{ds} = A(\beta)E_t(R_s) + g(\beta), \quad (8)$$

¹¹Since the jump-size J is assumed to be independent of the state variable, for any function $f(\cdot)$, the conditional expectation $E_t[f(r_t + J) - f(r_t)]$ can be replaced by the unconditional expectation $E_J[f(r_t + J) - f(r_t)]$. In addition, the first four moments of $J \sim \text{Uniform}[a, b]$ are respectively $E(J) = \frac{(b^2 - a^2)}{2(b - a)}$, $E(J^2) = \frac{(b^3 - a^3)}{3(b - a)}$, $E(J^3) = \frac{(b^4 - a^4)}{4(b - a)}$, and $E(J^4) = \frac{(b^5 - a^5)}{5(b - a)}$. They will be used later for calculating the conditional moments of the short rate.

where $g(\beta)$ is a 4×1 vector and $A(\beta)$ is a 4×4 lower-triangle matrix. These equations are nonlinear functions of the structural parameter vector $\beta = (\kappa, \theta, \sigma, \rho, a, b)'$, which is specified by the jump-diffusion process (1). The differential equation (8) may be more complicated if more general choices of jump size J are considered. Since the coefficients of this non-homogeneous linear first-order differential equation do not depend on time, one obtains the following vector autoregressive solution,

$$E_t(R_{t+1}) = e^{A(\beta)} R_t + A^{-1}(\beta)(e^{A(\beta)} - I)g(\beta), \quad (9)$$

where I is the 4×4 identity matrix. The discrete time sampling frequency is normalized to be one. Explicit expressions of $A(\beta)$ and $g(\beta)$, in terms of the structural parameters, are easily verified as

$$A(\beta) = \begin{bmatrix} -\kappa & 0 & 0 & 0 \\ 2\kappa\theta + \sigma^2 + 2\rho E_J(J) & -2\kappa & 0 & 0 \\ 3\rho E_J(J^2) & 3\kappa\theta + 3\sigma^2 + 3\rho E_J(J) & -3\kappa & 0 \\ 4\rho E_J(J^3) & 6\rho E_J(J^2) & 4\kappa\theta + 6\sigma^2 + 4\rho E_J(J) & -4\kappa \end{bmatrix}$$

and

$$g(\beta) = \begin{bmatrix} \rho E_J(J) + \kappa\theta \\ \rho E_J(J^2) \\ \rho E_J(J^3) \\ \rho E_J(J^4) \end{bmatrix},$$

where the moments of J are nonlinear functions of jump-size parameters a and b .

The matrix form solution to conditional moments developed here is exactly analogous to the Riccati type ODE solution obtained through the separation of variables approach (Duffie et al., 2000), although our method does not rely on the knowledge of the characteristic function of the underlying process. It is not difficult to realize that our matrix solution also applies to a different kind of jump-diffusion process $dr_t = \kappa(\theta - r_t)dt + (\sigma_0 + \sigma_1 r_t)dW_t + JdN(\rho t)$, where the volatility is a quadratic instead of an affine function of the short rate.¹² The moment solution is in the same form as equation (9), except that A will be a full, rather than a lower triangular matrix. Admitting that the closed-form solution to term structure is still unresolved for the quadratic volatility jump-diffusion model, the empirical application in Section 4 will focus on the affine jump diffusion model (1).

¹²A complete characterization of the types of jump-diffusion processes that sustain such a series representation of their distributions can be carried out by following Wong (1964), who studied the case of pure diffusion processes.

3.2 Estimation and Inference

The moment condition (9) generated by Itô's formula above justifies the nonlinear regression hypothesis $R_{t+1} = D(\beta)R_t + C(\beta) + U_{t+1}$, where $U_{t+1} = [u_{1t+1}, u_{2t+1}, u_{3t+1}, u_{4t+1}]'$ is a vector of errors and where $D(\beta) = e^{A(\beta)}$ and $C(\beta) = A^{-1}(\beta)(e^{A(\beta)} - I)g(\beta)$ are nonlinear functions of the structural parameter vector β . The Multivariate Weighted Nonlinear Least Square for Jump-Diffusion (MWNLS-JD) estimator constructed here is in the spirit of Gallant (1987) and Wooldridge (1994),¹³ and solves the minimization problem as follows:

$$\min_{\beta \in \mathcal{B}} Q_T(\beta) \equiv \frac{1}{T} \sum_{t=1}^{T-1} \frac{1}{2} U_{t+1}(\beta)' W_T(\bar{\beta})^{-1} U_{t+1}(\beta), \quad (10)$$

where $\bar{\beta}$ is some consistent estimator of β and where the weighting matrix $W_T(\bar{\beta})$ is a 4×4 symmetric, positive definite matrix with probability one, usually constructed by $W_T(\bar{\beta}) = \frac{1}{T} \sum_{t=1}^{T-1} U_{t+1}(\bar{\beta}) U_{t+1}(\bar{\beta})'$. Either a two-step or an iterated estimator may apply.

3.2.1 Consistency

Given the data-generating process (1) and Assumption (a)-(c) of Section 2.1, it clearly follows that $0 < R_t < \infty$ and $E|R_t| < \infty$.¹⁴ This is equivalent to the strict stationary and ergodic condition.¹⁵ The nonlinear functions $D(\beta)$ and $C(\beta)$ are obviously continuous. The parameter space \mathcal{B} for $\beta \in \mathcal{B} \subset \mathcal{R}^l$ is usually assumed to be compact. Thus the consistency result is primarily driven by the identification condition:

Condition 1 (Identification) *There exists an unique $\beta_0 \in \mathcal{B} \subset \mathcal{R}^l$ such that $E_t(U_{t+1}) = 0$.*

This condition is equivalent to requiring the Hessian of (10) to be full rank. If R_t is stationary, it is typically assumed that $W_T(\bar{\beta}) \xrightarrow{a.s.} W_0$, where W_0 is a symmetric positive definite matrix. The last building block of the consistency result is a uniform law of large numbers. Assumptions (a)-(c) of Section 2.1 deliver an invariant probability measure $P(R_{t+1}, R_t | \beta)$, which is Markovian and stationary. Let $Q(\beta) = \lim_{T \rightarrow \infty} \int Q_T(\beta) dP(\cdot | \beta)$ be the population

¹³Our estimator is indeed a Classical Method of Moments. Singleton (2001) also discusses several estimators for the affine diffusion processes based on the Conditional Characteristic Function.

¹⁴Some of the empirical results in Section 4 do not guarantee the non-negativity of interest rates, although going from weekly to daily data seems to alleviate this problem. Also, specifying the jump size only by its moments may avoid violating the non-negativity restriction.

