THE JOURNAL OF FINANCE • VOL. LIX, NO. 1 • FEBRUARY 2004

The Statistical and Economic Role of Jumps in Continuous-Time Interest Rate Models

MICHAEL JOHANNES*

ABSTRACT

This paper analyzes the role of jumps in continuous-time short rate models. I first develop a test to detect jump-induced misspecification and, using Treasury bill rates, find evidence for the presence of jumps. Second, I specify and estimate a nonparametric jump-diffusion model. Results indicate that jumps play an important statistical role. Estimates of jump times and sizes indicate that unexpected news about the macroeconomy generates the jumps. Finally, I investigate the pricing implications of jumps. Jumps generally have a minor impact on yields, but they are important for pricing interest rate options.

THERE IS STRONG ANECDOTAL EVIDENCE that jumps play an important role in determining the dynamics of interest rate movements, although most models are based on diffusions and explicitly rule out their possibility. For example, Figure 1 displays the daily changes in the 3-month Treasury bill (T-bill) rate from 1991 to 1993, a period during which the short rate fell from 7 to just more than 3 percent; of note are the relatively infrequent but large spikes, interpreted here as "jumps."

Given the magnitude of the jumps, the obvious question is: what generated these movements? It turns out that each of the large movements in Figure 1 coincided with the arrival to the Treasury market of significant information regarding the current or future state of the economy. This observation also holds over other time periods during which news about the macroeconomy generated the largest movements in yields.[1] Thus, at least anecdotally, jumps appear to be an important conduit through which macroeconomic information enters the term structure.

*Johannes is with the Finance and Economics Division, Graduate School of Business, Columbia University. This is a revised version of my Ph.D. dissertation at the University of Chicago. I thank Torben Andersen, Federico Bandi, John Cook, Pete Kyle, Nick Polson, Pietro Veronesi, Brian Viard, and seminar participants at Cirano/University of Montreal, Columbia, Duke, Northwestern, Chicago, and the Board of Governors of the Federal Reserve for their comments. I would especially like to thank my dissertation committee: Lars Hansen, John Cochrane, and José Scheinkman. The comments of the Editor and the anonymous referees led to significant improvements. All errors are mine.

[1] Fleming and Remolona (1997), Balduzzi, Elton, and Green (2001) and Furfine (2001) find that as many as 20 different macroeconomic announcements significantly impact yield changes in the Treasury market and, moreover, the largest yield movements are generated by these macroeconomic announcements.

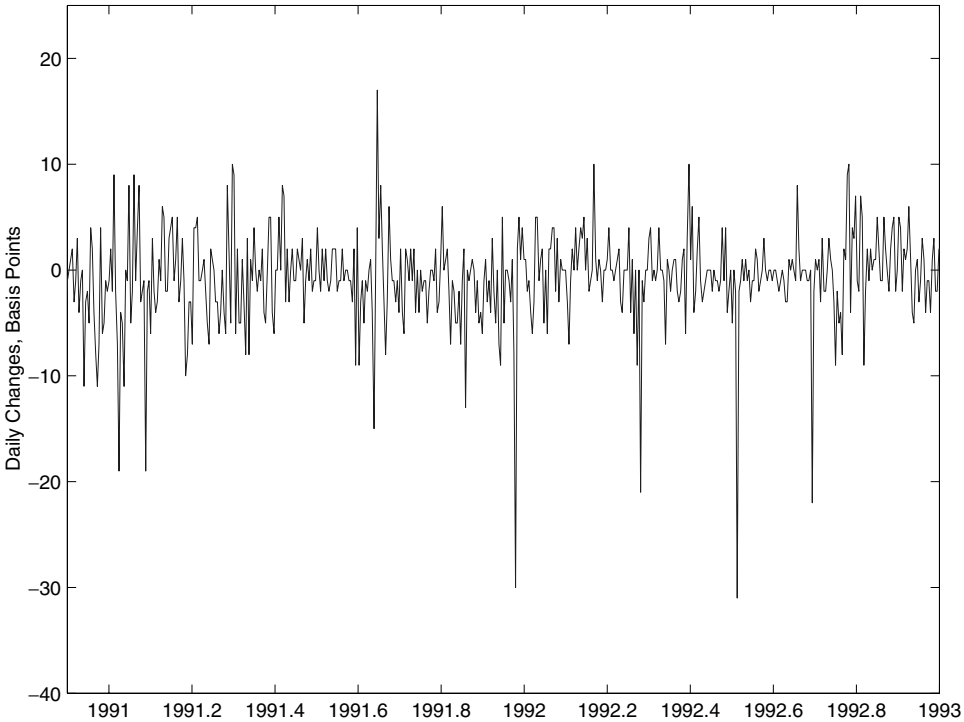

Figure 1. Daily changes in basis points in the 3-month Treasury bill rate from 1991 to 1993.

In this paper, I examine the statistical and economic role of jumps in continuous-time interest rate models. Statistically, the presence of jumps implies that diffusion models are misspecified. I develop a new procedure to test for jump-induced misspecification and results provide strong evidence for the presence of jumps. Next, I propose and estimate a flexible jump-diffusion model to quantify the statistical role of jumps in interest rates. Finally, to gauge economic impact, I analyze the connection between jumps and macroeconomic news arrivals, and I also explore how jumps affect the pricing of bonds and some simple interest rate derivatives.

To test for the presence of jumps, I compare the unconditional and conditional nonnormalities in interest rates with analogs generated by candidate diffusion models.[2] I compute the finite sample distribution of these statistics under the null using bootstrapping or Monte Carlo methods. As benchmark or null models, I consider the nonparametric single-factor model, which has received considerable attention in the literature, and the multifactor models of Andersen and Lund (1997, 1998) that incorporate time-varying central tendency and stochastic volatility.

[2] This provides an alternative to the test developed in Aït-Sahalia (2002). The test developed here depends on a given null model while Aït-Sahalia's test is model independent.

Results indicate that none of the diffusion models can generate nonnormalities consistent with those of observed Treasury rates. Andersen and Lund (1997) reach a similar conclusion suggesting that it "remains difficult to replicate the fat-tailed, or non-Gaussian, innovations of the conditional distribution" (p. 19). While it may be possible that extra diffusive factors or more flexible factor dynamics can remedy this misspecification, recent work indicates that this is unlikely, at least with the models currently in use. For example, Ahn, Dittmar, and Gallant (2002) analyze a battery of three-factor affine and quadratic models and conclude that none of the models "are able to capture the ARCH and non-Gaussian features of the observed data" (p. 275). Moreover, none of the models they study pass their omnibus chi-square test, a measure of overall fit.

Why do these diffusion models fail? Diffusion models induce interest rate increments that are approximately normal over short time intervals while the increments of actual data are very nonnormal; thus, it is difficult for diffusion models to fit the observed data. Huang (1985) provides an alternative, more intuitive interpretation. He notes that the information structures in diffusion models are generated by Brownian motions and these filtrations have the property that "no events can take us by surprise" (Huang (1985), p. 60). Since jumps are precisely the events that take market participants by surprise, diffusion models are misspecified.

These results, combined with the anecdotal evidence presented earlier linking jumps and macroeconomic events, point to the importance for formally estimating models with jumps. Toward this end, I introduce a simple yet flexible model of the short rate that incorporates the lessons of the single-factor diffusion literature, while also allowing for jumps.[3] I develop and justify an estimation procedure that provides nonparametric estimates of the drift, diffusion, and jump intensity, as well as the parameters of the jump distribution. This procedure extends Stanton's (1997) estimator to multivariate jump-diffusions, and the asymptotic properties are derived in Bandi and Nguyen (2003). Due to concerns regarding potentially inaccurate inferences resulting from the use of approximations to asymptotic distributions, I rely on Monte Carlo simulations that provide finite sample confidence bands.

Estimation results indicate that jumps play a dominant role in interest rate dynamics. At low rates, jumps generate more than half of the conditional variance of interest rate changes; this proportion increases to almost two-thirds at high rates. These results are consistent with the intuition that follows from Figure 1, which shows that typical diffusive movements are small (as a proportion of variance) and that the infrequent but large movements dominate. The probability of a jump on a given day is about 6 percent at low rates and more than 20 percent at high rates, with a three standard deviation jump-move corresponding to approximately 50 basis points at low rates and as much as

[3] Ahn and Gao (1999), Aït-Sahalia (1996a, 1996b), Stanton (1997), Jiang (1998), and Bandi (2000) all find that the diffusion coefficient is a nonlinear function of the short rate. The evidence for nonlinearities in the drift is more tenuous as there is some question regarding the reliability of statistical methods for testing nonlinearity in the drift or diffusion (see, e.g., Pritsker (1998), Chapman and Pearson (2000), and Jones (2003)).

150 basis points at high rates. The jump-diffusion model easily generates patterns of conditional and unconditional kurtosis that are consistent with the data. Together, these results imply that jumps play an important statistical role in interest rate dynamics.

To examine the economic implications of jumps in interest rates, I first consider the connection between model-implied jumps and macroeconomic shocks. To do this, I estimate jump times and jump sizes and examine the periods from 1991 to 1993 and 1979 to 1982 in detail. In the 1991 to 1993 period, the filter identifies all of the large moves in Figure 1 as jumps. Moreover, each of these moves coincided with unexpected macroeconomic news arrivals such as surprise Federal Reserve target changes, unemployment announcements, the Soviet coup, the outbreak of the Gulf War, and the 1992 Bush–Clinton presidential debates. Similar results hold for the 1979 to 1982 period, although the particular news items that coincided with the large moves were naturally different.

These results are the first to provide model-based evidence that jumps are a primary conduit through which information about the macroeconomy enters the term structure. While jumps are interesting in their own right, they pose new modeling challenges. For example, the evidence indicates a high-frequency relation between macroeconomic variables and yields, in contrast to the lower-frequency relations typically modeled in the macroeconomic literature (see, e.g., Ang and Piazzesi (2003) who analyze interactions at the monthly frequency). These results are also related to Piazzesi (2001, 2003) who develops term structure models that embed jump processes triggered by macroeconomic announcements. In the Piazzesi models, jumps occur only on days when macroeconomic variables such as unemployment or inflation are announced. Our results, on the other hand, indicate that jumps occur only when the announcements contain significant unexpected components. That is, it is not the announcement that matters per se, but rather the surprise component.

Finally, I examine the pricing implications of jumps in interest rates. First, I find that unless investors place substantial risk premiums on jump components, jumps have little impact on the yield curve. This is not surprising as the yield curve is a static, cross-sectional entity that depends on the distribution of the average interest rate. Due to this fact, jump-diffusion and diffusion models with the same conditional mean and variance will generate similar yield curves. Does this imply that jumps are irrelevant for pricing? Not necessarily. I also compute the price of call options on the 3-month T-bill rate in diffusion and jump-diffusion models with the same instantaneous first and second moments. For short maturities, the pricing difference can be quite large for out-of-the-money options, which implies that jumps in interest rates have a similar impact as jumps in equity index prices. In both markets, the impact of jumps is most clearly seen in higher-moment departures from conditional normality, and the effect appears in drastically different prices for out-of-the-money options (see, e.g., Bakshi, Cao, and Chen (1997) or Pan (2002)).

This paper is related to a number of other papers that study the impact of jumps in interest rates. On the theoretical side, Ahn and Thompson (1985),

Duffie and Kan (1996), Baz and Das (1996), Duffie, Pan, and Singleton (2000), and Chacko and Das (2002) generate closed- or near-closed-form formulas for bond and bond option prices in the presence of jumps. On the empirical side, this paper is closely related to Das (2001). Das (2001) incorporates jumps into generalizations of the Vasicek model and finds strong evidence for jumps in the daily federal funds rate. However, as pointed out by Hamilton (1996), the federal funds rate is subject to severe regulation-induced microstructure effects (settlement Wednesdays and the last-day-of-the-quarter/year effects) that spuriously induce jumps in the daily federal funds rates.[4]

The following section, Section I, develops and implements diagnostics to detect when diffusion models are misspecified due to the presence of jumps. Section II develops a jump-diffusion model of the short rate. Section III develops a nonparametric estimation procedure and provides estimation results. Section IV considers the economic implications of jumps and Section V concludes.

I. Jumps and Diffusion Models

The main difference between models with and without jumps is the continuity of the sample path. By construction, diffusion models are continuous functions of time whereas jump models have occasional discontinuities. If the short rate were observed continuously through time, $\{r_t\}_{t\geq 0}$, then jump sizes would also be observed as discontinuities: at a jump time τ, the jump size, Z_τ, would be $r_\tau = r_{\tau-} + Z_\tau$, where $r_{t-} = \lim_{s\uparrow t} r_s$. In practice, observations are available only discretely and the above limit cannot be computed. This implies that statistical metrics based on discretely observed interest rates must be used to detect jumps.

