Affine Jump Term Structure Models: Expectation Puzzles, and Conditional Volatility*

Haitao Li†
Stephen M. Ross School of Business
University of Michigan

Zhaogang Song†
Board of Governors of
the Federal Reserve System

October 2012

ABSTRACT

The celebrated affine term structure models (ATSMs) of Duffie and Kan (1996) and Dai and Singleton (2000) face two empirical challenges. First, they ignore well-documented jumps in interest rates as the state variables follow affine diffusions. Second, they fail to simultaneously capture violations of the “expectation hypothesis” and time variations in conditional volatility of bond yields. We develop affine jump term structure models (AJTSMs), in which the state variables follow affine jump-diffusions, and provide a comprehensive empirical analysis of three-factor AJTSMs using Libor swap rates from 1990 to 2008. Our empirical results show that jumps are essential for modeling term structure dynamics, as AJTSMs significantly outperform ATSMs in capturing conditional moments of yields, especially the skewness and kurtosis. By allowing the jump size to be state dependent, our model also captures jump effects at the long end of the maturity spectrum, which is consistent with observable jumps in long-term yields in the data. Moreover, we show that jump risk premiums lead to flexible time-varying market prices of risks without restricting time variations in conditional volatility. Consequently, two sub-classes of three-factor AJTSMs simultaneously capture violations of the “expectation hypothesis” and time variations of the conditional volatility of Libor swap rates.

*Special thanks to Bob Jarrow and Andrew Karolyi for their encouragement and numerous helpful comments. We are also grateful to Warren Bailey, Alan Bester, David Bowman, Ren-Raw Chen, Stefania D’Amico, Michiel DePooter, Dobri Dobrev, George Gao, Janet Gao, Bob Kimmel, Yuriy Kitsul, Canlin Li, Hedibert Lopes, Nick Polson, Pamela Moulton, Min Wei, Robert Whitelaw, Hao Zhou, seminar participants at Cornell University, Federal Reserve Board, Fordham University, Chicago Booth, the 5th SoFiE annual conference, and the 2012 CICF conference for their helpful discussions. All remaining errors are solely ours. A large portion of this article appeared previously under the title “Expectation Puzzles, Time-Varying Conditional Volatility and Jumps in Affine Term Structure Models.” The views expressed herein are the author’s and do not necessarily reflect the opinions of the Board of Governors of the Federal Reserve System.

†Haitao Li, Stephen M. Ross School of Business, University of Michigan, Ann Arbor, MI 48109. E-mail: htli@umich.edu.

‡Zhaogang Song, Monetary and Financial Market Analysis Section, Board of Governors of the Federal Reserve System, Mail Stop 165, 20th Street and Constitution Avenue, Washington, DC, 20551. E-mail: Zhaogang.Song@frb.gov.
Affine Jump Term Structure Models: Expectation Puzzles and Conditional Volatility

October 8, 2012

ABSTRACT

The celebrated affine term structure models (ATSMs) of Duffie and Kan (1996) and Dai and Singleton (2000) face two empirical challenges. First, they ignore well-documented jumps in interest rates as the state variables follow affine diffusions. Second, they fail to simultaneously capture violations of the “expectation hypothesis” and time variations in conditional volatility of bond yields. We develop affine jump term structure models (AJTSMs), in which the state variables follow affine jump-diffusions, and provide a comprehensive empirical analysis of three-factor AJTSMs using Libor swap rates from 1990 to 2008. Our empirical results show that jumps are essential for modeling term structure dynamics, as AJTSMs significantly outperform ATSMs in capturing the conditional moments, especially the skewness and kurtosis, of bond yields at both short and long maturities. Moreover, we show that jump risk premiums lead to flexible time-varying market prices of risks without restricting time variations in conditional volatility. Consequently, two sub-classes of three-factor AJTSMs simultaneously capture violations of the “expectation hypothesis” and time variations of the conditional volatility of Libor swap rates.
1 Introduction

The affine term structure models (ATSMs) of Duffie and Kan (1996), in which zero coupon bond yields are linear functions of latent state variables that follow affine diffusions, have become arguably the most popular state-of-the-art multifactor dynamic term structure models (DTSMs) due to their analytical tractability and empirical flexibility. The comprehensive empirical analysis of ATSMs by Dai and Singleton (2000) has established the canonical form of ATSMs and the standard empirical framework for analyzing multifactor DTSMs. Following the pioneering works of Duffie and Kan (1996) and Dai and Singleton (2000), a huge literature has been developed in the past decade analyzing ATSMs using bond yields, Libor swap rates, or other fixed income derivatives (see Dai and Singleton (2003) and Piazzesi (2010) for excellent reviews of the literature).