¹⁵To ensure ergodicity, no state should be absorbing for r_t . The most likely scenarios for zero to become an absorbing state are: (A) $\theta = 0$ with $\rho = 0$ and/or $J = 0$ (reducible), (B) $\theta = 0$ with $\rho > 0$ and $J > 0$ (periodic). Both of these scenarios are clearly ruled out by assumptions (a)-(c).

limit of the objective function; then a standard proof gives the result $Q_T(\beta) \xrightarrow{a.s.} Q(\beta)$ uniformly on \mathcal{B} . In addition, assume that $Q(\beta)$ is continuous on \mathcal{B} . Now we have the consistency result

$$\hat{\beta} \xrightarrow{a.s.} \beta_0, \text{ where } \hat{\beta} = \arg \min_{\beta \in \mathcal{B}} Q_T(\beta). \quad (11)$$

3.2.2 Asymptotic Normality

Due to the Markov property of our jump-diffusion specification (1), the derived moment condition (9) is essentially a martingale difference sequence (MDS). It is straightforward to verify that:

Condition 2 (No Serial Correlation) For all t , $E_t(U_t U_{t+j}') = 0$, for any $j \geq 1$.

Consequently, the scores $s_{t+1}(\beta) = \frac{\partial}{\partial \beta} U_{t+1}(\beta)' W_T(\bar{\beta})^{-1} U_{t+1}(\beta)$ are not serially correlated (Wooldridge, 1994). Assume that β_0 is an interior point of a convex set \mathcal{B} . Define the expected outer product of the scores as $J = \int s_{t+1}(\beta) s_{t+1}(\beta)' dP(\cdot|\beta)$ and its empirical counterpart as $J_T = \frac{1}{T} \sum_{t=1}^{T-1} s_{t+1}(\hat{\beta}) s_{t+1}(\hat{\beta})'$. Standard argument leads to $\sqrt{T} \frac{\partial}{\partial \beta} Q_T(\beta_0) \xrightarrow{D} N(0, J_0)$, where J_0 is J evaluated at the true parameter and is a positive semi-definite matrix. The Hessian is defined as $A = \int \frac{\partial}{\partial \beta} U_{t+1}(\beta)' W_T(\bar{\beta})^{-1} \frac{\partial}{\partial \beta} U_{t+1}(\beta) dP(\cdot|\beta)$ and its empirical estimate as $A_T = \frac{1}{T} \sum_{t=1}^{T-1} \frac{\partial}{\partial \beta} U_{t+1}(\hat{\beta})' W_T(\bar{\beta})^{-1} \frac{\partial}{\partial \beta} U_{t+1}(\hat{\beta})$. Assume that $A_T(\beta) \xrightarrow{P} A$ uniformly in a neighborhood of β_0 and that A is continuous at β_0 . If we let A_0 be the Hessian evaluated at the true parameter, we arrive at the asymptotic normality result

$$\sqrt{T}(\hat{\beta} - \beta_0) \xrightarrow{D} N(0, A_0^{-1} J_0 A_0^{-1}). \quad (12)$$

To make inferences about the parameter value, the asymptotic variance should be estimated by the usual sandwich formula $\widehat{\text{Avar}}(\hat{\beta}) = \frac{1}{T} A_T^{-1} J_T A_T^{-1}$, which is a White's heteroscedasticity robust estimator. Because of Condition 2, no serial correlation needs to be considered.

3.2.3 Conditional Moment Test

A conditional moment-type test (Newey, 1985; Tauchen, 1985) can be constructed from the errors of the fitted moments. It only requires the estimation of the restricted model. The error vector of the first two moments is $\mathcal{U}_{t+1} = [r_{t+1} - E_t(r_{t+1}), r_{t+1}^2 - E_t(r_{t+1}^2)]'$, which has a distribution with $E_t(\mathcal{U}_{t+1}) = [0, 0]'$ and $\text{Var}_t(\mathcal{U}_{t+1}) = \Omega_{t+1}$. With the first four conditional moments solved by (9), we can easily calculate Ω_{t+1} as

$$\Omega_{t+1} = \begin{bmatrix} E_t(r_{t+1}^2) & E_t(r_{t+1}^3) \\ E_t(r_{t+1}^3) & E_t(r_{t+1}^4) \end{bmatrix} - \begin{bmatrix} E_t(r_{t+1})E_t(r_{t+1}) & E_t(r_{t+1})E_t(r_{t+1}^2) \\ E_t(r_{t+1}^2)E_t(r_{t+1}) & E_t(r_{t+1}^2)E_t(r_{t+1}^2) \end{bmatrix}.$$

By construction, Ω_{t+1} is symmetric and positive definite. Applying a suitable version of the central limit theorem, one has

$$\frac{1}{T} \sum_{t=1}^{T-1} \mathbf{u}'_{t+1} \Omega_{t+1}^{-1} \mathbf{u}_{t+1} \xrightarrow{D} \text{Chi-Square}(2). \quad (13)$$

The regularity condition for the central limit theorem is very mild, since the error vector is not serially dependent and since the conditional heteroscedasticity, as a function of stationary r_t , is well-bounded.¹⁶

4 Empirical Application

This section conducts extensive empirical comparisons between the square-root diffusion and jump-diffusion processes. Four data sets—the weekly Federal Funds rate (Weekly FF), daily Federal Funds (Daily FF) rate, weekly 3-month T-bill rate (Weekly 3MTB), and daily 3-month T-bill rate (Daily 3MTB)—from 1954 to 1999 are used to estimate the interest rate models. To control for the potential structural breaks, subsample analyses of 1954-1977, 1978-1987, and 1988-1999 are also performed. To ensure robustness, both the semiparametric specification of jump-size and the public announcement of rate changes are used for cross validations. Finally, the risk-neutral pricing errors are extracted from the CIR and CIR-Jump models and are examined for any predictive pattern.

4.1 Data Description

Table 1 presents the summary statistics for the weekly Federal Funds rate (Weekly FF), daily Federal Funds (Daily FF) rate, weekly 3-month T-bill rate (Weekly 3MTB), and daily 3-month T-bill rate (Daily 3MTB) data from 1954 to 1999, which are obtained from the public website of the Federal Reserve System. On average, the Federal Funds rate is 60 basis points higher than the 3-month T-bill rate, with more volatility, larger skewness, and higher

¹⁶As pointed out by Newey (1985) and Tauchen (1985), the asymptotic variance of the test statistic (13) is understated if the estimated parameter value is used instead of the population truth. Hence the conditional moment test is more conservative. Other standard tests are also available but are not adopted here for the following concerns. The Wald test lacks invariance since the parameter enters the objective function nonlinearly in the model considered here. The likelihood ratio type test requires the correct specification of the conditional variance matrix—5th to 8th moments in this case, since the first four moments are being matched—and the scaling of a matrix is so unbalanced that its inversion has unsurmountable numerical difficulties. The Rao’s score test or LM test, although robust to misspecification of the weighting matrix, is arbitrary in specifying the dimensions of the alternative model. An earlier version of this paper documented that the LM test was unable to distinguish between the pure diffusion and jump-diffusion models.

kurtosis. Between daily and weekly observations, the unconditional mean and variance are almost the same, although the high frequency data is slightly more skewed and leptokurtotic. These features are clearly seen from the time series plots in Figure 1.