To test for the presence of jumps, I use the information contained in the unconditional and conditional distributions of interest rate increments, $p(r_{t+\Delta} - r_t)$ and $p(r_{t+\Delta} - r_t \mid r_t)$, respectively. The true continuous-time model generates these distributions over an observation interval of length Δ as the solution to a stochastic differential equation. The idea is simple: any given single-factor or multifactor diffusion model induces a distribution of increments, and it is straightforward to compare the distributional properties of the model with those of the data.

The following discrete-time example provides the intuition for the tests. Suppose that interest rates evolve according to

$$r_{t+\Delta} - r_t = \mu(r_t)\Delta + \sigma(r_t)\varepsilon_{t+\Delta} + J_{t+\Delta}Z_{t+\Delta}, \tag{1}$$

where $J_t = 1$ (with probability $\lambda\Delta$) indicates the arrival of a jump, $\varepsilon_t \sim N(0, \Delta)$ and $Z_t \sim N(0, \sigma_z^2)$. In this model, the conditional mean is $\mu(r_t)\Delta$ and the conditional variance is $\sigma^2(r_t)\Delta + \lambda\Delta\sigma_z^2$. Based only on the first two moments, the

[4] For example, the microstructure effects often induce daily changes of more than 100 basis points and, as Hamilton notes, some are as large as 800 basis points.

Table I
Summary Statistics

Summary statistics for daily Treasury bill rates from January 1965 to February 1999.

	Mean	Standard Deviation	Skewness	Kurtosis	First Autocorrelation
r_t	6.47	2.57	1.41	5.26	0.99
$r_{t+\Delta} - r_t$	7.5×10^{-5}	0.11	0.20	24.30	0.14

jump model is observationally equivalent to a model without jumps with a drift coefficient of $\mu(r_t)\Delta$ and a diffusion coefficient equal to $\sigma^2(r_t)\Delta + \lambda\Delta\sigma_z^2$. This suggests that to test for the presence of jumps, one needs to analyze higher-moment departures from normality in $p(r_{t+\Delta} - r_t)$ and $p(r_{t+\Delta} - r_t \mid r_t)$. These departures from normality are commonly measured by the skewness and kurtosis statistics.

While it is easy to compute statistics such as the unconditional or conditional kurtosis, it is difficult to use such statistics for formal hypothesis testing of diffusion models. The reason is as follows. Suppose that interest rate increments are very nonnormal, as reflected by the kurtosis statistic (Table I indicates that the kurtosis of daily 3-month T-bill rate increments is more than 24).[5] A standard test for nonnormalities would compare the kurtosis statistic to its value under the null hypothesis that the data is normally distributed. This is not valid in our setting because diffusion models induce increments that are generally not normally distributed. For example, in the Cox, Ingersoll, and Ross (1985) model,

$$r_{t+\Delta} - r_t = \int_t^{t+\Delta} \kappa_r(\theta_r - r_s)\,ds + \int_t^{t+\Delta} \sigma_r \sqrt{r_s}\,dW_s, \tag{2}$$

where W_s is a scalar Brownian motion, κ_r is the speed of mean reversion, θ_r is the long-run mean, and σ_r is the volatility. Due to the randomization of the Brownian increment by $\sqrt{r_s}$, the stochastic integral $\int_t^{t+\Delta} \sigma_r \sqrt{r_s}\,dW_s$ is not normally distributed (it is a scale mixture of normals). This implies that the unconditional and conditional distributions of $r_{t+\Delta} - r_t$ are fat-tailed relative to a normal distribution.

To test diffusion models based on measures of nonnormalities, we need to compute the distribution of the test statistic under the null hypothesis of a given diffusion model. To do this, I use Monte Carlo hypothesis testing, or, bootstrapping. This procedure, which has a long history in statistics, is typically used in two settings: when the distribution of the test statistic is unknown,

[5] This result is not an artifact of the short end of the Treasury curve, as longer maturity Treasury and Libor rates also exhibit extremely high kurtosis. For example, daily changes in 3-month Libor from 1991 to 2000 has a kurtosis of 19 (similar for the period from 1970 to 2000) and the daily changes in the 5-year, constant-maturity Treasury yield from 1962 to 2000 (the longest period for which the Treasury data is available) has a kurtosis of nearly 16.

or, when there is reason to believe finite sample approximations to limiting distributions are inaccurate.[6] Here, both of these concerns are relevant.

To test for a given model's ability to generate unconditional or conditional nonnormalities consistent with observed data, I follow Conley et al. (1997) and Pritsker (1998) and implement the following procedure:

1. Estimate the statistic from a sample of interest rate data. For example, with T observations, equally spaced at length Δ, $\{r_1, \ldots, r_T\}$, compute the unconditional kurtosis coefficient, $\hat{k}$;
2. For a given diffusion model, simulate a large number, G, of sample paths from the true continuous-time model $\{r_1^g, \ldots, r_T^g\}_{g=1}^G$;
3. For each sample path, recompute the test statistic. This produces a sample of size G, $\{\hat{k}^{(g)}\}_{g=1}^G$, from the distribution of the statistic conditional on a given T and Δ; and,
4. Use the quantiles of the empirical distribution to obtain the critical values of the test statistic under the null, and compare the statistic estimated from observed data to the critical values.

Provided G is large enough, this procedure provides an exact finite sample hypothesis test based on the null of a given model. I use $G = 1{,}000$, which is twice as many simulations as used by Conley et al. (1997) and Pritsker (1998), and which guarantees that we have a reliable picture of the finite sample distribution of the test statistics.

To measure departures from normality, I compute unconditional kurtosis statistics over various time horizons (daily, weekly, and monthly) and the daily conditional kurtosis. I also compute the corresponding skewness statistics, but these have no ability to discriminate across models as I cannot reject any of the models at conventional significance levels. This should not be a surprise as there is little noticeable asymmetry in the distribution of T-bill rate changes.

The conditional kurtosis over a fixed interval of length Δ is defined as the fourth conditional moment divided by the second conditional moment squared:

$$k_\Delta(r) \triangleq \frac{E\left[(r_{t+\Delta} - r_t)^4 \,\middle|\, r_t = r\right]}{\left(E\left[(r_{t+\Delta} - r_t)^2 \,\middle|\, r_t = r\right]\right)^2}. \tag{3}$$

Given that $k_\Delta(r)$ is a function of the short rate level, hypothesis testing is more difficult in this case. The reason is that Monte Carlo simulations provide the pointwise distribution of $k_\Delta(r)$ under the null and it is not valid to simply compare two functions point-by-point. To construct a formal test, I follow Härdle and Mammen (1993) and consider integrated metrics.[7] Specifically, I estimate the integrated conditional kurtosis, $\int k_\Delta(r)\,dr$, and the integrated squared kurtosis, $\int k_\Delta^2(r)\,dr$, which provide two metrics of total conditional kurtosis. I also compute the density-weighted statistics, $\int k_\Delta(r)p(r)\,dr$ and $\int k_\Delta(r)^2 p(r)\,dr$, where

[6] These tests are typically attributed to Barnard (1963).

[7] I would like to thank a referee for suggesting this procedure.

$p(r)$ is the marginal distribution of r. However, they gave the same result as the Lebesgue-measure weighted estimates and do not warrant separate reporting.

There are two main advantages to the above testing procedure. First, it does not rely on $\Delta \to 0$ or $T \to \infty$ asymptotics, nor on the approximation of the conditional distribution by a normal distribution, as is the case in an Euler discretization in which the discretization interval matches the observation frequency. Rather, the testing procedure merely compares the tail behavior of interest rate increments with the increments generated by a candidate continuous-time diffusion model. Second, provided G is large enough, this procedure provides exact finite sample hypothesis testing.

Aït-Sahalia (2002) develops an alternative test to detect jumps, which relies on a positivity condition of derivatives of the conditional density. The approach is unlikely to work in an interest rate setting due to difficulties that arise in estimating the conditional density—unlike the equity option application in Aït-Sahalia (2002), the sampling error in estimating interest rate densities is typically quite high (see, e.g., Conley et al. (1997) and Pritsker (1998)).

A. *Diffusion Benchmarks*

The previous section discussed a general methodology for testing the ability of a diffusion model to generate tail behavior consistent with observed interest rate increments. This section describes the benchmark diffusion models used in formal testing. I first consider the popular nonparametric single-factor model

$$dr_t = \mu(r_t)\,dt + \sigma(r_t)\,dW_t, \tag{4}$$

where $\mu(\cdot)$ and $\sigma^2(\cdot)$ are flexible unknown functions of r_t. The nonparametric and other nested parametric models have been advocated in a number of recent papers (see, e.g., Ahn and Gao (1999), Aït-Sahalia (1996a, 1996b), Bandi (2000), Chan et al. (1993), Jiang (1998), and Stanton (1997)). Many of these papers find nonlinearities in the diffusion, which result in a greater randomization of the Brownian increment and thus a greater potential to generate nonnormalities. The model in equation (4) nests all single-factor models in which the spot rate is an invertible function of a state variable. In this regard, it does not nest certain quadratic models; I will discuss this latter set of models below.

As a multifactor benchmark, I consider the stochastic volatility model of Andersen and Lund (1997, 1998),

$$\begin{aligned} dr_t &= \kappa_r(\mu_r - r_t)\,dt + r_t^{\gamma}\sqrt{V_t}\,dW_t^1 \\ d\log(V_t) &= \kappa_v(\mu_v - \log(V_t))\,dt + \sigma_v\,dW_t^2, \end{aligned} \tag{5}$$

where the two Brownian motions are uncorrelated. I also implement Andersen and Lund's (1997) three-factor model with a time-varying central tendency factor. However, as these authors note, this additional factor provides no help

in generating nonnormalities and the results of the three-factor model are the same as those of the two-factor model.

The stochastic volatility model is an important benchmark as time-varying volatility is a popular method to generate departures from conditional normality. The stochastic volatility model also allows for a nonlinear interest rate level effect, r_t^{γ}, in the diffusion function, and for nonlinear mean reversion in the drift of the volatility process.[8] An additional motivation to use the Andersen and Lund specification is the fact that they model the 3-month T-bill rate, the rate that is often used in the analysis of single-factor models (e.g., Stanton (1997)). It is important to note that there exist many other multifactor models, especially in the affine and quadratic classes. I discuss the performance of these models below.

To implement the specification tests, I need to estimate the model parameters and nonnormality statistics. For the multifactor models, I use Andersen and Lund's (1997) parameter estimates. For the single-factor model, I use the first-order nonparametric estimation scheme of Stanton (1997) whose asymptotic properties are derived in Bandi and Phillips (2002) and Bandi (2000). The estimators are given by

$$\frac{1}{\Delta}\hat{E}\left[(r_{t+\Delta}-r_t)^j \mid r_t = r\right] = \frac{\sum_{i=1}^{T} K\left(\frac{r_{i\Delta}-r}{h_j}\right)\frac{(r_{(i+1)\Delta}-r_{i\Delta})^j}{\Delta}}{\sum_{i=1}^{T} K\left(\frac{r_{i\Delta}-r}{h_j}\right)} = \begin{cases} \hat{\mu}(r) : j = 1 \\ \hat{\sigma}^2(r) : j = 2, \end{cases} \tag{6}$$

where $\{r_{i\Delta}\}_{i=1}^{T}$ are the observed increments, $K(\cdot)$ is a Gaussian kernel, and (h_1, h_2) are bandwidths. Given the moment estimates, forming the appropriate ratios of kernel estimates of the conditional moments provides estimates of the conditional kurtosis.

Nonparametric estimators have a number of advantages. First, they require little prior information relating to the functional form of the conditional expectations. Asymptotically, the nonparametric estimates will recover the true conditional expectation function if it is a smooth (continuous and differentiable) function of the underlying state variable (see Bandi (2000) for formal conditions). Second, nonparametric estimators are local. This implies that conclusions about high interest rate environments depend very little on data from low interest rate environments and vice versa. This implies that removing part of the sample, for example, the volatile 1979 to 1982 period, does not substantively change any of the conclusions. Finally, kernel estimators are computationally easy to evaluate and thus it is feasible to construct Monte Carlo confidence bands.

To implement the kernel estimates, the bandwidth parameters must be chosen. When the data are independent and identically distributed there are theoretically optimal bandwidth choices, but there are no related results for data generated from diffusion models. Bandwidth selection is important, however,

[8] The process violates the linear growth condition. As shown by Downing (2000), the model does not have any degeneracies and a unique solution appears to exist.

for identifying nonlinearities in the drift function, but this is not our objective and thus it is less of a concern here. Following Chapman and Pearson (2000), I allow the drift and diffusion bandwidths to differ and choose each bandwidth using the following internally consistent procedure. For given bandwidths h_1 and h_2, equation (3) provides the estimates of the drift and diffusion. Given these estimates, I simulate sample paths from the implied diffusion model and then, using the same bandwidth choices, reestimate the drift and diffusion for each simulated path.[9] I consider bandwidth combinations for the drift and diffusion coefficients of the form $h = c \times \hat{s}$, where $\hat{s}$ is the estimated standard deviation of the sample path, and c varies from 0.3 to 1.5 in increments of 0.05. Finally, I report estimates based on the bandwidths that result in the smallest finite sample bias.