Yet notwithstanding their successes in the literature, ATSMs still face two major empirical challenges. First, by assuming the state variables and consequently the spot rate follow affine diffusions, ATSMs ignore well-documented jumps in interest rates.\(^1\) Many studies, such as Das (2002), Anderson et al. (2004), and Johannes (2004), have documented the important role of jumps in interest rates as “surprise elements.” Despite the overwhelming evidence of jumps in interest rates, Dai and Singleton (2003) point out that “state variables with jumps have received relatively less attention in the empirical literature on DTSMs.” Johannes and Polson (2009) also note that “there is little work of jump-diffusion term structure models.” While several studies have tried to incorporate jumps into term structure models (e.g., Ahn and Thompson (1988), Jiang and Yan (2009), Piazzesi (2005), and Zhou (2001)), they tend to focus on some special cases of multifactor DTSMs and do not consider maximally flexible and econometrically identifiable models in the spirit of Dai and Singleton (2000).\(^2\) As a result, these models cannot answer the many important questions raised by Johannes and Polson (2009): “Do multiple factors jump, or is it only the short rate? Does the market price diffusive and jump risks differently in the term structure? How do predictable jumps affect the term structure?”

Second, it has been widely documented in the literature that ATSMs fail to simultaneously

\(^1\)Though Duffie and Kan (1996) sketch out the extension of their model to include jumps, they do not provide a formal analysis of jump term structure models.

\(^2\)Piazzesi (2005), Cheng and Scaillet (2007), and Jiang and Yan (2009) cannot be covered by AJTSMs directly since their models have a quadratic component in the state variables. But as jumps can happen only in the affine state variables, these models can be regarded as special cases of AJTSMs as far as jumps are concerned.
capture time variations in bond risk premium and conditional volatility of bond yields. The classic papers of Fama and Bliss (1987) and Campbell and Shiller (1991) among others have shown that U.S. Treasury yields violate the “expectation hypothesis,” a result that points to time variations of bond risk premium. Moreover, there is overwhelming evidence that bond yields exhibit time-varying conditional volatility (see, for example, Aït-Sahalia (1996), Gallant and Tauchen (1998), and Andersen and Lund (1997)). In fact, the term structure of the unconditional volatility of both the levels and changes of bond yields and Libor-Swap rates tend to be hump-shaped (Litterman et al. 1991; Dai and Singleton 2000, 2003). Unfortunately, Dai and Singleton (2002), Duffee (2002), and Duarte (2004) have shown that ATSMs that are flexible enough to capture time variations in bond risk premium are incapable of generating any time variations in interest rate volatility. Therefore, serious tensions exist in ATSMs between matching the first and second moments of yield data.\(^3\)

In this paper, we provide a comprehensive theoretical and empirical analysis of multifactor affine jump term structure models (AJTSMs). Under AJTSMs, the spot rate is a linear function of latent state variables, which follow affine jump-diffusions. Since our AJTSMs are within the analytical framework of Duffie, Pan, and Singleton (2000) and Chacko and Das (2002), they share the same level of analytical tractability as traditional ATSMs and yield closed-form solutions for a wide range of fixed income securities. More important, our AJTSMs can simultaneously address the two empirical challenges facing traditional ATSMs: While jumps in the state variables help to capture the “surprise elements” in interest rates, jump risk premiums help to break the tension of traditional ATSMs in modeling the time-varying risk premium and conditional volatility of bond yields. Our paper makes three important contributions to the current literature on DTSMS.