In the literature it is typical to use the 3-month T-bill rate as a proxy for the unobservable short rate, as the trading of the 3-month T-bill is the most liquid on the secondary market. Certain idiosyncratic liquidity and microstructure effects pertaining to shorter maturity proxies, such as the 1-month T-bill rate, the 7-day Eurodollar rate, or overnight Federal Funds rate, can be avoided by using the 3-month T-bill rate. On the other hand, using proxies may introduce significant bias in estimating the volatility function (Chapman et al., 1999). Since jumps to the short rate are in large sizes by construction, their impacts may be dramatically understated when using longer maturity proxies. In addition, the overnight Federal Funds rate market is completed by derivatives like futures and options, so its role as literally the short end of the term structure is much more appreciated by the investing public.

Another consideration in estimating short-rate dynamics is the need to use high frequency data to identify the jump impact. Jumps are rare and in large size. Low frequency data—monthly observations, for example—can not distinguish if a large change in the short rate is coming from a single jump or from successive small adjustments. Of course the problem of using high frequency data—daily observations, for example—is the contamination of calendar effects or microstructure noise.

4.2 Estimation Results and Moment-Based Diagnostics

Table 2 reports the main estimation results. The diffusion (CIR) model is strongly rejected for all four interest-rate data sets. The jump-diffusion (CIR-Jump) model is rejected for the Daily 3MTB and Weekly 3MTB (marginally) rates but not for the two Federal Funds rates. All the parameters are estimated significantly, except the local variance σ in the CIR-Jump model for the two 3-month T-bill rates. As pointed out by Chapman et al. (1999), certain systematic biases can be introduced when using the 3-month rate as a proxy for the instantaneous short rate. Also, the jump-size lower bound is negative for the two 3-month T-bill rates, thus violating the regularity assumption (c) in Section 2.1 and possibly resulting in negative interest rates. Therefore it is not surprising that the jump-diffusion models for the 3-month T-bill rates are rejected, while they are not rejected for the Federal Funds rate estimates that do not violate the non-negativity constraint.

The parameter estimates of diffusion (CIR) and jump-diffusion (CIR-Jump) processes in Table 2 are further utilized to deduce the unconditional moments, which are compared with their data counterparts. Further, the jump-rate and jump-size estimates can provide some guidance on the annualized jump frequencies and on jump-size mean and standard deviation.

Table 3 shows that the deviation between the sample mean and the model-implied mean has a wide range from 5 basis points (CIR for Weekly FF) to 100 basis points (CIR for Daily 3MTB).¹⁷ For higher-order moments—standard deviation, positive skewness, and excess kurtosis—both CIR and CIR-Jump model-implied estimates are understating the sample averages, regardless whether the data are weekly and daily. However, one important feature is consistently remarkable across all four data sets: the jump-diffusion estimate of unconditional volatility is much higher than the pure-diffusion estimate and is much closer to the sample average. Improved volatility fitting is indeed the key motivation of the jump-diffusion modeling of this paper.

The non-rejected jump-diffusion models for the two Federal Funds rates predict one jump every two years by weekly data and four jumps per year by daily data. Their implication for jump sizes are quite homogeneous: about 200 basis points in jump mean and 50 basis points in jump standard deviation. Later subsample analysis confirms that a large jump-size mean is predominantly driven by the irregular monetary activity of the early 1980s; also, later comparison with the public announcement of rate changes verifies that four jumps per year is quite the norm for Federal Funds rate targeting. For the rejected two 3-month T-bill estimates of jump-diffusion model, there is about one jump per year, while the jump size has contradicting signs from the weekly and daily data. Other defects in daily 3-month T-bill (Daily 3MTB) estimation are negative jumps, negative interest rates, and negative skewness (contrary to the data). Note that the jump frequency and jump magnitudes implied by the Federal Funds rate market reflect the volatile liquidity shocks to the large banks at the very short end of the term structure.

¹⁷It is common knowledge that for a stationary yet persistent time series, sample average is a very inaccurate estimate of the unconditional mean. For a CIR-type interest rate process, the precision of the sample average as an estimate of unconditional mean θ decreases dramatically when the mean reversion parameter κ is close to zero. Pritsker (1998) provides related asymptotic theory and Monte Carlo evidence.

4.3 Structural Breaks, Semiparametric Jumps, and Public Announcements

The sample period 1954-1999 experienced several structural breaks in monetary policy targeting, thus raising a well-known issue of parameter stability (Brown and Dybvig, 1986; Chan et al., 1992). The jump specification in this paper does not distinguish between structural shifts and regular jumps. Although the full sample jump-diffusion estimate (especially for Daily FF) is successful, the jump size of 2-3% seems too large for the last decade yet too small for the early 1980s. This judgment can be verified by controlling the structure breaks. An objective subsample analysis (Table 4) reinforces the rejection of the pure CIR diffusion process in each short sample. The parameter estimates across the three periods, albeit some differences exist, are quite stable—do not reflect the potential structural breaks. On the other hand, the jump-diffusion (CIR-Jump) subsample estimation is more revealing: the model is rejected for the monetary experiment era of 1978-1987, while this model is not rejected before 1978 (marginally) and after 1987 (soundly). The parameter estimates, especially those of the jump component, are dramatically different across the three time frames. Structural shifts are certainly the chief suspect, but one can not rule out the small sample effect, since the jump-diffusion process estimate for the entire sample is not rejected.

A legitimate criticism of the Uniform jump-size assumption is that it is unrealistic: the central tendency of jumps may well be around 3% in 1980s but close to 25 basis points in 1990s, as suggested by the above subsample estimation results. Also since the realized jumps are rare, the parametric jump size can be poorly estimated and sensitive to outliers. The real power of the moment-based MWNLS-JD estimator for the jump-diffusion model here is to allow a semiparametric modeling of the jump size in up to its fourth moments. The estimation results turn out to be very reasonable (see Table 5). Similar to the parametric results in Table 2, the Weekly FF and Daily FF are not rejected with very significant parameter estimates. The Daily 3MTB is strongly rejected with implausible sign and magnitude for jump parameters, while the Weekly 3MTB is not rejected with insignificant estimates of local variance and jump-size mean. It is arguable that using the 3-month rate as a proxy for the short rate may overlook some important microstructure effects, hence missing some jumps at the short end of the term structure. In addition, the non-rejected Federal Funds rate estimates produce similar jump rates and sizes as the parametric cases. The bottom line is that the semiparametric jump-size method should be more robust to the specification error and, at the same time, can provide a sanity check on the parametric jump models.