To implement the Monte Carlo procedure, I generate $G = 1{,}000$ Monte Carlo sample paths. To match the observed frequency between observations, the diffusion must be sampled at a fixed interval $\Delta = 1/252$. This is achieved by simulating paths at intervals of length $\Delta/5$, and sampling every fifth point to reduce any discretization bias. The median across the simulations and the quantiles of the simulated drift, diffusion, skewness, and kurtosis summarizes the Monte Carlo simulation. I average simulations across their empirical support and perform no extrapolation, to avoid drawing conclusions over regions outside those reached by the simulations.

B. Empirical Results

I use secondary market quotes for the 3-month T-bill to estimate and test the models. The 3-month T-bill has a number of advantages over other short maturity interest rates. These bills are very liquid, have small bid–ask spreads, and are free of idiosyncratic effects (Duffee (1996) and Fleming and Sarkar (1999)) that could spuriously induce nonnormalities. In using 3-month T-bill data, I am directly modeling the yield as opposed to the instantaneous spot rate. As noted by Honore (1998) and Chapman, Long, and Pearson (1999), there is an important difference between the instantaneous spot rate and the yield on a short maturity instrument. These authors show that estimates of the drift and diffusion may be biased in parametric models when yields proxy the short rate. This is not a concern at this stage as I am not invoking a parametric term structure model.[10]

Table I summarizes and Figure 2 plots the data. With daily data, it is important to document the effects induced by omitting weekends and holidays. To measure the impact of omitting pairs of observations split by weekends or

[9] For certain simulated sample paths (those whose maximum value is relatively small), the estimates of the second moment are approximately zero for large values of the state. To avoid the numerical problems of dividing by numbers close to zero, in all simulations I add a small number (10^{-15}) to the moment estimates.

[10] Provided the yield is an invertible and twice-differentiable function of the spot rate, a single-factor diffusion model for the yield is fully consistent with a single-factor model for the spot rate. To see this, assume $dr_t = \mu(r_t)\,dt + \sigma(r_t)\,dW_t$. Ito's lemma implies that the yield y_t is given by $dy_t = \mu_y(y_t)\,dt + \sigma_y(y_t)\,dW_t$ for some drift and diffusion if the spot rate can be inverted from the yields. Since I do not impose a parametric model on the drift and diffusion, there is no bias.

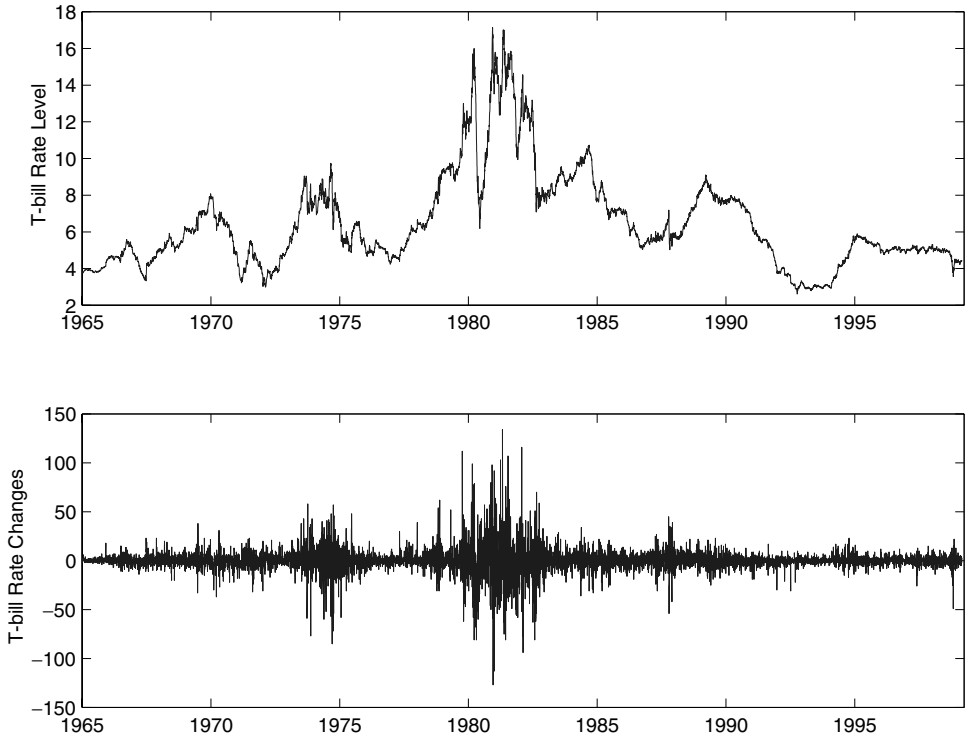

Figure 2. The time series of the level and changes (in basis points) in the 3-month Treasury bill rate, January 1965 to February 1999.

holidays, I remove these observations which leave, ignoring holidays, Monday/Tuesday, Tuesday/Wednesday, Wednesday/Thursday, and Thursday/Friday observation pairs. The original and reduced data sets are virtually identical in terms of their statistical properties, and none of the conclusions regarding either the shortcomings of the diffusion models or the qualitative structure of the jump processes change. I also considered the effects of rounding (Treasury yields are rounded to a given basis point); all conclusions are robust to rounding.[11]

B.1. The Single-Factor Diffusion Model

Drift and diffusion estimates for the single-factor model and their Monte Carlo confidence bands are given in the top panels of Figure 3.[12] I report estimates for $r_t \in [0.029, 0.16]$, which cover the (0.5, 99.5) quantiles of the data.

[11] I thank a referee for suggesting this experiment.

[12] The bandwidths used for the drift and diffusion are $h_1 = 1.25 \times \hat{s}$ and $h_2 = 0.4 \times \hat{s}$. These choices are consistent with the findings of both Chapman and Pearson (2000) and Bandi and Nguyen (1999) who recommend oversmoothing the drift relative to the diffusion. The higher moments are estimated using the same bandwidth as the second moment.

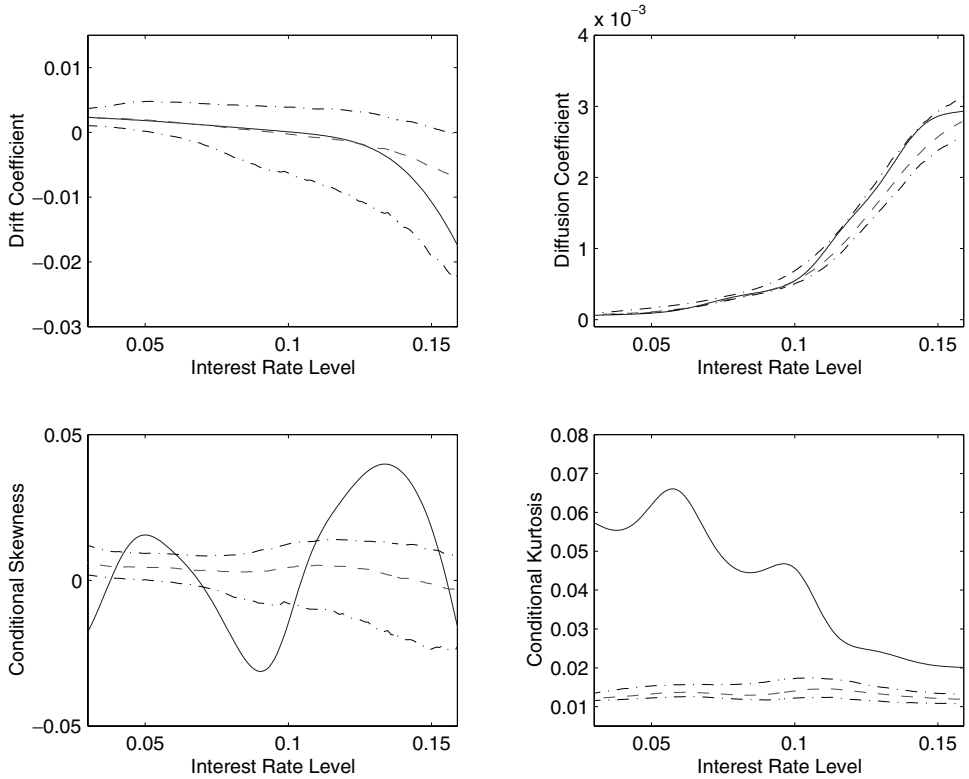

Figure 3. Nonparametric estimation results for the single-factor diffusion model. The solid line is the function estimated from short rate data, the dashed line is the Monte Carlo median, and the dash–dot lines are the (10%, 90%) Monte Carlo confidence bands.

The simulation results indicate that the estimates are unbiased, at least for rates less than 12 percent. Since there are few observations at high rates, the relatively wide confidence intervals should not be surprising.

The diffusion coefficient estimates are more precise than the drift coefficient estimates. Bandi and Phillips (2002) provide an explanation for this result. They prove that consistent estimation of the drift requires a long time span (large T) and a high sampling frequency (small Δ). The results in this study imply that daily data are a short-enough sampling frequency to estimate the diffusion, but 34 years are not a long-enough time span to estimate the drift. There is an additional subtle point: unlike Chapman and Pearson (2000), who document that there is significant bias in drift estimates at higher interest rates, I do not find this result.[13]

[13] A potential explanation for this is that I did *not* extrapolate off the support of the simulated sample paths when computing the Monte Carlo estimates. Extrapolation could cause the sort of bias reported in Chapman and Pearson (2000). I would like to thank Matt Pritsker for extensive conversations regarding the patterns of bias in diffusion models.

Table II
Statistical Tests

Panel A provides the conditional and unconditional kurtosis statistics from the observed data. Panels B, C, and D provide the quantiles of the conditional and unconditional kurtosis statistics under the null of the nonparametric single-factor model, the two-factor stochastic volatility model of Andersen and Lund (1997), and the nonparametric jump-diffusion model, respectively.

	Unconditional Kurtosis			Conditional Kurtosis	
	Daily	Weekly	Monthly	Absolute	Squared
		Panel A: 3-Month Treasury Bill Rate Data			
	24.3015	17.1501	16.1618	1.9459	0.0896
		Panel B: Nonparametric Diffusion Model			
50%	6.7860	6.5015	6.1384	0.4403	0.0057
75%	8.8370	8.7195	8.1935	0.5889	0.0080
90%	10.9935	11.1683	11.1018	0.6623	0.0096
95%	12.6059	12.6580	12.9536	0.6969	0.0106
99%	15.4791	15.7653	17.7360	0.7790	0.0131
		Panel C: Stochastic Volatility Model			
50%	6.8121	6.4437	5.5203	0.5903	0.0138
75%	7.7988	7.5810	6.7470	0.7046	0.0180
90%	9.3758	9.0718	8.7403	0.8300	0.0235
95%	10.9601	11.4194	10.5507	0.9455	0.0287
99%	16.5754	17.2357	17.0542	1.2712	0.0576
		Panel D: Jump-Diffusion Model			
50%	22.9943	10.1425	7.3117	1.8406	0.0705
75%	28.4758	12.5991	9.6654	2.1455	0.1300
90%	35.1220	16.5203	13.0704	2.4656	0.1651
95%	39.8182	20.1456	15.3049	2.7073	0.1929
99%	46.6100	23.2665	20.0280	3.1219	0.2748

The single-factor diffusion model is misspecified. Table II summarizes the nonnormality tests. Panel A provides the sample statistics and Panel B provides quantiles of the finite sample distribution of the test statistics under the null of the nonparametric diffusion model. By all metrics, the model is unable to generate kurtosis close to that in the observed data. For example, at a daily level, the median of the kurtosis distribution is less than 7 and the 99th percentile is 15.5, compared to a kurtosis statistic in the sample of more that 24. The conditional tests (which should have more power) provide an even stronger rejection.

Graphically, the inability of the model to generate realistic amounts of conditional kurtosis can be seen in the bottom-right panel in Figure 3. The conditional kurtosis is about 10 to 20 percent smaller than its sample counterpart over much of the support of the data. Figure 3 indicates that the diffusion model comes closer to matching the conditional kurtosis at high interest rates.

The reasons for this are intuitive: the local variance of the process, $\sigma^2(r)$, is more than 20 times larger at high rates than at low rates. This nonlinearity creates additional randomization of the Brownian increment at the high rates that generates greater conditional kurtosis. The model has the greatest difficulty generating nonnormalities at low rates where $\sigma^2(r)$ is roughly constant and interest rate increments are approximately a random walk. The conditional skewness plot in the lower-left panel is included for completeness, but, again, provides little additional insight.[14] These conclusions do not depend on the bandwidth parameters as similar patterns of kurtosis occur for all of the possible bandwidth choices, although the smoothness of the estimates does vary. In the limit, as the bandwidth increases, the nonparametric estimates converge to the unconditional kurtosis, which yields the same conclusion.