First, we study theoretically the advantages of AJTSMs over ATSMs in capturing term structure dynamics. Following Dai and Singleton (2000), we develop canonical forms of multifactor AJTSMs that are maximally flexible and econometrically identifiable. We identify four sub-classes of three-factor “essentially” AJTSMs, which are studied in our empirical analysis. Moreover, we show that jump risk premiums generalize the market prices of risk without restricting the time-varying condi-

\(^3\)This mean-volatility tension also exists in other DTSMs, such as the semi-affine models of Duarte (2004) and the quadratic Gaussian models of Ahn et al. (2002).
tional volatility.\textsuperscript{4} In ATSMs, more flexible time-varying conditional volatility specification lead to more restrictive time-varying risk premium. In contrast, for AJTSMs, the jump risk premiums brings more flexibility for the market prices of risk without imposing any restrictions on conditional volatility. Therefore, theoretically the tight link between the market prices of risk and the time-varying conditional volatility in ATSMs is broken with the introduction of jump risk premiums.

Second, we provide a comprehensive empirical analysis of three-factor “essentially” AJTSMs using daily Libor swap rates at 6-month, 2-year, 3-year, 5-year, 7-year and 10-year maturities from August 13, 1990, to December 31, 2008. The daily Libor swap rates are readily available and are more relevant for studying jumps than the commonly used monthly Treasury yields. Our empirical analysis shows that similar to Treasury yields, Libor swap rates exhibit violations of the “expectation hypothesis” and time-varying conditional volatility. For the AJTSMs, “essentially” affine specifications for the market prices of diffusive risk are adopted (Duffee 2002). Moreover, we specify the jump risk premiums to compensate for the constant jump intensity uncertainty and state-dependent jump size risk. Our empirical analysis shows that most of the jump parameters are significantly different from zero. These estimates imply state-dependent jump size and significant jumps in the state variables under both the $\mathcal{P}$ and $\mathcal{Q}$ measures. Overall, we find positive jump risk premiums. We also find that the “essentially” AJTSMs significantly outperform the “extended” ATSMs of Cheridito et al. (2007) in capturing the conditional moments, especially the skewness and kurtosis, of bond yields at both short and long maturities.\textsuperscript{5} Overall, incorporating jumps significantly improves the fit of yield curve dynamics.

Finally, we demonstrate empirically that the flexibility offered by jump risk premiums in AJTSMs helps to break the tension between modeling the time-varying risk premium and conditional volatility for Libor swap rates. To examine whether the AJTSMs can capture the time varying risk premium, following Dai and Singleton (2002), we compare the model-implied and empirical versions of several standard “expectation hypothesis” regression coefficients. We also simulate time series of Libor swap

\textsuperscript{4}To make the AJTSMs empirically implementable, we follow Pan (2002) and Jarrow et al. (2007) to specify the risk factor dynamics under both the physical and risk-neutral measures. Then jump risk premiums are calculated as the difference of the physical and risk-neutral dynamics of jumps.

\textsuperscript{5}We do not choose to generalize the “extended” ATSMs of Cheridito et al. (2007) with jumps as the risk premium of “essentially” AJTSMs covers that of “extended” ATSMs as special cases. Therefore, “extended” ATSMs with jumps may have identification issues; see Section 2.3 for details.
rates from the estimated models and compare the volatility of the simulated data with that of the actual data. In sharp contrast to “essentially” and “extended” ATSMs, we find that two three-factor AJTSMs simultaneously match time variations in both the risk premium and conditional volatility of Libor swap rates. In fact, all sub-classes of three-factor “essentially” AJTSMs, including the one with the most flexible time-varying conditional volatility specification, can closely match the time-varying risk premium of Libor swap rates. Therefore, the empirical evidence suggests that the tension between matching the first and second moments of interest rates, which exists in ATSMs, is indeed eliminated in AJTSMs.

The rest of the paper is organized as follows. In Section 2, we develop multifactor AJTSMs and discuss their advantages over traditional ATSMs. Section 3 briefly discusses the estimation methods used in our analysis. Section 4 describes the data and stylized facts of time-varying risk premium and conditional volatility of Libor swap rates. In Section 5, we report parameter estimates and model diagnostic results of three-factor ATSMs and AJTSMs. Section 6 shows that several three-factor AJTSMs can simultaneously match time-varying risk premium and conditional volatility of Libor swap rates. Section 7 concludes.