A comparison between the estimated jump intensity and jump size with the official jump announcements can be informative and may shed some light on (1) why the Daily FF rate CIR-Jump model is rejected for the time period 1978-1987 and (2) why the estimated mean jump size is as large as 2-3%. Table 6 summarizes the publicly announced short-rate changes or jumps. Comparing with the daily Federal Funds rate (Daily FF) estimates in Table 2, the estimated jump intensity—four times a year—match with the public announcements quite well. Interestingly, for the time period 1978-1987, the maximum announced rate jump is as large as 9.25%, whereas for the recent era 1988-1999, the average announced rate change is 38 basis points. This explains why the average jump size in both the parametric Uniform jump model (Table 2) and the semiparametric moment-based jump model (Table 5) ranges from 200 to 300 basis points for the entire 1954-1999 time span. In addition, 1978-1987 represents an experimental period in which the Federal Funds rate is not a monetary policy target; rather, the rate changes or jumps resulted from a targeting of the borrowed or nonborrowed reserves. This may be the reason why the jump-diffusion model for the best performing daily Federal Funds rate (Daily FF) is still rejected in this period.

Such a casual comparison should be considered with some caution. In the case of jump intensities, some announced rate changes are not jumps per se, since they are well anticipated by the market participants and are fully discounted long before the announcement dates; on the other hand, many jumps perceived by market due to drastic changes in economic conditions—such as the numerous occasions of emerging market crises—are not always reflected in announced rate changes. Similar concerns apply to the jump size as well, since the market-driven rate jumps may be on average larger or smaller than the publicly announced rate changes.

4.4 Pricing Error and Predictability of Risk Premia

Once the short-rate process is estimated, it can be used to price the bond yield. Of immediate interest is the comparison of the pricing errors between the CIR and CIR-Jump short-rate models. A further look at the extracted risk premia from the point-in-time yield curves also provides useful diagnostics on the various pricing models. Table 7 performs risk-neutral bond pricing exercise. The monthly bond-yield data for maturities at months 1, 3, 6, 9 and years 1, 2, 3, 4, 5 are retrieved from the Center for Research in Security Prices (CRSP) between June 1964 and December 1999. There are a total of 427 months of observations. The risk premia extracted from the time-average yield curve have reasonable signs and magnitudes, with the

jump-diffusion model producing a smaller pricing error. Extracting the risk premia from point-in-time yield curves involve 427 nonlinear optimizations for both CIR and CIR-Jump models. The upshot is that, at most representative quantiles, pricing error in the CIR-Jump model is about half the size of that in the CIR model.

Table 8 looks at whether there is any predictive pattern of the extracted point-in-time risk premia. If the model is correct, the extracted market prices of risks (parameters) should not be correlated with the underlying risk factors or observed yields. This is the flip side of extracting the time-invariant risk premia and requiring the pricing errors to be void of any dynamic structure (Chen and Scott, 1993; Duffie and Singleton, 1997; Jagannathan et al., 2000). It is clear from Figure 2 that the predictability of $\lambda_W(t)$ in the CIR model by observed yields clearly violates the risk-neutral pricing rule, because any systematic covariance should be explicitly modeled into the pricing kernel process. At the same time, the risk premia extracted from the jump-diffusion (CIR-Jump) model lack any systematic trend or pattern (see the lower panels in Figure 2), and the time variations seem to be caused by some random and idiosyncratic pricing errors.

5 Conclusion

This paper studies a jump-diffusion interest-rate model under a no-arbitrage condition. The contribution is to design a computationally efficient estimation strategy with a diagnostic specification test. Analytical solutions are provided for both the conditional moments of the short rate and the entire term structure of different maturities.

When applied to the weekly and daily Federal Funds rates and 3-month T-bill rates, the pure diffusion model is rejected in all occasions, while the jump-diffusion process finds clear support in the short end maturity and even more confirmation at the level of daily frequency. Not only is the volatility better matched, but the jump timing and magnitude also become more reasonable. Subsample analysis suggests that jump-diffusion modeling is solid, except for in the monetary experiment period of 1978-1987. These findings are further corroborated by estimating a less parametric jump process and by comparing the results with the publically announced short-rate changes. Finally, the risk-neutral pricing exercise reveals that jump augmentation can reduce the predictive pattern of the extracted point-in-time risk premia.

The challenge of fitting the short-term interest rate is to accommodate the rich volatility

feature. Jump-diffusion modeling is one of numerous efforts that are working in this direction. An important lesson from this study is that incorporating jumps and using higher-order moments can help to model and estimate volatility more accurately. This necessity is due to the fact that the fat-tail characteristics of interest-rate data cannot be well explained by a smooth sample path distribution with stationarity restrictions. The remaining irregularities of parameter estimates may indicate some specification errors, which are introduced by this univariate jump-diffusion model of interest rates.

References

- Ahn, Chang Mo and Howard E. Thompson (1988), “Jump-Diffusion Process and the Term Structure of Interest Rates,” *Journal of Finance*, vol. 43, 155–174.
- Aït-Sahalia, Yacine (1996a), “Nonparametric Pricing of Interest Rate Derivatives,” *Econometrica*, vol. 64, 527–560.
- Aït-Sahalia, Yacine (1996b), “Testing Continuous-Time Models of the Spot Interest Rate,” *The Review of Financial Studies*, vol. 9, 385–426.
- Andersen, Torben G. and Jesper Lund (1996), “Stochastic Volatility and Mean Drift in the Short Rate Diffusion: Source of Steepness, Level, and Curvature of the Yield Curve,” *Working Paper*, Kellogg School of Management, Northwestern University.
- Andersen, Torben G. and Jesper Lund (1997), “Estimating Continuous Time Stochastic Volatility Models of the Short Term Interest Rate,” *Journal of Econometrics*, vol. 77, 343–378.
- Balduzzi, Pierluigi, Sanjiv R. Das, Silverio Foresi, and Rangarajan K. Sundaram (1996), “A Simple Approach to Three Factor Affine Term Structure Models,” *Journal of Fixed Income*, vol. 6, 43–53.
- Bandi, Federico M. and Peter C. B. Phillips (2000), “Fully Econometric Estimation of Scalar Diffusion Models,” *Working Paper*, Graduate School of Business, University of Chicago.
- Baz, Jamil and Sanjiv Ranjan Das (1996), “Analytical Approximation of the Term Structure for Jump-Diffusion Process: A Numerical Analysis,” *Journal of Fixed Income*, vol. 6, 78–86.
- Brown, Stephen J. and Philip H. Dybvig (1986), “The Empirical Implications of the Cox, Ingersoll, Ross Theory of the Term Structure of Interest Rates,” *Journal of Finance*, vol. 41, 617–630.
- Chacko, George and Sanjiv Das (1999), “Pricing Interest Rate Derivatives: A General Approach,” *Working Paper*, Graduate School of Business, Harvard University.
- Chacko, George and Luis M. Viceira (1999), “Spectral GMM Estimation of Continuous-Time Processes,” *Working Paper*, Graduate School of Business, Harvard University.