B.2. Multi-factor Models

This section repeats the exercise of the previous section using the stochastic volatility model and parameter estimates of Andersen and Lund (1997). The conclusions of this experiment are similar to those in the single-factor diffusion model. Panel C of Table II indicates that the stochastic volatility model generates roughly the same amount of unconditional kurtosis as the nonparametric diffusion model, and is therefore rejected. Moreover, while the stochastic volatility model does generate more conditional kurtosis than the nonparametric diffusion model, it is still strongly rejected by the formal tests. The same conclusions hold for the three-factor model of Andersen and Lund (1998), which adds a stochastic central tendency factor.

There is one further point that needs to be addressed. Andersen and Lund (1997, 1998) use weekly data to estimate their models, whereas I use daily data. It is safe to conclude that the misspecification they find with regard to fitting the nonnormalities would be more pronounced if they used daily data for two reasons. First, as indicated by Table II, the nonnormalities in the short rate are more pronounced in daily data than in weekly data. Second, diffusion models become more "normal," in a distributional sense, over daily intervals than weekly (see, e.g., Duffie and Pan (1997) and Das and Sundaram (1999)). Thus, with the change to daily frequencies, the data become less normal while the models become more normal, which generates even greater misspecification.

C. The Shortcoming of Diffusion Models

The results indicate that the single-factor and certain multifactor diffusion models do not generate enough kurtosis, conditional or unconditional, to match estimates from short rate data. A number of points are in order. First, the single-factor diffusion model can fit the first two conditional moments of the short

[14] In the jump-diffusion model given later, the error bands on the conditional skewness are much wider and contain the sample estimates.

rate but cannot fit the higher moments as there are no more degrees of freedom available. Second, it is not the case that there does not exist a multifactor diffusion model that can match the conditional or unconditional kurtosis. For example, increasing σ_v in Andersen and Lund's model by 75 percent increases the kurtosis to levels close to those of the T-bill data. However, the cost of this modification is severe: with the higher σ_v, the maximum of many of the simulated sample paths is greater than 60 percent, and the volatility of daily movements is more than double that in the data.

It appears that none of the multifactor diffusion models that have been estimated to date can fit the observed nonnormalities in interest rates. Ahn et al. (2002) analyze numerous three-factor affine and quadratic models and find that none of them can fit the tails of the distribution. Dai and Singleton (2000) find further that a number of three-factor affine models cannot fit the nonnormalities in swap rates as indicated by their specification tests.

One way to see the shortcoming of diffusion models is to approximate the conditional kurtosis using stochastic Taylor series expansions.[15] The idea, based on Mihlstein (1974), is to expand conditional moments,

$$E[f(r_{t+\Delta})\,|\,r_t = r] = f(r_t) + \mathcal{L}f(r_t)\Delta + \tfrac{1}{2}\mathcal{L}^2 f(r_t)\Delta^2 + \tfrac{1}{6}\mathcal{L}^3 f(r_t)\Delta^3 + O(\Delta^4), \quad (7)$$

where $\mathcal{L}$ is the generator of r_t applied to a given function f in its domain. Applying a second-order expansion to a univariate diffusion,

$$E\big[(r_{t+\Delta} - r_t)^2 \,\big|\, r_t = r\big] = \sigma^2(r)\Delta + O(\Delta^2) \quad (8)$$

and

$$E\big[(r_{t+\Delta} - r_t)^4 \,\big|\, r_t = r\big] = 3\sigma^4(r)\Delta^2 + O(\Delta^3), \quad (9)$$

which implies that, to order Δ, the conditional kurtosis over a Δ-interval of a single-factor diffusion is approximately 3; thus, deviations from normality are necessarily small.

Next, consider an affine stochastic volatility model,

$$dr_t = \kappa_r(\theta_r - r_t)\,dt + \sqrt{V_t}\,dW_t^r \quad (10)$$

and

$$dV_t = \kappa_v(\theta_v - V_t)\,dt + \sigma_v\sqrt{V_t}\,dW_t^v, \quad (11)$$

and assume the Brownian motions are uncorrelated. In this case, the second-order expansions imply that

$$E\big[(r_{t+\Delta} - r_t)^2 \,\big|\, r_t = r\big] = E(V_t)\Delta + O(\Delta^2)$$

[15] I would like to thank a referee for suggesting this procedure.

and

$$E\left[(r_{t+\Delta} - r_t)^4 \,|\, r_t = r\right] = 3E\left(V_t^2\right)\Delta^2 + O(\Delta^3).$$

Since $E(V_t^2) = \theta_v^2\kappa_v + \frac{\theta_v\sigma_v^2}{2}$ and $E(V_t) = \theta_v$,

$$kurt_\Delta(r) = 3\left(\kappa_v + \frac{\sigma_v^2}{2\theta_v}\right) + O(\Delta^3).$$

Using parameter estimates from Collin-Dufresne, Goldstein, and Jones (2002), the second-order expansion implies that $\text{kurt}_\Delta(r) = 3.34$. A third-order approximation is a very complicated expression as it depends on the level of the short rate. In the case in which the short rate is equal to its long-run average, $r_t = \theta_r$, the third-order expansion implies that $\text{kurt}_\Delta(r) = 3.42$. This should not be a surprise as the short rate is effectively a random walk when $r_t \approx \theta_r$. The kurtosis is slightly greater, $\text{kurt}_\Delta(r) = 3.88$, if we consider the polar case of $r = 0$ as the mean reversion contributes. In a stochastic volatility model, the only way to generate large nonnormalities is to have $E[V_t^2 \,|\, r_t] > (E[V_t \,|\, r_t])^2$, which requires strong contemporaneous interactions between volatility and the spot rate; such interactions are not supported in the data.

The problem with low-dimensional diffusion-based models is that there is not a factor that can move rapidly enough to generate the large but isolated movements often seen in short rate movements. In all of these models, the factor that drives the fourth conditional moment also drives the second moment. Thus, if volatility is high, the fourth moment is high also, and their effects to a large extent cancel.

II. A Jump-Diffusion Model of the Short Rate

Given the inability of popular diffusion models to generate fat tails, I consider a simple generalization incorporating jumps. Suppose the interest rate solves

$$dr_t = \mu(r_t)\,dt + \sigma(r_{t-})\,dW_t + d\left(\sum_{n=1}^{N_t} r_{\tau_n-}(e^{Z_n} - 1)\right), \tag{12}$$

where W_t is a scalar Brownian motion, N_t is doubly stochastic point process with stochastic intensity $\lambda(r_t)$, and $Z_n \sim \mathcal{N}(\mu_z, \sigma_z^2)$ are the marks of the point process that arrive at time τ_n. I assume that μ, σ, λ, and Z have sufficient regularity for the solution to equation (12) to be well defined. The Appendix provides a formal statement of these conditions.

The jump specification in (12) implies that the probability of a jump arriving over a short time interval is approximately $\lambda(r_t)\,dt$. At a jump time τ_n, the jump size induces a discontinuity in the sample path whose size depends on the current spot rate, $r_{\tau_n} = r_{\tau_n-}e^{Z_n}$. Thus, the $\mathcal{F}_{\tau_n-}$ *conditional* distribution of

jump sizes is log-normal and r_t cannot jump negative. The jump size specification and the nonparametric jump intensity provide the flexibility to capture periods such as 1979 to 1982. Since the drift and diffusion coefficients are nonparametric, the model nests the entire single-factor class of diffusion models and also the jump-diffusion models considered by Ahn and Thompson (1985), Duffie and Kan (1996), Baz and Das (1996), Chacko and Das (2002), and Zhou (2001).

The model in (12) implies that jumps occur at random, unpredictable times, which is an implication of both the time-homogeneity and the continuity of the jump intensity. This may at first appear to be at odds with the observation that many of the events that induce jumps occur at regularly scheduled intervals. For example, the Federal Open Market Committee (FOMC) meets every 6 weeks and federal agencies release macroeconomic data at monthly or quarterly frequencies.

Note first that it is not the fact that an announcement occurs that matters, but rather that the announcement contains a large, unanticipated component. Every week there are consensus forecasts for a large number of announcements and only if a reported number is significantly different than anticipated is there a response in the market. This is especially true for Federal Reserve target changes, which are often fully anticipated.

Second, it is not possible to explicitly model all of the regularly scheduled events that cause jumps. The reason is that there are too many announcements that have been shown to be important. Fleming and Remolona (1997) and Balduzzi et al. (2001) find that more than 25 monthly announcements have significant impacts on Treasury yields. Additionally, every 6 weeks the FOMC meets and every quarter other macroeconomic data such as GDP growth are released. Often there are multiple announcements per day since there are only 22 trading days a month, on average. Without ultra-high frequency data (5-minute intervals) it is impossible to disentangle these announcements.

A shortcoming of the model is that, although there are three random factors, there is only one single-state variable. However, my goal is to study the role of jumps and not to find the model with the lowest pricing errors. Thus the model in (12) is useful for a number of purposes. First, as noted in the previous section, additional diffusive factors cannot remove the misspecification in the tails of the distribution, it is important to test whether a parsimonious jump specification can remove the misspecification. Second, it is straightforward to estimate the jump times and sizes in this model, which allows us to connect jump arrivals with macroeconomic events. Finally, principal components analysis of zero-coupon Treasury yields from 1959 to 1998 indicates that a single factor, the short rate, explains about 95 percent of the variation in yields.[16] Thus, from a time-series perspective, the vast majority of interest rate moves are due to a single factor, and, consequently, it is important to understand the dynamics of this dominant factor.

[16] The data used are an updated version of the McColluch–Kwon data covering 18 yields with maturity up to 10 years. I would like to thank Michael Brandt for providing the data.

III. Estimation and the Statistical Role of Jumps

This section provides a general methodology for the nonparametric identification and estimation of jump-diffusion processes. Although the results apply for multivariate jump-diffusions, I consider the univariate model in (12) transformed into logarithms,

$$d\log(r_t) = \mu(r_t)\,dt + \sigma(r_{t-})\,dW_t + d\left(\sum_{n=1}^{N_t} Z_n\right), \tag{13}$$

where I have redefined the drift and diffusion coefficients. For the current application, I assume that the mean jump size is zero. Originally, the parameter was estimated but it was indistinguishable from zero and added significant noise into the estimation of the other components. As Figure 2 shows, the large movements are quite symmetric and thus this assumption is of little consequence.

The first step for estimation is to identify $\mu(r), \sigma^2(r), \lambda(r)$, and σ_z^2. Under regularity conditions given in the Appendix,

$$\lim_{\Delta\downarrow 0} \frac{1}{\Delta} E[\log(r_{t+\Delta}/r_t) \,|\, r_t = r] = \mu(r), \tag{14}$$

$$\lim_{\Delta\downarrow 0} \frac{1}{\Delta} E\left[\log(r_{t+\Delta}/r_t)^2 \,\middle|\, r_t = r\right] = \sigma^2(r) + \lambda(r)E[Z^2], \tag{15}$$

and

$$\lim_{\Delta\downarrow 0} \frac{1}{\Delta} E\left[\log(r_{t+\Delta}/r_t)^j \,\middle|\, r_t = r\right] = \lambda(r)E[Z^j]. \tag{16}$$

These limiting moment conditions extend the well-known instantaneous moment conditions of a diffusion. Specifying the process in logarithms with mean-zero jumps ensures that $\mu(r)$ retains its interpretation as the local mean of the process. It also implies that the jump components $\lambda(r)$ and σ_z^2 can be identified using the fourth and sixth moments, since

$$\lim_{\Delta\downarrow 0} \frac{1}{\Delta} E\left[\log(r_{t+\Delta}/r_t)^4 \,|\, r_t = r\right] = 3\lambda(r)\left(\sigma_z^2\right)^2 \tag{17}$$

and

$$\lim_{\Delta\downarrow 0} \frac{1}{\Delta} E\left[\log(r_{t+\Delta}/r_t)^6 \,|\, r_t = r\right] = 15\lambda(r)\left(\sigma_z^2\right)^3. \tag{18}$$

Once the jump components are identified, the second moment identifies the diffusion coefficient, $\sigma^2(r)$, and the first moment identifies the drift.

In practice, the limiting moment conditions cannot be directly estimated as data are observed at fixed intervals. Instead, I consider an approximate identification scheme. For the drift, I assume that for small Δ,

$$\frac{1}{\Delta} E[\log(r_{t+\Delta}/r_t) \,|\, r_t = r] \approx \mu(r) \tag{19}$$

is a reasonable approximation. In the case of single-factor diffusions, Stanton (1997) argues that for Δ as large as 1 month, the error in the above approximations is negligible. I provide Monte Carlo evidence below that shows that the discretization bias is small in the jump-diffusion model. To estimate the conditional moments, I use nonparametric kernel estimators, as in the diffusion setting. The only problem is that the ratio of the sixth-to-fourth moments is not constant.[17] I estimate the jump variance by integrating the ratio of the sixth-to-fourth moments over the stationary density. Alternatively, one could use the Lebesgue-measure weighted average; the results are substantively the same.