2 Affine Jump Term Structure Models

2.1 General Model Specification

In this section, we discuss the specification of multifactor AJTSMs, which include the risk-neutral and physical dynamics of the state variables, as well as the market prices of diffusion and jump risks. Our AJTSMs are defined on a filtered probability space \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathcal{P})\), with the instantaneous risk-free rate \(r_t\) as

\[
r_t = \delta_0 + \delta_1' X_t,
\]

where \(\delta_0\) is a scalar and \(\delta_1\) is a \(d \times 1\) vector in \(\mathbb{R}^d\). The risk-neutral dynamics of the multivariate state variable \(\{X_t\}\) is given by the following stochastic differential equation (SDE):

\[
dX_t = \tilde{\mathcal{K}} \left( \bar{\theta} - X_t \right) dt + \Sigma \sqrt{S_t} dW_t^Q + dJ_t^Q,
\]

\[\text{(2.1)}\]
where $\tilde{\theta}$ is a $d \times 1$ vector, $\tilde{K}$ and $\Sigma$ are $d \times d$ matrices (could be non-diagonal and asymmetric), $S_t$ is a diagonal matrix with the $i$-th diagonal element given by

$$[S_t]_{ii} = \alpha_i + \beta_i' X_t,$$

(2.3)

$W_t^Q$ is a $d \times 1$ standard Brownian motion under the risk-neutral measure $Q$, and $J_t^Q$ is a pure jump process with a constant jump arrival intensity

$$\lambda^Q (X_t) = \lambda_0^Q > 0$$

(2.4)

and a $d$-dimensional random jump size vector $\xi^Q$ with mean $\mu^Q (X_t) = \begin{bmatrix} \mu_{10}^Q & \mu_{11}^Q \\ \vdots & \vdots \\ \mu_{d0}^Q & \mu_{d1}^Q \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ X_{1t} \\ \vdots \\ X_{dt} \end{bmatrix}$.

The term “affine” refers to the fact that the instantaneous short rate in (2.1), the drift term in (2.2), the conditional variance in (2.3), the jump intensity in (2.4), and the mean of the jump size are all affine functions of the state variables $X_t$. ATSMs are special cases of AJTSMs without jumps in $X_t$.

Note that the jump intensity is restricted to be nonnegative. However, the affine specifications in (2.1) cannot restrict the short rate $r_t$ to be positive. Negative nominal short rates are undesirable, as they may lead to arbitrage opportunities in economies with money. Following Piazzesi (2005, 2010), we view the AJSTMs as a tool to approximate true bond prices in that $r_t^{true} = \max \{ r_t, 0 \} = \max \{ \delta_0 + \delta_1' X_t, 0 \}$. As a result, AJSTMs in (2.1)-(2.4) become an approximating model that ignores the truncation induced by the max operators. The approximation is good as long as the probability that $r_t$ takes on negative values is small.

Compared with ATSMs, AJTSMs capture the potential jump activities via the pure-jump process $J_t^Q$ with a state-dependent jump size. Suppose the jump-occurrence times are $\{ \Gamma_i, i \geq 1 \}$. Given the occurrence of the $i$-th jump, the state variable $X_t$ jumps from $X(\Gamma_{i-})$ to $X(\Gamma_i-) \times_i^Q$, where $\times_i^Q$ is

---

5

---

6Our framework of AJTSMs follows Dai and Singleton (2000), which, however as shown by Collin-Dufresne, Goldstein, and Jones (2008), do not have the most general conditional volatility dynamics for $N>3$. However, we do not adopt the framework of Collin-Dufresne, Goldstein, and Jones (2008) not only because of the comparison with Dai and Singleton (2002) and Duffee (2002), but also because the Dai and Singleton (2000) framework allows simple derivations of the population regression coefficients in (6.1) and (6.2) for the expectation hypothesis.
independent of \( W_t^Q \) and \( \xi_j^Q \) for \( i \neq j \). The intuition is that the conditional probability at time \( t \) of another jump before time \( t + \Delta \) is approximately \( \lambda_0^Q \Delta \) for small \( \Delta \) and, given a jump-occurrence, the mean relative jump size is \( \mu^Q = E[\xi^Q - 1] \). Hence, the last term \( (\lambda_0^Q \mu^Q) \, dt \) in (2.2) combines the effects of random jump timing and sizes and acts as the compensator for the instantaneous change in \( X_t \) induced by the pure jump process \( J_t^Q \).