- 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*, vol. 47, 1209–1227.
- Chapman, David A., John B. Long, and Neil D. Pearson (1999), “Using Proxies for the Short Rate: When Are Three Months Like an Instant,” *Review of Financial Studies*, vol. 12, 763–806.
- Chen, Lin (1996), “Stochastic Mean and Stochastic Volatility - A Three-Factor Model of the Term Structure of Interest Rates and Its Applications in Derivatives Pricing and Risk Management,” *Financial Markets, Institution and Instruments*, vol. 5, 1–88.
- Chen, Ren-Raw and Louis Scott (1993), “Maximum Likelihood Estimation for a Multifactor Equilibrium Model of the Term Structure of Interest Rates,” *Journal of Fixed Income*, vol. 3, 14–31.
- Christoffersen, Peter F. and Francis X. Diebold (2000), “How Relevant Is Volatility Forecasting for Financial Risk Management?” *Review of Economics and Statistics*, vol. 82, 12–22.
- Conley, Tim, Lars Peter Hansen, Erzo Luttmer, and Jose Scheinkman (1997), “Short Term Interest Rates as Subordinated Diffusions,” *Review of Financial Studies*, vol. 10, 525–578.
- Cox, John C., Jonathan E. Ingersoll, and Stephen A. Ross (1985a), “A Theory of the Term Structure of Interest Rates,” *Econometrica*, vol. 53, 385–407.
- Cox, John C., Jonathan E. Ingersoll, and Stephen A. Ross (1985b), “An Intertemporal General Equilibrium Model of Asset Prices,” *Econometrica*, vol. 53, 363–384.
- Dai, Qiang and Kenneth J. Singleton (2000), “Specification Analysis of Affine Term Structure Models,” *Journal of Finance*, vol. 55, 1943–1978.
- Das, Sanjiv Ranjan (1998), “Poisson-Gaussian Process and the Bond Markets,” *Working Paper*, Harvard Business School.
- Duarte, Jefferson (1999), “The Relevance of the Parameterization of Price of Risk in Affine Term Structure Models,” *Working Paper*, Graduate School of Business, University of Chicago.

- Duffee, Gregory (2000), “Forecasting Future Interest Rates: Are Affine Models Failures?” *Working Paper*, Haas School of Business, University of California at Berkeley.
- Duffie, Darrell, Jun Pan, and Kenneth Singleton (2000), “Transform Analysis and Asset Pricing for Affine Jump-Diffusions,” *Econometrica*, vol. 68, 1343–1376.
- Duffie, Darrell and Kenneth Singleton (1993), “Simulated Moments Estimation of Markov Models of Asset Prices,” *Econometrica*, vol. 61, 929–952.
- Duffie, Darrell and Kenneth Singleton (1997), “An Econometric Model of the Term Structure of Interest-Rate Swap Yields,” *Journal of Finance*, vol. 52, 1287–1321.
- Feller, William (1951), “Two Singular Diffusion Problems,” *Annals of Mathematics*, vol. 54, 173–182.
- Fisher, Mark and Christian Gilles (1996), “Estimating Exponential Affine Models of the Term Structure,” *Working Paper*, Finance and Economics Discussion Series, Federal Reserve Board.
- Gallant, A. Ronald (1987), *Nonlinear Statistical Models*, Wiley, New York.
- Gallant, A. Ronald and George Tauchen (1996), “Which Moment to Match?” *Econometric Theory*, vol. 12, 657–681.
- Gallant, A. Ronald and George Tauchen (1998), “Reprojecting Partially Observed Systems with Application to Interest Rate Diffusions,” *Journal of the American Statistical Association*, vol. 93, 10–24.
- Gibbons, Michael R. and Krishna Ramaswamy (1993), “A Test of the Cox, Ingersoll, and Ross Model of the Term Structure,” *Review of Financial Studies*, vol. 6, 619–658.
- Hansen, Lars Peter and José Alexandre Scheinkman (1995), “Back to the Future: Generalized Moment Implications for Continuous Time Markov Process,” *Econometrica*, vol. 63, 767–804.
- Hansen, Lars Peter, José Alexandre Scheinkman, and Nizar Touzi (1998), “Spectral Methods for Identifying Scalar Diffusions,” *Journal of Econometrics*, vol. 86, 1–32.

- Heston, Steven (1993), “A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options,” *Review of Financial Studies*, vol. 6, 327–343.
- Ingersoll, Jonathan E. (1987), *Theory of Financial Decision Making*, Rowman & Littlefield.
- Jagannathan, Ravi, Andrew Kaplin, and Guoqiang Sun (2000), “An Evaluation of Multi-Factor CIR Models Using LIBOR, Swap Rates, and Cap and Swaption Prices,” *Working Paper*, Kellogg Graduate School of Management, Northwestern University.
- Johannes, Michael S. (1999), “Jumps in Interest Rates: A Nonparametric Approach,” *Working Paper*, Department of Economics, University of Chicago.
- Lo, Andrew W. (1988), “Maximum Likelihood Estimation of Generalized Itô Process with Discretely Sampled Data,” *Econometric Theory*, vol. 4, 231–247.
- Merton, Robert C. (1971), “Optimum Consumption and Portfolio Rules in a Continuous Time Model,” *Journal of Economic Theory*, vol. 3, 373–413.
- Merton, Robert C. (1976), “Option Pricing When Underlying Stock Returns Are Discontinuous,” *Journal of Financial Economics*, vol. 3, 125–144.
- Newey, Whitney K. (1985), “Maximum Likelihood Specification Testing and Conditional Moment Tests,” *Econometrica*, vol. 53, 1047–1070.
- Pearson, Neil D. and Tong-Sheng Sun (1994), “Exploiting the Conditional Density in Estimating the Term Structure: An Application to the Cox, Ingersoll, and Ross Model,” *Journal of Finance*, vol. 49, 1279–1304.
- Piazzesi, Monika (2000), “An Econometric Model of the Yield Curve with Macroeconomic Jump Effects,” *Working Paper*, Graduate School of Business, Stanford University.
- Pritsker, Matt G. (1998), “Nonparametric Density Estimation and Tests of Continuous Time Interest Rate Models,” *Review of Financial Studies*, vol. 11, 449–487.
- Singleton, Kenneth (2001), “Estimation of Affine Asset Pricing Models Using the Empirical Characteristic Function,” *Journal of Econometrics*, forthcoming.

- Stanton, Richard (1997), “A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk,” *The Journal of Finance*, vol. 52, 1973–2002.
- Tauchen, George (1985), “Diagnostic Testing and Evaluation of Maximum Likelihood Models,” *Journal of Econometrics*, vol. 30, 415–443.
- Tauchen, George (1997), “New Minimum Chi-Square Methods in Empirical Finance,” in “Advances in Econometrics, Seventh World Congress,” (edited by Kreps, D. and K. Wallis), Cambridge University Press, Cambridge UK.
- Vasicek, Oldrich A. (1977), “An Equilibrium Characterization of the Term Structure,” *Journal of Financial Economics*, vol. 5, 177–188.
- Wong, Eugene (1964), “The Construction of a Class of Stationary Markoff Processes,” in “Sixteenth Symposium in Applied Mathematics - Stochastic Process in Mathematical Physics and Engineering,” (edited by Belleman, R.), pages 264–276, Providence, RI.
- Wooldridge, Jeffrey M. (1994), *Handbook of Econometrics*, vol. 4, chap. 45, Estimation and Inference for Dependent Processes, pages 2642–2738, Elsevier Science B. V.