Bandi and Nguyen (2003) derive the limiting distribution under the assumption that the realizations are recurrent and that $\mu(r)$, $\sigma(r)$, and $\lambda(r)$ are smooth functions. I use Monte Carlo simulations to justify the estimation procedure and provide finite sample confidence bands. I use an Euler approximation to generate sample paths and, as in the diffusion case, I simulate r_t at a higher frequency to replicate the observed samples from the continuous-time model.

Figure 4 plots the nonparametric jump-diffusion estimation results.[18] The upper-left panel gives the drift estimation results that are similar to the results in the diffusion model. As indicated by the wide Monte Carlo confidence bands, it is again difficult to make any firm conclusions regarding the shape of the drift coefficient.

The top-right panel of Figure 4 plots the instantaneous second moment. Results indicate that it is accurately estimated. If the data were generated by a single-factor diffusion model, the second moment would estimate $\sigma^2(r)$. With jumps, the second moment estimates $\sigma^2(r) + \lambda(r)\sigma_z^2$. Consistent with earlier results, it is easier to estimate the second moment than the drift coefficient. The lower-left panel reports estimates of $\lambda(r)$, scaled to daily units (the vertical axis gives the daily probability of a jump). The intensity varies between 6 percent and 25 percent depending on the short rate level. The Monte Carlo results indicate the estimation procedure provides approximately unbiased estimates, with the width of the confidence bands increasing with the short rate.

[17] Although it is not constant, the function (in a pointwise sense) is not statistically different from a constant function as the Monte Carlo confidence bands are relatively wide. I thank a referee for raising this point.

[18] The bandwidths used for the drift and diffusion are $h_1 = 1.25 \times \hat{s}$ and $h_2 = 0.4 \times \hat{s}$. The bandwidths for the fourth and sixth moments are $0.75 \times \hat{s}$. Note that using a larger bandwidth for the higher moments does not introduce any systematic bias and is formally justified in Bandi and Nguyen (2003).

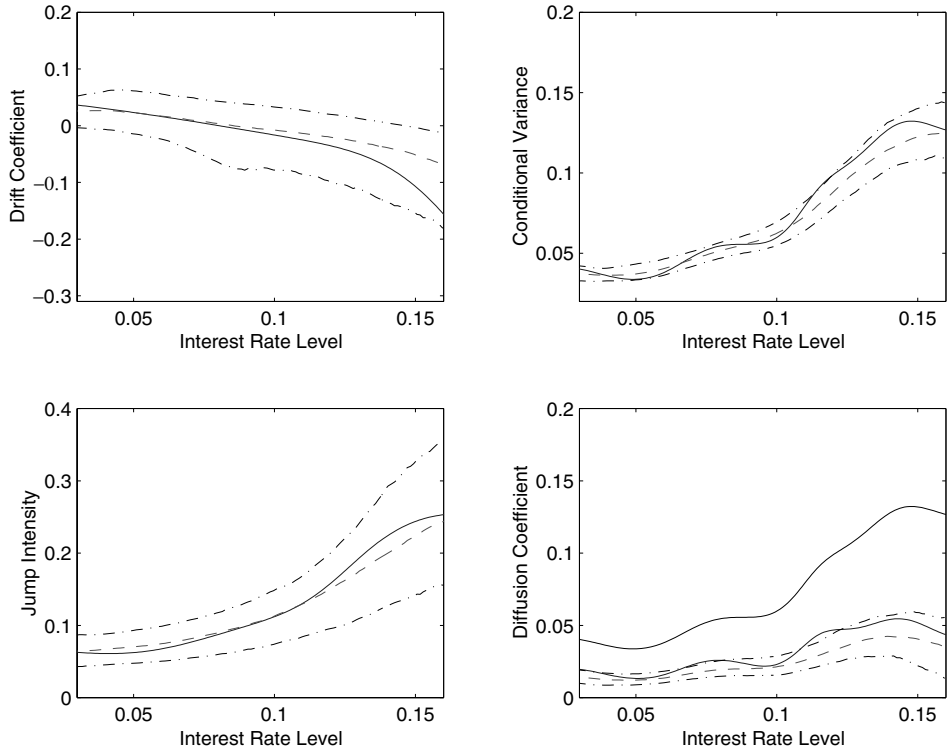

Figure 4. Nonparametric estimation results for the jump-diffusion model. The solid line is the function estimated from short rate data, the dashed line is the Monte Carlo median, and the dash–dot lines are the (10%, 90%) Monte Carlo confidence bands. To frame the results, the lower-right panel contains an additional solid line that is the total conditional variance, $\sigma^2(r) + \sigma_\xi^2 \lambda(r)$ (the latter solid line is the higher solid line in the panel).

The bottom-right panel of Figure 4 plots the results for the diffusion coefficient estimates. Note that these estimates are not as smooth as those of the drift or jump intensity. This is necessary to avoid inducing any bias. If the bandwidth on the second moment is increased, the estimated diffusion coefficient will be biased and for large bandwidths, the bias can be severe. It is more difficult to estimate $\sigma^2(r)$ because it depends on three other estimates (the jump variance, jump intensity, and second moment), which magnifies any sampling error. The lower-right panel also provides a volatility decomposition, with the solid upper line giving the total second moment. The ratio of the diffusion coefficient to the total second moment, $\sigma^2(r)/[\sigma^2(r) + \lambda(r)\sigma_z^2]$, is the proportion of variance generated by the diffusion component, and the estimates imply that diffusive components explain about half the volatility at low rates and only about a third at high rates. Thus, jump components dominate the conditional volatility of interest rate changes. As a comparison, Eraker, Johannes, and Polson (2003) find that jump in equity indices typically generates about 10 to 15 percent of the variance of returns.

Table III
The Impact of Jumps

Confidence intervals for implied jump impacts as a function of the spot rate, r_t. For each of the level of the spot rate, one, two, and three standard deviation (σ_z) confidence intervals are given for the impact of a jump on the spot rate.

	$r_t(\exp(\pm\sigma_z)-1)$	$r_t(\exp(\pm 2\sigma_z)-1)$	$r_t(\exp(\pm 3\sigma_z)-1)$
$r_t = 0.05$	(−0.0018, 0.0018)	(−0.0035, 0.0037)	(−0.0051, 0.0057)
$r_t = 0.10$	(−0.0035, 0.0037)	(−0.0070, 0.0075)	(−0.0103, 0.0114)
$r_t = 0.15$	(−0.0053, 0.0055)	(−0.0104, 0.0112)	(−0.0154, 0.0171)

The point estimate of σ_z^2 is 0.0013 with a Monte Carlo median of 0.0013 and a (10, 90) percent confidence band of (0.0012, 0.0018). These estimates imply that the finite sample distribution is skewed slightly upward. To see the impact on the level of the short rate, Table III gives confidence intervals for jump impacts for various levels of the short rate. The implied jump impacts are quite plausible. For low levels of interest rates, a one standard deviation move is 18 basis points while a three standard deviation move is 50 basis points. At higher levels of interest rates, although not common, the model implies that with a jump intensity of 25 percent, 100 basis points movements in interest rate levels would occur about once every 100 days. In fact, from Figure 2, there have been a number of daily changes greater than 100 basis points when interest rates were at high levels.

Panel D in Table II indicates that the jump-diffusion model does provide an accurate fit to the unconditional kurtosis statistic at a daily, weekly, and monthly interval and also for the conditional measures. Thus, jumps are significant in a statistical sense and provide a simple and intuitive mechanism for capturing the tail behavior of interest rates. The next section investigates the economic implications of jumps.

IV. The Economic Role of Jumps

A. Jumps and Macroeconomic Events

In this section, I investigate the connection between the occurrence of jumps and macroeconomic news arrivals. To identify the events that caused the jumps, I estimate the filtering distribution of jump times and jumps sizes, which is given by $p(J_{t+\Delta} = 1 \mid r_{t+\Delta}, r_t, \hat{\theta}_t)$ and $E[Z_{t+\Delta} J_{t+\Delta} \mid r_{t+\Delta}, r_t, \hat{\theta}_t]$, where, $\hat{\theta}_t = \{\hat{\mu}(r_t), \hat{\sigma}^2(r_t), \hat{\lambda}(r_t), \hat{\sigma}_z^2\}$ are the estimated characteristics. Computing these densities is straightforward using the Gibbs sampler that iteratively samples from $p(J_{t+\Delta} \mid r_{t+\Delta}, r_t, \hat{\theta}_t, Z_{t+\Delta})$ and $p(Z_{t+\Delta} \mid r_{t+\Delta}, r_t, \hat{\theta}_t, J_{t+\Delta})$, both of which are standard distributions. The algorithm produces a sequence $\{\{J^g_{t+\Delta}\}_{t=1}^T, \{Z^g_{t+\Delta}\}_{t=1}^T\}_{g=1}^G$, which are draws from the joint distribution $p(J_{t+\Delta}, Z_{t+\Delta} \mid r_{t+\Delta}, r_t, \hat{\theta}_t,)$. Since there is no parameter uncertainty, the algorithm converges quickly. I chose $G =$ 5,000 and discard the first 2,000 iterations as a burn-in period.

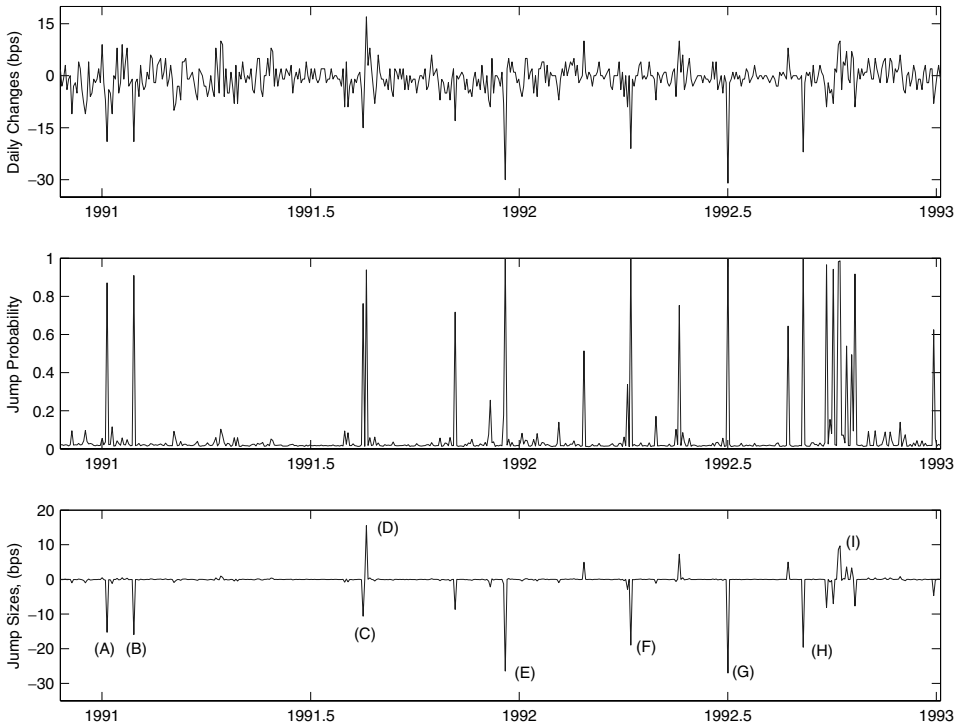

Figure 5. The time series of short rate changes, estimated jump probabilities, and estimated jump sizes, 1991–1993. The markers, A–I, represent the following dates and events: (A) 01/09/1991, the outbreak of the Gulf War; (B) 2/1/1991, a U.S. unemployment announcement and comments by the Federal Reserve; (C) 8/19/1991, the Kremlin coup and the collapse of the Soviet Union; (D) 8/21/1991, the emergence of Boris Yeltsin as leader of the remnants of the Soviet Union; (E) 12/20/1991, the Federal Reserve lowers the discount rate; (F) 4/9/1992, large Japanese equity market decline; (G) 7/2/1992, the Federal Reserve lowers the discount rate; (H) 9/4/92, a U.S. unemployment announcement; and, (I) October 1992, the Bush–Clinton presidential debates.