Absence of arbitrage opportunities, the time-\( t \) price of a zero-coupon bond that matures at time \( T \) is given by

\[
P(X_t, t, T) = E^Q \left[ \exp \left( - \int_t^T r_u \, du \right) \right] X_t,
\]

where the expectation is taken under the risk-neutral dynamics of \( X_t \) defined in (2.2). Following Duffie, Pan, and Singleton (2000), the bond price \( P \) satisfies the partial differential equation below:

\[
r_t P - \frac{\partial P}{\partial t} = \left[ \kappa \left( \bar{\theta} - X_t \right) - \lambda_0^Q \mu^Q \right]' \frac{\partial P}{\partial x} \bigg|_{x=X_t} \\
+ \frac{1}{2} \text{Trace} \left[ \Sigma_t \Sigma' \frac{\partial^2 P}{\partial x \partial x'} \bigg|_{x=X_t} \right] + \lambda^Q E \left[ P \left( X_t + \xi^Q, t, T \right) - P(X_t, t, T) \right],
\]

with the terminal condition \( P(X_t, T, T) = 1 \) for all \( T \).

Duffie, Pan, and Singleton (2000) show that the bond prices actually assume the exponential affine form:

\[
P(X_t, t, \tau) = \exp \left[ A(\tau) - B(\tau)' X_t \right],
\]

where \( A(\tau) \) and \( B(\tau) = [B_1(\tau), \cdots, B_d(\tau)]' \) satisfy the so-called complex-valued Riccati-type ordinary differential equations:

\[
\frac{dA(\tau)}{d\tau} = - \left( \bar{\kappa} \bar{\theta} \right)' B(\tau) + \frac{1}{2} \sum_{i=1}^N \left[ \Sigma' B(\tau) \right]^2_i \alpha_i - \delta_0 + \lambda_0^Q \left[ \zeta^Q (B(\tau)) - 1 \right],
\]

\[
\frac{dB(\tau)}{d\tau} = - \bar{\kappa}' B(\tau) - \frac{1}{2} \sum_{i=1}^N \left[ \Sigma' B(\tau) \right]^2_i \beta_i + \delta_1,
\]

where \( \tau = T - t \) is the bond’s time to maturity and \( \zeta^Q (u) = E \left[ \exp \left( u' \xi^Q \right) \right] \) is the moment-generating function of the random jump size vector \( \xi^Q \). These ordinary differential equations can be solved easily through numerical integration, starting from the initial conditions \( A(0) = 0 \) and \( B(0) = 0_{d \times 1} \). By
the yields of zero-coupon bonds, \( y(X_t, \tau) \equiv -\frac{1}{\tau} \log [P(X_t, t, \tau)] \), are also affine functions of the state variables:

\[
y(X_t, \tau) = \frac{1}{\tau} \left[ -A(\tau) + B(\tau)'X_t \right].
\] (2.9)

Therefore, AJTSMs share the same level of analytical tractability as ATSMs.

To obtain the physical dynamics of \( X_t \), we need to specify the market prices of both diffusion and jump risks. Dai and Singleton (2000) propose a “completely” affine specification for the market price of diffusion risks as \( \Lambda_t^C = (\sqrt{S_t})^{-1} (S_t \eta_1) \), where \( \eta_1 \) is a \( d \times 1 \) parameter vector. Duffee (2002) then generalizes \( \Lambda_t^C \) to the “essentially” affine specification

\[
\Lambda_t^E = \left( \sqrt{S_t} \right)^{-1} \left[ S_t \eta_1 + \sqrt{S_t S_t^-} \eta_2 X_t \right],
\] (2.10)

where \( S_t^- \) is a \( d \times d \) diagonal matrix with the \((i, i)\)-th element

\[
[S_t^-]_{ii} = \begin{cases} 
(\alpha_i + \beta_i'X_t)^{-1}, & \text{if } \inf (\alpha_i + \beta_i'X_t) > 0 \\
0, & \text{otherwise}
\end{cases},
\] 

and \( \eta_2 \) is a \( d \times d \) matrix. Furthermore, a more general “extended” affine specification has been proposed by Cheridito et al. (2007) as

\[
\Lambda_t^{EX} = \left( \sqrt{S_t} \right)^{-1} \eta \begin{bmatrix} 1 \\ X_t \end{bmatrix}
\] (2.11)

where \( \eta \) is a \( d \times (d + 1) \) matrix and for the maximum \( k \) such that \( \inf (X_{t,k+1}) \geq 0 \), \( \eta_{ij} = 0 \) for \( i \leq k \) and \( j > k + 1 \).