6 Tables and Figures

Table 1: Summary Statistics of Interest-Rate Data

The weekly Federal Funds rate (Weekly FF) is from 07/07/1954 to 12/29/1999, with a total of 2374 observations; the daily Federal Funds rate (Daily FF) is from 07/01/1954 to 12/31/1999, with a total of 16620 observations; the weekly 3 month T-bill rate (Weekly 3MTB) is from 01/08/1954 to 12/31/1999, with a total of 2400 observations; and the daily 3 month T-bill rate (Daily 3MTB) is from 01/04/1954 to 12/31/1999, with a total of 11488 observations. All data are obtained from the public website of the Federal Reserve System.

Interest Rate	Weekly FF	Daily FF	Weekly 3MTB	Daily 3MTB
Moments				
Mean	0.0609	0.0609	0.0550	0.0550
Std Deviation	0.0337	0.0339	0.0280	0.0280
Skewness	1.2364	1.2443	1.1316	1.1366
Kurtosis	5.0229	5.1081	4.7797	4.8087
Quantiles				
Minimum	0.0020	0.0013	0.0058	0.0055
05 percent	0.0189	0.0175	0.0207	0.0206
25 percent	0.0350	0.0350	0.0350	0.0349
50 percent	0.0543	0.0542	0.0505	0.0506
75 percent	0.0812	0.0804	0.0704	0.0703
95 percent	0.1255	0.1260	0.1065	0.1066
Maximum	0.2006	0.2236	0.1676	0.1714

Table 2: Parameter Estimation and Specification Test

This table reports the Multivariate Weighted Nonlinear Least Square (MWNLS) estimation result of the jump-diffusion interest-rate process (1) in Section 2 (labeled as CIR-Jump) and its square-root counterpart (labeled as CIR). Details on solving the moment conditions and constructing parameter estimate, standard error, and specification test can be found in Sections 3.1 and 3.2. Four data sets—Weekly FF, Daily FF, Weekly 3MTB, and Daily 3MTB—are used as proxies for the instantaneous short rate, and their time series characteristics are reported in Table 1 and Figure 1. White’s Heteroscedasticity Robust estimates of standard errors are reported in the parentheses.

Data	Weekly FF		Daily FF		Weekly 3MTB		Daily 3MTB	
Model	CIR	CIR-Jump	CIR	CIR-Jump	CIR	CIR-Jump	CIR	CIR-Jump
Parameter Estimates of Diffusion Component								
θ	0.0604 (0.0024)	0.0422 (0.0020)	0.0601 (0.0009)	0.0376 (0.0011)	0.0464 (0.0007)	0.0249 (0.0017)	0.0449 (0.0003)	0.0836 (0.0003)
κ	0.0116 (0.0008)	0.0117 (0.0026)	0.0119 (0.0005)	0.0120 (0.0013)	0.0053 (0.0014)	0.0056 (0.0020)	0.0011 (0.0002)	0.0012 (0.0001)
σ	0.0150 (0.0039)	0.0130 (0.0041)	0.0164 (0.0019)	0.0131 (0.0017)	0.0053 (0.0018)	2.129e-5 (0.0283)	0.0025 (0.0003)	1.312e-5 (0.0010)
Parameter Estimates of Jump Component								
a		0.0113 (0.0045)		0.0272 (0.0016)		-0.0141 (0.0022)		-0.0229 (0.0007)
b		0.0312 (0.0044)		0.0314 (0.0015)		0.0325 (0.0021)		0.0049 (0.0007)
ρ		0.0110 (0.0025)		0.0108 (0.0010)		0.0157 (0.0020)		0.0045 (0.0001)
Conditional Moment Specification Testing								
$\chi^2(2)$	16.8226	1.8405	73.084	3.0475	26.877	6.0373	30.003	21.799
p-Value	0.0002	0.3984	0.0000	0.2179	0.0000	0.0489	0.0000	0.0000

Table 3: Diagnostics of Estimated Model Specifications

The unconditional moments of interest rates are obtained from the formula of conditional moments (9) by letting the time span go to infinity, $\lim_{T \rightarrow \infty} E_t(R_T)$. The jump frequency is calculated by annualizing the weekly and daily jump-rate estimates, and the jump size moments are deduced from the Uniform distribution assumption.

Moments	Weekly Federal Funds			Daily Federal Funds		
	Data	CIR	CIR-Jump	Data	CIR	CIR-Jump
Data and Model Implied Moments for Short Rates						
Mean	0.0609	0.0604	0.0622	0.0609	0.0601	0.0638
Std. Dev.	0.0337	0.0241	0.0260	0.0339	0.0260	0.0289
Skewness	1.2364	0.7987	0.7640	1.2443	0.8650	0.8010
Kurtosis	5.0229	3.9570	3.7887	5.1081	4.1223	3.8259
Model Implied Moments for Jumps						
Jumps Per Year			0.5695			3.8880
Jump Size Mean			0.0213			0.0293
Jump Size Std. Dev.			0.0057			0.0052
Moments	Weekly 3-Month T-Bill			Daily 3-Month T-Bill		
	Data	CIR	CIR-Jump	Data	CIR	CIR-Jump
Data and Model Implied Moments for Short Rates						
Mean	0.0550	0.0464	0.0507	0.0550	0.0449	0.0488
Std. Dev.	0.0280	0.0111	0.0193	0.0280	0.0116	0.0168
Skewness	1.1316	0.4801	0.7519	1.1366	0.5161	-0.6763
Kurtosis	4.7797	3.3458	3.8006	4.8087	3.3996	3.5567
Model Implied Moments for Jumps						
Jumps Per Year			0.8164			1.1250
Jump Size Mean			0.0092			-0.0090
Jump Size Std. Dev.			0.0135			0.0080

Table 4: Subsample Estimates with the Daily Federal Funds Rate

All estimations are conducted in exactly the same way as in Table 2. The breaking of subsample periods is mechanically done by the tenures of the Federal Reserve chairmen—the Martin (and Burns) era 1954-1977, (Mill and) Volcker era 1978-1987, and Greenspan era 1988-1999.