In order to focus on specific time periods, I consider the period from 1991 to 1993 mentioned in the introduction, the 1979 to 1981 period of high interest rates, and the Fall 1998 period. Figure 5 displays the changes in the short rate (top panel), the estimated jump probabilities (middle panel), and jump sizes (bottom panel) from 1991 to 1993. The filter identifies a number of jumps (around 10) over this period. This may seem low as the arrival intensity over this period is between 5 and 10 percent. However, since the jumps are mean zero, many of them are too small to be identified. The jump size estimates indicate that many of the jumps were quite large, some approaching 30 basis points. For 1992 when the 3-month T-bill rate was less than 4 percent, the three largest declines, denoted F, G, and H, correspond to percentage changes of -5, -9, and -7 percent, respectively.

Given model-implied jump times and sizes, we can identify the events that generated the jumps. The following is a list of the jump dates and any major

news events on that date: (A) 01/09/1991, the outbreak of the Gulf War; (B) 2/1/1991, a U.S. unemployment announcement and comments by the Federal Reserve; (C) 8/19/1991, the Kremlin coup and the collapse of the Soviet Union; (D) 8/21/1991, the emergence of Boris Yeltsin as leader of the remnants of the Soviet Union; (E) 12/20/1991, the Federal Reserve lowers the discount rate; (F) 4/9/1992, large Japanese equity market decline; (G) 7/2/1992, the Federal Reserve lowers the discount rate; (H) 9/4/92, a U.S. unemployment announcement; and, (I) October 1992, the Bush–Clinton presidential debates.

These events indicate that there were three major sources of jumps: first, official announcements regarding the current state of the economy such as unemployment announcements; second, announcements by the Federal Reserve regarding monetary policy; and, third, exogenous political–economic events in the United States or other important countries. The fall of the Soviet Union, the realization that the Gulf War would indeed occur, and problems for major trading partners such as Japan all had large impacts on interest rates. Ignoring the political debates, the remaining jumps were generated equally by regularly scheduled announcements (unemployment and Federal Reserve target changes) and by surprise macroeconomic news (Gulf War, Kremlin Coup, etc.). These results reinforce the intuition conjectured earlier, namely, that jumps provide the mechanism through which information about the macroeconomy enters the Treasury market.

Figure 6 displays jump time estimates in 1979. As interest rates increased, the jump intensity increased and more jumps arrived. This is consistent with the following explanation. As jumps occurred when *unanticipated* information arrived, market participants' forecasts during both the monetary experiment and periods of high interest rates were less accurate, which led to more unanticipated information arrivals and more jumps. The events generating jumps in 1979 were slightly different than those for the period from 1991 to 1993. There were 12 jumps identified with high probability (more than 60 percent) and the news that generated these jumps included trade deficit, industrial production, GNP, consumer prices and retail sales announcements, events associated with gas prices (riots associated with gas shortages, a diesel fuel tax strike, and the announcement of energy quotas), President Carter's declaration of war on inflation, Ford Motor Company's layoffs, and the Iranian crises. Again, all of these events provided news regarding the current and future state of the macroeconomy. Only one of the movements identified as a jump in 1979 occurred without any major news (10/22/79). Furthermore, similar types of events generated the jumps observed in 1981.

These results point toward a different interpretation of the 1979 to 1982 period than the interpretation suggested by stochastic volatility or regime switching models. In both of these models, the 1979 to 1982 period of high interest rates is interpreted as a period of high volatility. In contrast to this explanation, the jump model argues that the source of the large movements was an increased rate of surprise information arrivals.

Finally, Figure 7 provides some evidence of misspecification via jump probabilities in the Fall of 1998. During the Long-Term Capital Management

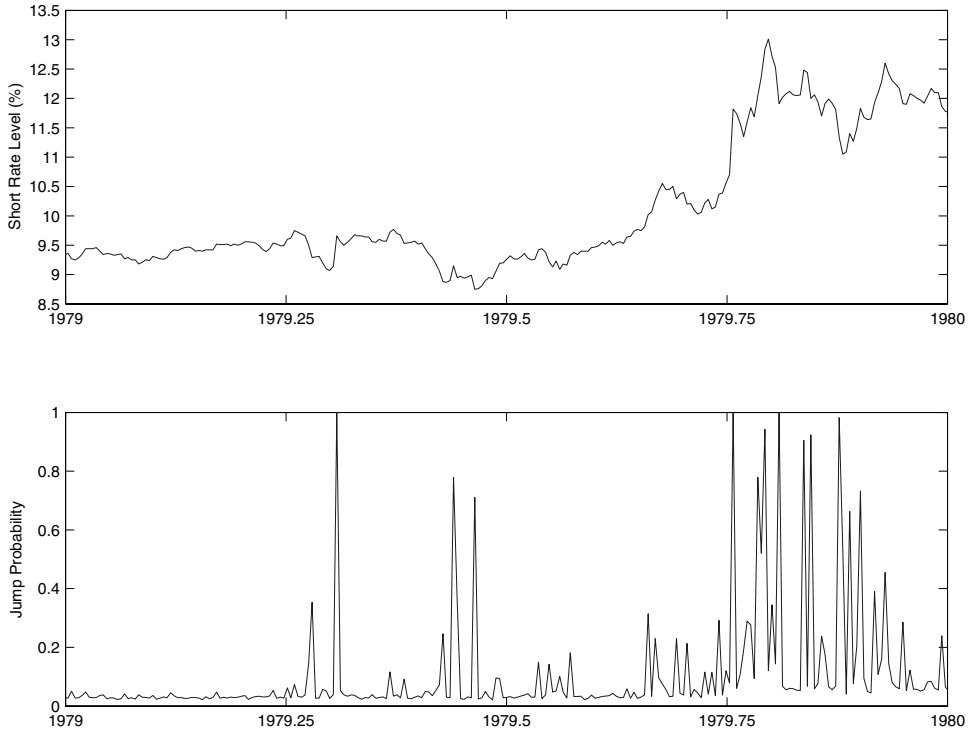

Figure 6. The time series of the short rate level and estimated jump probabilities in 1979.

crisis in September 1998, there is a period during which there are jumps on four consecutive days. This is extremely unlikely at the relatively low rates that were observed during this period, and these results indicate that this period of time might be more accurately modeled with a stochastic volatility factor.

B. Jumps and the Term Structure

This section considers the impact of jumps on bond prices. I compute the zero-coupon term structure using the jump-diffusion model in equation (13) as the model of the short rate and I compare the resulting term structure to one generated by the nonparametric diffusion model. This exercise may suffer from the proxy problem of Chapman et al. (1999). However, these authors show that the expected bias from this procedure is small unless there is a substantial market price of risk and nonlinearities in the drift or diffusion.

Pricing in the presence of jumps is more complicated than in the diffusion case. In general, continuously distributed jumps introduce an incompleteness in the market, as the discontinuity cannot be hedged with a finite number of securities. Standard arguments in Björk, Kubanov, and Runggaldier (1997)

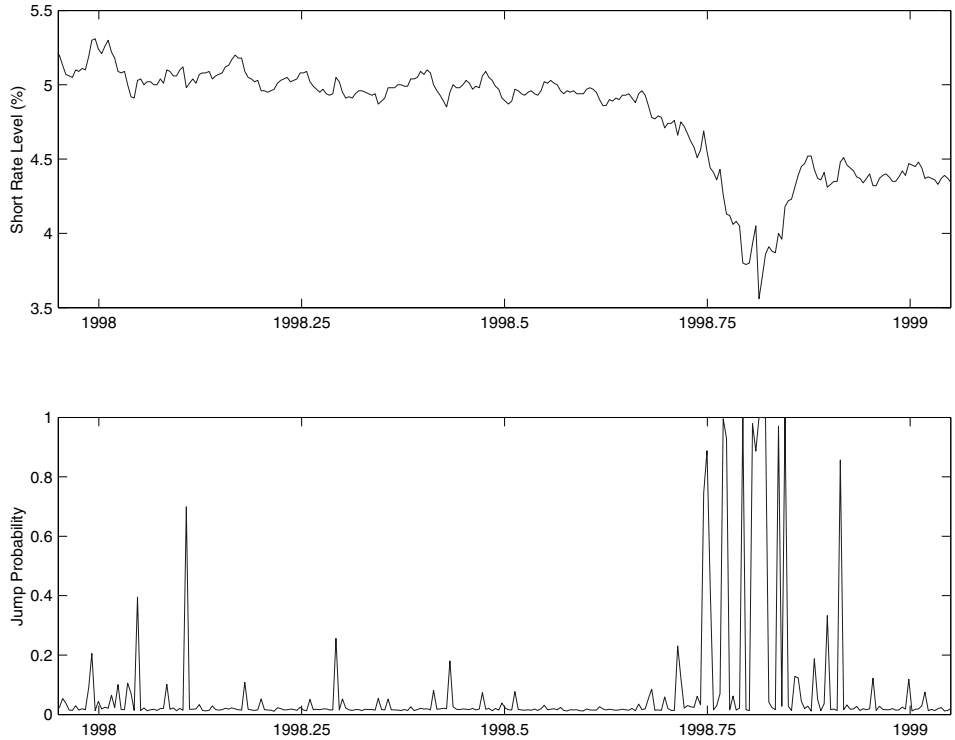

Figure 7. The time series of the short rate level and estimated jump probabilities in 1998.

indicate that there exists a new measure, $\mathbb{Q}$, equivalent to $\mathbb{P}$, such that the price of a zero-coupon bond is

$$P(r_t, \tau) = E_t^{\mathbb{Q}}\left[\exp\left(-\int_t^{t+\tau} r_s\, ds\right)\right].$$

Under $\mathbb{Q}$, the spot rate evolves according to

$$d\log(r_t) = [\mu(r_t) - \eta_t \sigma(r_t)]\, dt + \sigma(r_t)\, dW_t(\mathbb{Q}) + d\left(\sum_{n=1}^{N_t(\mathbb{Q})} Z_n(\mathbb{Q})\right),$$

where η_t is the market price of diffusive risks, $N_t(\mathbb{Q})$ is a doubly stochastic point process under $\mathbb{Q}$ with intensity $\lambda^{\mathbb{Q}}(r_t)$, the jump sizes are distributed $\Pi^{\mathbb{Q}}$, and $W_t(\mathbb{Q})$ is a standard Brownian motion. This measure is not unique and thus additional information is required to pin it down.

To focus on the role of jumps and their associated market prices of risk, I assume that there is no market price of risk for diffusive components. This assumption may not be critical as Stanton (1997) finds that the diffusive market price of risk is close to zero for both low and high interest rates. I assume the

market prices of jump risk take a simple form: the jump arrival rate under $\mathbb{Q}$ is $\lambda(r_t)^{\mathbb{Q}} = \theta_1\lambda(r_t)$ and the jump distribution under $\mathbb{Q}$ is $Z_n(\mathbb{Q}) \sim N(\theta_2, \theta_3\sigma_z^2)$. The parameters θ_1, θ_2, and θ_3 are interpreted as risk premia parameters. I use the estimates obtained earlier to compare the nonparametric diffusion and jump-diffusion models, which implies that the diffusion and jump-diffusion models have common first and second moments.

The usual procedure at this point is to calibrate the model to a given day's yield curve (Aït-Sahalia (1996a)). Instead, I consider a comparative statics exercise that documents how the yield curve depends on the given risk-neutral parameters. The parameters chosen are arbitrary and are given by $\theta_1 = (0.75, 1.25), \theta_2 = (-0.001, 0.001)$, and $\theta_3 = (0.75, 1.25)$. These risk premia alter the intensity and jump variance by 25 percent and change the jump mean by approximately 10 basis points. To compute bond prices, I used the Monte Carlo method with 10,000 simulations from the continuous-time model. Together, the number of simulations and the discretization interval (1 day) are chosen to render the Monte Carlo error negligible (less than 1 basis point at the long end of the yield curve).

Figures 8 and 9 give the resulting yield curves when the current short level is 5 and 10 percent, respectively. In both cases, the upper-left panel displays the term structure of interest rates for the models without risk premiums. The yield curves are quite similar, but this is not surprising. The bond price depends on the conditional distribution of the average interest rate over the life of a bond, $P(r_t, t, T) = E^{\mathbb{Q}}[e^{-(T-t)\bar{r}_{t,T}} \mid r_t]$, where $\bar{r}_{t,T}$ is the average interest rate. Since the diffusion and jump-diffusion models have the same conditional mean and volatility, they generate similar conditional distributions of the average interest rate, which implies that bond prices are not very different. The models do differ slightly at the long end of the curve, however, due to a Jensen's inequality-type effect.

The upper- and lower-right panels display the impact of the arrival (θ_1) and the jump variance (θ_3) risk premiums. When $r_t = 0.05$, these parameters affect the yield curves similarly as an increase in either results in a small but significant increase in yields. The largest yield increases occur in the middle of the curve indicating the jump variance effects the curvature of the yield curve. The lower-left panel gives the results for the mean jump-size risk premium. The yield curves are particularly sensitive to this parameter. A positive θ_2 increases the average jump size and the yields, as the average short rate over the life of the bond increases. In the case $r_t = 0.10$, the effects are largely similar, with the exception that the market price of jump volatility risk induces a hump at short horizons, while the yield curve falls below the risk-neutral curve for higher rates.