For the market prices of jump risks, we follow Pan (2002) and Jarrow, Li, and Zhao (2007) to first specify the dynamics of \( X_t \) under the physical measure and then take the difference in the jump dynamics under the risk-neutral and physical measures as the jump risk premium. Under both
specifications of $\Lambda_t$ in (2.10) and (2.11), we assume that $X_t$ is still affine under the physical measure:

$$
\text{d}X_t = \bar{\kappa} \left( \bar{\theta} - X_t \right) \text{d}t + \kappa \sqrt{S_t} \text{d}W_t + \sqrt{S_t} \Lambda_t \text{d}t + \text{d}J_t,
$$

(2.12)

where $W_t$ is a $d \times 1$ standard Brownian Motion under the physical measure $\mathcal{P}$, $J_t$ is a pure jump process with the jump arrival intensity $\lambda_0 > 0$ with a $d$-dimensional random jump size vector $\xi$ having the mean $\mu(X_t) = \begin{bmatrix} \mu_{10} & \mu_{11} \\ \vdots & \vdots \\ \mu_{d0} & \mu_{d1} \end{bmatrix} \begin{bmatrix} 1 & \cdots & 1 \\ X_{1t} & \cdots & X_{dt} \end{bmatrix}$, and $(\lambda_0 \mu) \text{d}t$ is the compensator for the pure jump process $J_t$ under the physical measure. The third term in (2.13), $\sqrt{S_t} \Lambda_t \text{d}t$, represents the market price of diffusive risk.

Comparing the physical dynamics of $X_t$ in (2.13) with the risk-neutral dynamics in (2.2), we see that by allowing the risk-neutral mean relative jump size $\mu^Q$ to differ from its physical counterpart $\mu$, a premium is assigned for jump size uncertainty. Similarly, a premium is also accommodated for jump timing risk by permitting $\lambda_0^Q$ in the risk-neutral jump-arrival intensity to differ from their physical counterparts $\lambda_0$. Therefore, the time-$t$ compensation for jump risk is $\lambda_0 \mu (X_t) - \lambda_0^Q \mu^Q (X_t)$. In contrast, Piazzesi (2005) sets the market prices of jump timing uncertainty as zero due to data limitations (only five years of data).

### 2.2 Contribution of Jump Risk Premiums

By incorporating jumps in the state variables, we accomplish two things. First, $\text{d}J_t$ captures well-documented discontinuous movements in $r_t$ through jumps in $X_t$. Second, jump risk premiums provide additional flexibility in modeling bond risk premium without restricting the conditional volatility of bond yields. In this section, we provide theoretical analysis on the advantages coming from the jump dynamics. In particular, we borrow the argument of Joslin, Le and Singleton (2011), Joslin, Singleton, and Zhu (2011), and especially Joslin and Le (2012) to show explicitly how jump dynamics help the model to capture both expectation puzzles and conditional volatility of yields.\(^7\)

Without loss of generality, we consider a three-factor AJTSM with one volatility vector that has

\(^7\)We are very grateful to an anonymous referee for referring us to this line of research.
the state variable \( X_t = (X_{1t}, X_{2t}, X_{3t})' \) and \( X_{1t} \) is the square root factor. The risk-neutral and physical dynamics of \( X_t \) are specified as (2.2) and (2.12), respectively. Such a model reduces to a three-factor ATSM with one volatility factor if we remove the jump dynamics:

\[
dX_t = \frac{\kappa}{\frac{\kappa}{2} + \rho} \left( \frac{\kappa}{\frac{\kappa}{2} + \rho} - X_t \right) dt + \sqrt{\frac{\kappa}{\frac{\kappa}{2} + \rho}} dW_t,
\]

(2.13)

which covers all three-factor ATSMs with one volatility factor including "completely affine", "essentially affine", and "extended affine" models depending on \( \Lambda_t \).