	1954-1977		1978-1987		1988-1999	
Model	CIR	CIR-Jump	CIR	CIR-Jump	CIR	CIR-Jump
Parameter Estimates of Diffusion Component						
θ	0.0611 (0.0032)	0.0236 (0.0012)	0.0950 (0.0009)	0.0480 (0.0011)	0.0527 (0.0005)	0.0273 (0.0005)
κ	0.0168 (0.0019)	0.0219 (0.0008)	0.0162 (0.0011)	0.0163 (0.0013)	0.0138 (0.0001)	0.0144 (0.0004)
σ	0.0082 (0.0022)	0.0150 (0.0021)	0.0142 (0.0019)	1.48e-5 (0.0064)	0.0100 (0.0008)	2.20e-5 (0.0014)
Parameter Estimates of Jump Component						
a		0.0119 (0.0031)		0.0244 (0.0033)		-0.0081 (0.0006)
b		0.0130 (0.0035)		0.0362 (0.0032)		0.0146 (0.0006)
ρ		0.0433 (0.0015)		0.0281 (0.0018)		0.1296 (0.0004)
Conditional Moment Specification Testing						
$\chi^2(2)$	1953.2	7.0771	31.745	29.945	18.791	4.1866
p-Value	0.0000	0.0291	0.0000	0.0000	0.0000	0.1233

Table 5: MWNLS-JD Estimation of Semiparametric Jump-Size Model

All estimations are conducted in exactly the same way as in Table 2, except that the jump size is only specified up to its first and second moments, while the skewness and kurtosis are fixed at 0 and 3. This guarantees identification yet retains comparability.

Data	Weekly FF	Daily FF	Weekly 3MTB	Daily 3MTB
Parameter Estimates of Diffusion Component				
θ	0.0369 (0.0021)	0.0357 (0.0013)	0.0427 (0.0024)	0.0469 (0.0008)
κ	0.0117 (0.0032)	0.0120 (0.0012)	0.0056 (0.0024)	0.0011 (0.0005)
σ	0.0131 (0.0037)	0.0131 (0.0023)	1.19e-5 (0.0106)	0.0024 (0.0004)
Parameter Estimates of Jump Component				
μ_J	0.0164 (0.0021)	0.0267 (0.0014)	0.0035 (0.0022)	4.62e-5 (0.0003)
σ_J	0.0040 (0.0067)	0.0021 (0.0018)	0.0174 (0.0040)	1.37e-5 (0.0048)
ρ	0.0179 (0.0019)	0.0127 (0.0012)	0.0126 (0.0023)	-0.0264 (0.0016)
Conditional Moment Specification Testing				
$\chi^2(2)$	2.0511	3.2360	1.4275	35.037
p-Value	0.3586	0.1983	0.4898	0.0000

Table 6: Comparison with Rate Changes Announced by Monetary Authority

The data about public announcements of changes in the discount rate and Federal Funds rate are obtained from the public website of the Federal Reserve Bank of New York. All data are based on available records of Federal Open Market Committee (FOMC) meetings since 1971. Note that the Federal Funds rate was not a primary monetary policy target from the late 1970s until the late 1980s. During that interim period, the official rate changes resulted from market forces interacting with operating targets such as nonborrowed reserves or borrowed reserves.

Absolute Changes in Discount Rate				
	1971-1999	1971-1977	1978-1987	1988-1999
Jumps/Year	2.5172	2.7143	3.5000	1.5833
Mean	0.0056	0.0042	0.0069	0.0047
Std. Dev.	0.0026	0.0019	0.0027	0.0022
Min.	0.0025	0.0025	0.0025	0.0025
Max.	0.0100	0.0100	0.0100	0.0100
Absolute Changes in Federal Funds Rate				
	1971-1999	1971-1977	1978-1987	1988-1999
Jumps/Year	4.5862	4.7143	5.9000	3.4167
Mean	0.0094	0.0098	0.0131	0.0038
Std. Dev.	0.0124	0.0081	0.0165	0.0026
Min.	0.0013	0.0013	0.0013	0.0013
Max.	0.0925	0.0375	0.0925	0.0138

Table 7: Extracting Market Prices of Risks

Pricing error is the standard deviation of the minimized difference between an average observed yield curve and a model-implied yield curve. Monthly discount bond yields for maturities of months 1, 3, 6, 9 and years 1, 2, 3, 4, 5 are obtained from the Center for Research in Security Prices (CRSP) for the period between June 1964 and December 1999. The top row reports the extracted risk premia from a time-average yield curve, while the bottom row from 427 months point-in-time yield curves. White's robust standard error is given in parentheses (as the q 's quantile).

	Diffusion Risk λ_W	Jump Risk λ_J	Absolute Pricing Error
Nonlinear Least Square from One Average Yield Curve			
CIR Model	-0.8348 (0.0196)		0.0008
CIR-Jump Model	-1.8525 (0.2531)	1.0015 (0.0010)	0.0006
Nonlinear Least Squares from 427 Point-in-Time Yield Curves			
CIR Model			
10% Quantile	-2.0220 (0.0166)		0.0007
30% Quantile	-1.1583 (0.0327)		0.0015
Medium	-0.5626 (0.0495)		0.0023
70% Quantile	0.1141 (0.0705)		0.0033
90% Quantile	1.2054 (0.1343)		0.0052
CIR-Jump Model			
10% Quantile	-3.6390 (0.1469)	0.9753 (0.0008)	0.0004
30% Quantile	-3.3386 (0.4124)	0.9868 (0.0018)	0.0008
Medium	-2.2747 (1.2470)	0.9984 (0.0045)	0.0011
70% Quantile	-1.5770 (6.3501)	1.0033 (0.0155)	0.0018
90% Quantile	9.7309 (25.2988)	1.0088 (0.5119)	0.0038

Table 8: Predictability of Extracted Market Prices of Risks

The correlations of extracted point-in-time risk premia with observed bond yield and yield spread are calculated here to search for any predicative or systematic pattern in the market prices of risks.

Correlation	CIR λ_W	CIR-Jump λ_W	CIR-Jump λ_J
1 Month Yield	-0.8206	-0.0153	0.0847
3 Month Yield	-0.8451	-0.0120	0.0769
6 Month Yield	-0.8781	-0.0166	0.0819
9 Month Yield	-0.8943	-0.0178	0.0828
1 Year Yield	-0.8943	-0.0275	0.0899
2 Year Yield	-0.9224	-0.0432	0.1003
3 Year Yield	-0.9276	-0.0542	0.1082
4 Year Yield	-0.9268	-0.0657	0.1159
5 Year Yield	-0.9248	-0.0673	0.1133
5 Year 1 Month Spread	-0.0628	-0.0961	0.0402