C. Jumps and Interest Rate Option Prices

Although jumps may not necessarily have a large impact on the cross-section of bond prices, they do have a large impact on the dynamics of interest rate movements. This implies that jumps may have the greatest impact on derivative

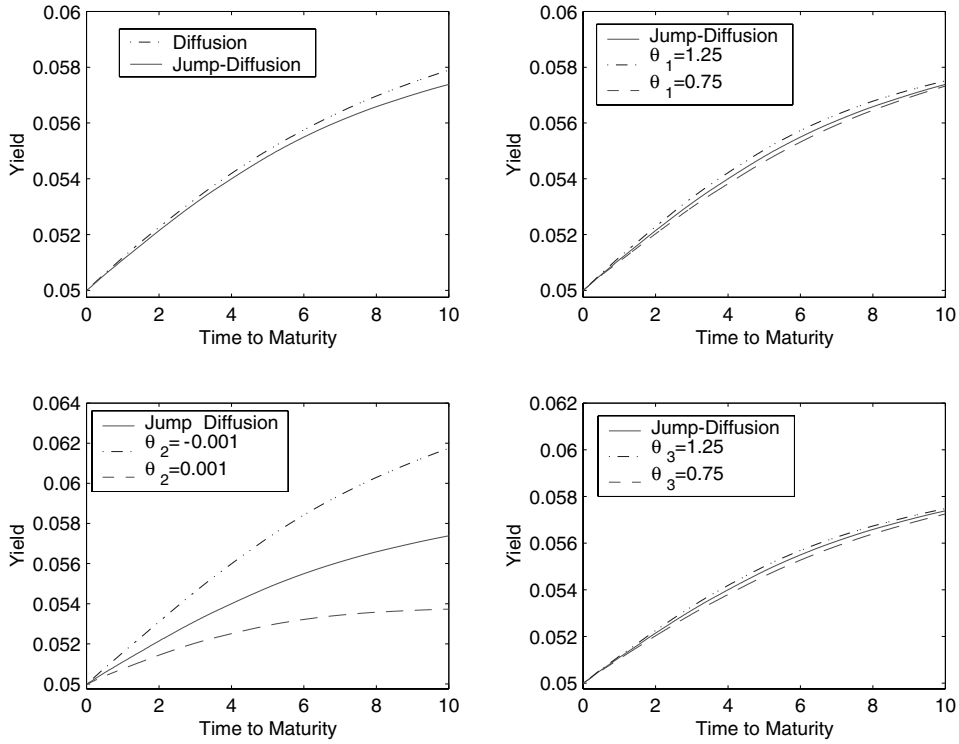

Figure 8. Term structures for the diffusion and jump-diffusion models assuming that the initial spot rate, r_t, is 0.05. The upper-left panel gives yield curves under the two models with no risk premiums. The upper-right panel shows the effect of the jump-intensity risk premium, θ_1, the lower-left panel shows the effect of the jump-mean risk premium, θ_2, and the lower-right panel shows the effect of the jump-volatility risk premium, θ_3.

contracts such as bond options, caps, or floors whose prices depend heavily on the tails of the conditional distribution of interest rate increments.

I consider the 3-month T-bill option traded on the Chicago Board of Exchange, which has a European-style exercise. The call option price is

$$C(K, t, T, r_t) = E_t^{\mathbb{Q}}\left[\exp\left(-\int_t^T r_s\, ds\right)\max(Y_{T,3} - K, 0)\right], \tag{20}$$

where $Y_{T,3}$ is the 3-month T-bill discount rate. Since the impact of jumps will likely be greatest for out-of-the-money contracts, I consider the difference in out-of-the-money option prices for short maturity options. As in the previous section, I consider option prices generated by the diffusion and jump-diffusion models, such that the first two moments are equal. The option prices are computed by Monte Carlo simulation.

Table IV provides the relative differences in option prices for contracts maturing in 2 weeks, 1 month, and 3 months. For each contract, I report the

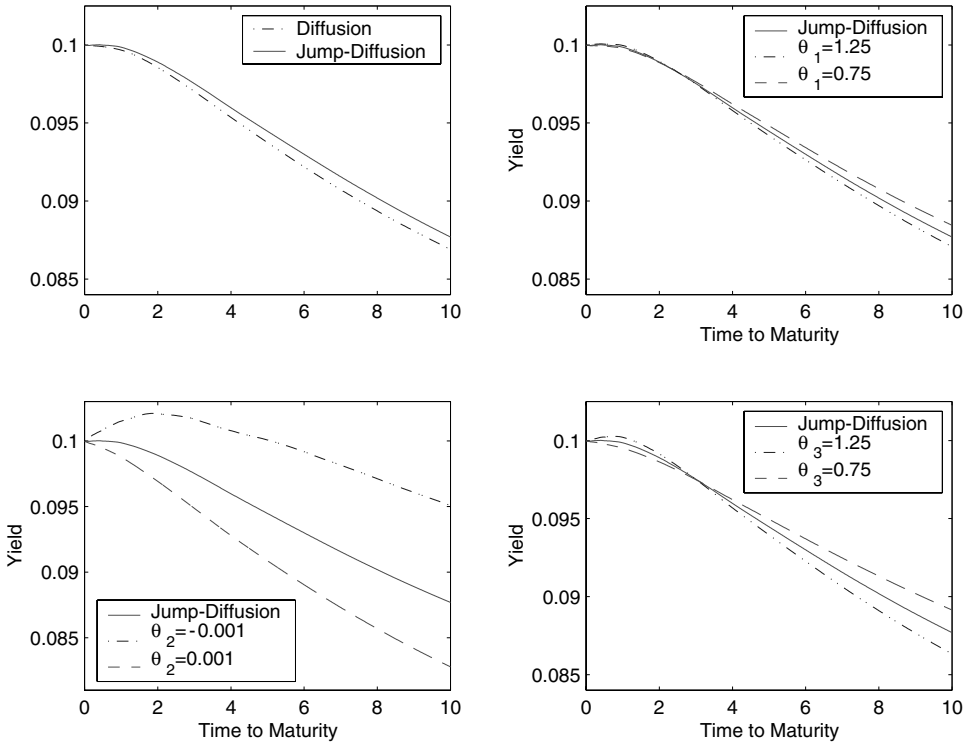

Figure 9. Term structures for the diffusion and jump-diffusion models assuming that the initial spot rate, r_t, is 0.10. The upper-left panel gives yield curves under the two models with no risk premiums. The upper-right panel shows the effect of the jump-intensity risk premium, θ_1, the lower-left panel shows the effect of the jump-mean risk premium, θ_2, and the lower-right panel shows the effect of the jump-volatility risk premium, θ_3.

Table IV

The Impact of Jumps on Option Prices

The relative differences between Treasury-bill call option prices under a diffusion and jump-diffusion model where the first two instantaneous moments are constrained to be equal.

	2-Weeks			1-Month			3-Months		
Strike	6.00	6.50	6.75	6.00	6.50	7.00	6.00	7.00	8.00
Difference	−0.0103	0.2837	0.6271	0.0189	0.0835	0.4091	−0.0035	0.0895	0.6748

relative difference for call option prices for the at-the-money and two out-of-the-money contracts. If C^D denotes the diffusion-based call price and C^{JD} the jump-diffusion call price, Table IV reports $(C^{JD} - C^D)/C^{JD}$. The initial spot rate for all of the contracts is 6 percent. In all cases I report option prices for strikes that are above the minimum tick level. Not surprisingly, the results indicate large differences in the call option prices. While at-the-money option prices are

similar, options with 2 weeks to maturity and only 50 basis points out-of-the-money differ by almost 30 percent. For longer maturities, the differences can be as great provided the contracts are sufficiently out-of-the-money. Glasserman and Kou (2003) argue that implied volatility skews embedded in cap prices are well captured by incorporating jumps.

These pricing results are similar to those in Aït-Sahalia (1996a), who finds only small differences in bond prices generated by the Vasicek (1978) model, the Cox et al. (1985) model, and a nonparametric diffusion models, but relatively large differences for bond option prices. Also, and not surprisingly, the above results are similar to those in equity option pricing for which jumps have the greatest impact on short maturity, out-of-the-money options (see Bakshi, Cao, and Chen (1997) and Pan (2002)). The main difference between the interest rate options and equity options is that interest rate jumps are symmetric whereas equity index jumps tend to be negative, which leads to the implied volatility smirk.

Some option-implied evidence supporting the importance of jumps comes from Christiansen and Hansen (2002). They find that implied volatilities for short maturity options have a strong smile shape. For example, for options with maturities between 7 and 90 days, Christiansen and Hansen find that implied volatility for out-of-the-money call options is 19.45 percent, 6.98 percent for at-the-money, and 19.67 percent for in-the-money options. This indicates that there is a severe volatility smile, and mean zero jumps are capable of generating this sort of pattern.

V. Conclusion

This paper analyzes the role of jumps in continuous-time models of the short interest rate. Results indicate that jumps are both economically and statistically important. Diffusion models ignore jumps and are misspecified in the sense that they cannot accurately capture the tail behavior of interest rate changes. Estimates imply that jumps remove the tail misspecification and, moreover, generate more than half the conditional variance of interest rate changes. For pricing purposes, jumps have an important impact on interest rate option prices, but they play a lesser role in determining the cross section of bond prices.

Using model-implied jump times and sizes, I find that jumps are typically generated by the surprise arrival of news about the macroeconomy. This poses new challenges for interest rate modeling, and, in particular, points to the importance of the approach of Piazessi (2001, 2003), who formally incorporates the Federal Reserve and macroeconomic announcements into a term structure model. One problem with this approach, however, is that there are a large number of announcements that generate jumps, and it is intractable to model more than a few announcements.

Jumps provide an interesting avenue for investigating the impact of macroeconomic variables on the term structure. For example, if a jump arrives only when a surprise macroeconomic announcement is released, the conditional

distribution of jump sizes provides a view of how the market translates macroeconomic news into the yield curve. The next step is to explicitly analyze multifactor models incorporating predictable and unpredictable jumps in the spot rate.

Appendix: Regularity and Moment Conditions

This appendix states the regularity conditions that guarantee the existence and uniqueness of the stochastic differential equations used in the text and also derives the limiting moment conditions used for identifying the drift, diffusion, jump intensity, and jump size volatility. References on stochastic differential equations of this type include Gihkman and Skorohod (1972) and Jacod and Shiryaev (1987).

The general form of the jump-diffusion model is given by

$$r_{t+\Delta} = r_t + \int_t^{t+\Delta} \mu(r_s)\,ds + \int_t^{t+\Delta} \sigma(r_{s-})\,dW_s + \sum_{n=N_t+1}^{N_{t+\Delta}} c(r_{\tau_{n-}}, Z_n), \qquad \text{(A1)}$$

where $\{\tau_n\}_{n\geq 1}$ are the jump times, N_t is a counting process with stochastic intensity $\lambda(r_t)$, Z_n are the marks of the point process, Π is the jump size distribution, and c is the jump impact function that translates the marks into jumps in r_t. Setting $c = 0$ or $\lambda = 0$ provides the results for the diffusion case.

I impose the following regularity conditions: (A) μ, σ^2, c, and $\lambda \geq 0$ are continuous functions of r and Π is independent of r_t; (B) μ, σ^2, c, and λ satisfy the linear growth condition

$$|\mu(x)| + |\sigma(x)| + \lambda(x) \int |c(x,z)|\Pi(dz) \leq k(1 + |x|), \qquad \text{(A2)}$$

for a positive constant k, and the local Lipschitz condition, for any $N > 0$ and for both $|x| \leq N$ and $|y| \leq N$, there is a k_N such that

$$|\mu(x) - \mu(y)| + |\sigma(x) - \sigma(y)| + \lambda(x) \int |c(x,z) - c(y,z)|\Pi(dz) \leq k_N|x - y|; \quad \text{(A3)}$$

(C) $\lambda(x) \int |c(x,z)|^p \Pi(dz) \leq k(1 + |x|^p)$; and, (D) $E(r_0^p) < \infty$ for $p > 2$. These conditions guarantee the existence and uniqueness of a strong solution to the stochastic differential equation (see Gihkman and Skorohod (1972), Chapter 2). In addition, conditions B, C, and D also imply that $E|r_t|^p \leq \alpha E(1 + |r_0|^p) < \infty$, where α is a finite constant depending on t and p but not N. Moreover, this guarantees finite conditional moments and provides the bound used below in a dominated convergence argument.