According to Joslin, Singleton, and Zhu (2011) and Joslin and Le (2012), the model (2.13) can be restated using observed yields. In particular, let the yields \( y_t = A_z + B_z X_t \), where \( A_z \) and \( B_z \) are conformable matrices. Define further the principle components of yields as \( P_t = W y_t = W A_x + W B_x X_t \). Assuming \( W B_x \) is of full rank, the model (2.13) can be represented as

\[
dP_t = (\kappa_0 + \kappa_1 P_t) dt + \sqrt{\Sigma_0 + \Sigma_1 P_t} dW_t,
\]

(2.14)

As a result, the yield volatility is \( V_t = (dP_t)^2 = \alpha + \beta P_t \), i.e., a generic linear function of \( P_t \); see Joslin and Le (2012, equation (2)). Moreover, we have

\[
E_t [P_{t+1}] = K_0 + K_1 P_t,
\]

(2.15)

\[
E_t^Q [P_{t+1}] = K_0^Q + K_1^Q P_t,
\]

where \( K_0 \) and \( K_1 \) are functions of \( \kappa_0, \kappa_1, \Sigma_0, \) and \( \Sigma_1 \), and \( K_0^Q \) and \( K_1^Q \) are functions of \( \kappa_0^Q, \kappa_1^Q, \Sigma_0, \) and \( \Sigma_1 \). Therefore, the conditional forecast of the yield volatility is

\[
E_t [V_{t+1}] = \alpha + \beta K_0 + \beta K_1 P_t
\]

\[
E_t^Q [V_{t+1}] = \alpha + \beta K_0^Q + \beta K_1^Q P_t
\]

\footnote{The specification (2.14) is essentially equivalent to equation (8) of Joslin and Le (2012) that has the square factor \( X_{1t} \) in the diffusion term because \( X_{1t} \) is a linear function of \( P_t \) in affine models.}
To ensure the model is admissible, we need $V_{t+1}$ to be autonomous, i.e., the conditional forecast of yield volatility can only depend on its lagged value:

$$
\beta K_1 = a \beta \text{ and } \beta K_Q^1 = a^Q \beta \text{ for some positive } a \text{ and } a^Q.
$$

As a result, $K_Q^1$ and $K_1$ must share the same left eigenvector. As shown in Joslin and Le (2011), $K_Q^1$ is strongly identified by the cross section of bond yields and hence the admissibility conditions of ATSMs impose $K_1$ to have a pre-specified left eigenvector that may be very different from the one obtained by only estimating the physical dynamics. This restriction illustrates the trade-off of ATSMs in capturing risk premiums and conditional volatility dynamics simultaneously, which further explains the documented empirical failure of ATSMs with volatility factors in capturing expectation puzzles (Dai and Singleton, 2002; Duffee (2002). That is, more factors driving conditional volatility dynamics lead to more restricted specifications of risk premiums.

By similar arguments leading to (2.14), our AJTSM can be represented as

$$
dP_t = (\kappa_0 + \kappa_1 P_t) dt + \sqrt{\Sigma_0 + \Sigma_1 P_t} dW_t + dJ_t
$$

$$
dP_t = (\kappa_Q^0 + \kappa_Q^1 P_t) dt + \sqrt{\Sigma_0 + \Sigma_1 P_t} dW_Q^t + dJ_Q^t
$$

where the $P_t$ dynamics still follow (2.15) but $K_1$ and $K_Q^1$ incorporate the contribution of $dJ_t$ and $dJ_Q^t$ to yields now and the yield volatility $V_t = \alpha + \beta P_t + J_t^v$, with $J_t^v$ as the volatility component due to jumps in yields satisfying

$$
E_t [J_{t+1}^v] = \gamma P_t, \text{ and } E_t^Q [J_{t+1}^v] = \gamma^Q P_t, \quad (2.16)
$$

given our assumption about the affine jump dynamics.\(^9\) Therefore, the conditional forecast of yield

\(^9\)Some quadratic terms of $P_t$ due to jumps are omitted in (2.16) as they are restricted to be positive automatically, and hence we mainly focus on the linear terms.
volatility for AJTSMs is

\[
E_t [V_{t+1}] = \alpha + \beta K_0 + \beta K_1 P_t + \beta \gamma P_t,
\]

\[
E_t^Q [V_{t+1}] = \alpha + \beta K_0^Q + \beta K_1^Q P_t + \beta \gamma^Q P_t.
\]

Now the admissibility conditions we need are

\[\beta (K_1 + \gamma) = a\beta \quad \text{and} \quad \beta (K_1^Q + \gamma^Q) = a^Q \beta \] for some positive \(a\) and \(a^Q\).