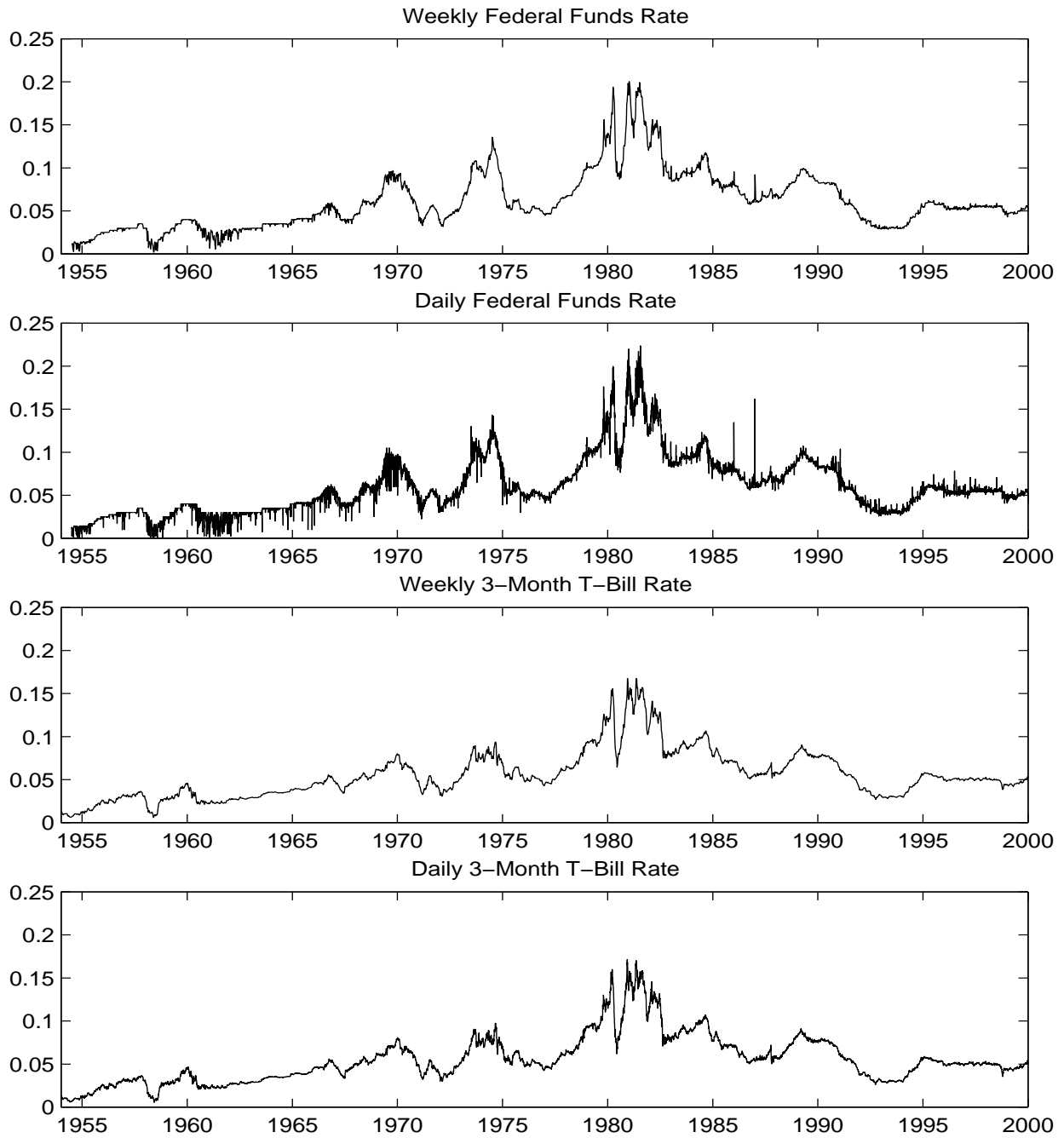


Figure 1: Time Series Plots of Interest-Rate Data.

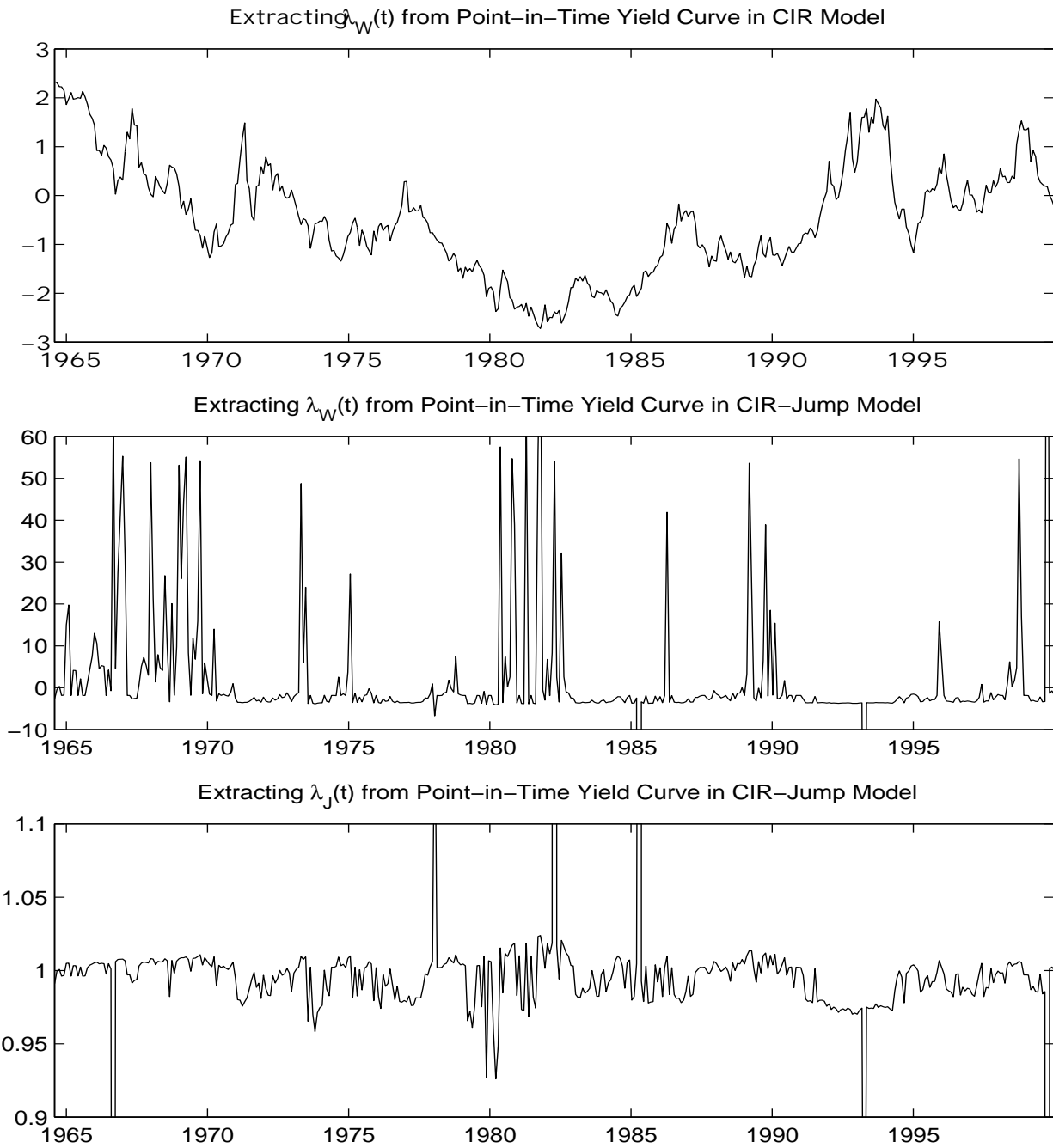


Figure 2: Time Series Plots of Extracted Risk Premia