The derivation of the limiting moment conditions in the diffusion case is given in Gikhman and Skorohod (1972, pp. 68–69). Consider the conditional

first moment in the diffusion model

$$\frac{1}{\Delta}E_{t,r}[r_{t+\Delta}-r_t]=\frac{1}{\Delta}E_{t,r}\left[\int_t^{t+\Delta}\mu(r_s)\,ds\right]=E_{t,r}\left[\int_0^1\mu(r_{t+\Delta v})\,dv\right]$$
$$=\left[\int_0^1 E_{t,r}[\mu(r_{t+\Delta v})]\,dv\right], \tag{A4}$$

where $E_{t,r}[r_{t+\Delta}-r_t]=E[r_{t+\Delta}-r_t\,|\,r_t=r]$, the change of variables $v=\frac{s-t}{\Delta}$ is used, and Fubini's theorem is applied to exchange the order of integration and expectation. The solution $r_t(\omega)$, is continuous (r_t is a diffusion) for all $\omega\in A$, with $\mathbb{P}(A)=1$, and $\lim_{\Delta\downarrow 0} r_{t+v\Delta}=r_t$ with probability 1. By the assumed continuity of the drift function, $\lim_{\Delta\downarrow 0}\mu(r_{t+v\Delta})=\mu(r_t)$ with probability 1. The linear growth condition implies that $\mu(r_{t+\Delta v})\le k(1+|r_{t+\Delta v}|)$, which in turn implies that the right-hand side is dominated by an $E_{t,r}$-integrable random variable. Dominated convergence implies that $E_{t,r}[\mu(r_{t+\Delta v})]$ converges to $E_{t,r}[\mu(r_t)]=\mu(r_t)$. Another application of dominated convergence implies that $\lim_{\Delta\downarrow 0}E_{t,x}[\frac{r_{t+\Delta}-r_t}{\Delta}]=\mu(r)$. The higher instantaneous moments are derived in a similar manner by passing the limits under the integral after an application of Ito's Lemma:

$$\frac{1}{\Delta}E_{t,r}\left[\left(R_t^{t+\Delta}\right)^j\right]$$
$$=\frac{1}{\Delta}E_{t,r}\int_t^{t+\Delta}\left[j\left(R_t^s\right)^{j-1}\mu(r_s)+\frac{1}{2}j(j-1)\left(R_t^s\right)^{j-2}\sigma^2(r_s)\right]ds, \tag{A5}$$

where $R_t^s=(r_s-r_t)$. The conditional moments are finite, again by (A) and (B), and the same dominated convergence arguments apply.

In the jump-diffusion model, Ito's Lemma implies that if r_t solves A1, then if μ,σ,c,λ, and Π are sufficiently regular and g is twice differentiable,

$$dg(r_t)=\mathcal{L}g(r_t)\,dt+\mathcal{A}g(r_t)\,dt+g_x(r_t)\sigma(r_t)\,dW_t+d\left(\sum_{n=1}^{N_t}[g(r_{\tau_n})-g(r_{\tau_{n-}})]\right)$$
$$-\lambda(r_t)\int_Z\left[g(r_t+c(r_t,z)-g(r_t))\right]\Pi(dz)\,dt, \tag{A6}$$

where $\mathcal{L}$ and $\mathcal{A}$ are differential operators for the diffusion and jump component

$$\mathcal{L}g(x)=g_x(x)\mu(x)+\tfrac{1}{2}g_{xx}(x)\sigma^2(x) \tag{A7}$$

and

$$\mathcal{A}g(x)=\lambda(x)\int_Z[g(x+c(x,z)-g(x)-g_x(x)c(x,z)]\Pi(dz), \tag{A8}$$

respectively.

To derive the limiting moments of the jump-diffusion model, proceed as in the diffusion case noting that r_t is now right-continuous. To identify the diffusion and jump parameters, we apply Ito's Lemma to functions of the form $(r_{t+\Delta} - r)^j$ for $j > 2$. This results in

$$
\begin{aligned}
&E_{t,r}\left[(r_{t+\Delta} - r_t)^j\right] \\
&\quad = E_{t,r}\left[\int_t^{t+\Delta}\left[j\left(R_s^t\right)^{j-1}\mu(r_s) + \frac{1}{2}j(j-1)\left(R_s^t\right)^{j-2}\sigma^2(r_s)\right]ds\right] \\
&\qquad + E_{t,r}\left[\int_t^{t+\Delta}\lambda(r_s)\int_Z\left[\left(R_s^t + c(r_s,z)\right)^j - \left(R_s^t\right)^j - j\left(R_s^t\right)^{j-1}c(r_s,z)\right]\Pi(dz)\,ds\right].
\end{aligned}
\tag{A9}
$$

For example, the limiting value of terms such as $E_{t,r}\frac{1}{\Delta}[\int_t^{t+\Delta}(r_s - r_t)\mu(r_s)\,ds]$ must be evaluated. Since $\lim_{s\downarrow t} r_s = r_t$, $\lim_{\Delta\downarrow 0}\ \mu(r_{t+v\Delta}) = \mu(r_t)$ with probability 1. The finite moments imply that higher-order conditional moments are finite, Fubini's theorem and the same dominated convergence argument together imply

$$
\lim_{\Delta\downarrow 0}\frac{1}{\Delta}E_{t,r}\left[\int_t^{t+\Delta}(r_s - r_t)\mu(r_s)\,ds\right] = 0.
$$

REFERENCES

Ahn, C.M., and Howard Thompson, 1985, Jump-diffusion processes and the term structure of interest rates, *Journal of Finance* 43, 7155–7174.

Ahn, Dong-Hyun, Robert Dittmar, and A. Ronald Gallant, 2002, Quadratic term structure models: Theory and evidence, *The Review of Financial Studies* 15, 243–288.

Ahn, Dong-Hyun, and Bin Gao, 1999, A parametric nonlinear model of term structure dynamics, *The Review of Financial Studies* 12, 721–763.

Aït-Sahalia, Yacine, 1996a, Nonparametric pricing of interest rate dependent 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, 2002, Telling from discrete data whether the underlying continuous-time model is a diffusion, *Journal of Finance* 57, 2075–2112.

Andersen, Torben, and Jesper Lund, 1997, Estimating continuous-time stochastic volatility models of the short rate, *Journal of Econometrics* 77, 343–377.

Andersen, Torben, and Jesper Lund, 1998, A 3-factor model of short term interest rates, Working paper, Graduate School of Business, Northwestern University.

Ang, Andrew, and Monika Piazzesi, 2003, A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables, *Journal of Monetary Economics* 50:4, 745–787.

Bakshi, Gurdip, Charles Cao, and Zhiwu Chen, 1997, Empirical performance of alternative option pricing models, *Journal of Finance*, 52, 2003–2049.

Balduzzi, Pierluigi, Edwin Elton, and T. Clifton Green, 2001, Economic news and bond prices: Evidence from the U.S. treasury market, *Journal of Financial and Quantitative Analysis* 36, 523–543.

Bandi, Federico, 2000, Non-parametric fixed income pricing: Stanton's (1997, JF) approach revisited, unpublished paper, University of Chicago.

Bandi, Federico, 2002, Short-term interest rate dynamics: A spatial approach, *Journal of Financial Economics* 65, 73–110.

Bandi, Federico, and Thong Nguyen, 1999, Fully nonparametric estimators for diffusions: A small sample analysis, Working paper, University of Chicago.

Bandi, Federico, and Thong Nguyen, 2003, On the functional estimation of jump-diffusion processes, *Journal of Econometrics* 116, 293–328.

Bandi, Federico, and Peter C.B. Phillips, 2002, Fully nonparametric estimation of scalar diffusion models, *Econometrica* (forthcoming).

Barnard, George, 1963, Discussion of Professor Bartlett's paper, *Journal of the Royal Statistical Association, Series B,* 25, 294.

Baz, Jamil, and Sanjiv Das, 1996, Analytical approximations of the term structure for jump-diffusion processes: A numerical analysis, *Journal of Fixed Income* 6, 78–86.

Björk, Tomas, Yuri Kabanov, and Wolfgang Runggaldier, 1997, Bond market structure in the presence of marked point processes, *Mathematical Finance* 7, 211–239.

Chacko, George, and Sanjiv Das, 2002, Pricing interest rate derivatives: A general approach, *Review of Financial Studies* 15, 195–241.

Chan, K.C., Andrew Karolyi, Francis Longstaff, and Anthony Sanders, 1993, An empirical comparison of alternative models of the short term interest rate, *Journal of Finance* 47, 1209–1227.

Chapman, David, Jonathan Long, and Neil Pearson, 1999, Using proxies for the short rate: When are three months like an instant, *The Review of Financial Studies* 12, 763–807.

Chapman, David, and Neil Pearson, 2000, Is the short rate drift actually nonlinear, *Journal of Finance* 55, 1, 355–388.

Christiansen, Charlotte, and Charlotte Hansen, 2002, Implied volatility of interest rate options: An empirical investigation of the market model, *Review of Derivatives Research* 5, 51–79.

Collin-Dufresne, Pierre, Robert Goldstein, and Christopher Jones, 2002, Identification and estimation of 'maximal' affine term structure models: An application to stochastic volatility, Working paper, Carnegie-Mellon University.

Conley, Timothy, Lars Hansen, and Wen-Fang Liu, 1997, Bootstrapping the long-run, *Macroeconomic Dynamics* 1, 279–311.

Cox, John, Jonathan Ingersoll, and Stephen Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 363–384.

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

Das, Sanjiv, 2001, The surprise element: Jumps in interest rates, *Journal of Econometrics* 106, 27–65.

Das, Sanjiv, and Rangarajan Sundaram, 1999, Of smiles and smirks: A term structure perspective, *Journal of Financial and Quantitative Analysis* 34, 211–240.

Downing, Chris, 2000, Nonparametric estimation of multifactor continuous-time interest rate models, Working paper, Federal Reserve Board.

Duffee, Greg, 1996, Idiosyncratic variation of Treasury bill yields, *Journal of Finance* 51, 527–551.

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

Duffie, Darrell, and Jun Pan, 1997, An overview of value at risk, *Journal of Derivatives* 4, 7–49.

Duffie, Darrell, Jun Pan, and Kenneth Singleton, 2000, Transform analysis and asset pricing for affine jump-diffusions, *Econometrica*, 68, 1343–1376.

Eraker, Bjorn, M. Johannes, and N. Polson, 2003, The impact of jumps in volatility and returns, *Journal of Finance* 58, 1269–1300.

Fleming, Michael, and Eli Remolona, 1997, What moves the bond market, *FRBNY Economic Policy Review* 3:4, 31–50.

Fleming, Michael, and Ansar Sarkar, 1999, Liquidity in U.S. Treasury spot and futures markets, Working paper, Federal Reserve Bank of New York.

Furfine, Chris, 2001, Do macroeconomic announcements still drive the Treasury market, *BIS Quarterly Review*, June, 49–57.

Gihkman, I., and A.V. Skorohod, 1972, *Stochastic Differential Equations*. Springer–Verlag, New York.

Glasserman, Paul, and Steven Kou, 2003, The term structure of simple forward rates with jump risk, *Mathematical Finance* 383–410.

Hamilton, James, 1996, The daily market for Federal funds, *Journal of Political Economy* 104, 26–56.

Hardle, Wolfgang, and Edwin Mammen, 1993, Comparing nonparametric versus parametric regression fits, *Annals of Statistics* 21, 1926–1947.

Honore, Peter, 1998, Maturity induced bias in estimating spot-rate diffusion models, Working paper, Aarhus School of Business, Denmark.

Huang, Chi-Fu, 1985, Information structure and equilibrium asset pricing, *Journal of Economic Theory* 35, 33–71.

Jacod, Jean, and A. Shiryaev, 1987, *Limit Theorems for Stochastic Processes*, Grundlehren der Mathematischne Wissenschaften, Vol. 288, Springer Verlag, New York.

Jiang, George, 1998, A nonparametric approach to the estimation 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, and J. 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.

Jones, Christopher, 2003, Nonlinear mean reversion in the short term interest rate, *Review of Financial Studies* 16, 793–843.

Milshtein, G., 1974, Approximate integration of stochastic differential equations, *Theory of Probability and Its Applications* 23, 396–401.

Pan, Jun, 2002, The jump-risk premia implicit in options: Evidence from an integrated time-series study, *Journal of Financial Economics* 63, 3–50.

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

Piazzesi, Monika, 2001, A linear quadratic jump diffusion model with scheduled and unscheduled announcements, Working paper, UCLA.

Pritsker, Matthew, 1998, Nonparametric density estimation and tests of continuous-time interest rate models, *Review of Financial Studies* 11, 449–487.

Stanton, Richard, 1997, A nonparametric model of term structure dynamics and the market price of interest rate risk, *Journal of Finance* 52, 1973–2002.

Vasicek, Oldrich, 1978, An equilibrium characterization of the term structure, *Journal of Financial Economics* 5, 17–88.

Zhou, Hao, 2001, Jump-diffusion term structure and the Itô conditional moment generator, Working paper, Duke University.