As the difference between \(\gamma\) and \(\gamma^Q\) represents the jump risk premium that is usually large (see Section 5 for some estimates of jump parameters), our AJTSM is not subject to the restriction that \(K_1^Q\) and \(K_1\) must share the same left eigenvector. In fact, \(K_1\) should be different from \(K_1^Q\) that is identified strongly by the cross section of yields. This is exactly the key that allows AJTSMs to simultaneously match time variations in both the risk premium and conditional volatility.

### 2.3 Canonical Three-Factor AJTSMs

Dai and Singleton (2000) impose admissibility conditions on the model parameters to ensure the existence of the process \(X_t\). They show that there exist \(d + 1\) disjoint admissible regions of the parameter space for each \(d\). Similarly, with \(d\) factors, there are \(d + 1\) non-nested families of AJTSMs corresponding to \(M = 0, 1, \ldots, d\) and denoted as \(\text{AJ}_{M}(d)\). We follow Dai and Singleton (2000) to consider the canonical representation for each family of AJTSMs with the matrix \(\Sigma\) normalized as the identity matrix.

We now present the canonical representations of \(d\)-dimensional “essentially” AJTSMs. We shall classify all admissible three-factor “essentially” AJTSMs (denoted as \(\text{EAJ}_{M}(3)\)) into maximal subfamilies, which are three-factor “essentially” ATSMs (denoted as \(\text{EA}_{M}(3)\)) of Duffee (2002) augmented by jumps. For all subfamilies, the sport interest rate is specified as

\[r_t = \delta_0 + \delta_{11} X_{1t} + \delta_{12} X_{2t} + \delta_{13} X_{3t} \]

We focus on the specifications of jump processes.
• EAJ0(3): In this model, the dynamics of $X_t$ under the physical measure are given as

$$
\begin{aligned}
d \begin{bmatrix}
X_{1t} \\
X_{2t} \\
X_{3t}
\end{bmatrix} &=
\begin{bmatrix}
\kappa_{11} & 0 & 0 \\
\kappa_{21} & \kappa_{22} & 0 \\
\kappa_{31} & \kappa_{32} & \kappa_{33}
\end{bmatrix}
\begin{bmatrix}
-X_{1t} \\
-X_{2t} \\
-X_{3t}
\end{bmatrix}
dt +
\begin{bmatrix}
dW_{1t} \\
dW_{2t} \\
dW_{3t}
\end{bmatrix}
+dJ_t,
\end{aligned}
\tag{2.17}
$$

where $J_t$ is a three-dimensional pure jump process, with the jump arrival intensity $\lambda(X_t) = \lambda_0 > 0$ and the random jump size vector $\xi$ having the mean vector $\mu = (\mu_{10} + \mu_{11}X_{1t}, \mu_{20} + \mu_{21}X_{2t}, \mu_{30} + \mu_{31}X_{3t})'$. Since $M = 0$, none of the elements in $X_t$ affect the volatility of $X_t$ and hence the state variables are homoskedastic. The corresponding risk neutral dynamics for EAJ0(3) are

$$
\begin{aligned}
d \begin{bmatrix}
X_{1t} \\
X_{2t} \\
X_{3t}
\end{bmatrix} &=
\begin{bmatrix}
\kappa_{11} & 0 & 0 \\
\kappa_{21} & \kappa_{22} & 0 \\
\kappa_{31} & \kappa_{32} & \kappa_{33}
\end{bmatrix}
\begin{bmatrix}
-X_{1t} \\
-X_{2t} \\
-X_{3t}
\end{bmatrix}
dt -
\begin{bmatrix}
\eta_{11} \\
\eta_{12} \\
\eta_{13}
\end{bmatrix}
dt
-
\begin{bmatrix}
\eta_{21,11} & \eta_{21,12} & \eta_{21,13} \\
\eta_{22,11} & \eta_{22,12} & \eta_{22,13} \\
\eta_{23,11} & \eta_{23,12} & \eta_{23,13}
\end{bmatrix}
\begin{bmatrix}
X_{1t} \\
X_{2t} \\
X_{3t}
\end{bmatrix}
dt +
\begin{bmatrix}
dW_{1t}^Q \\
dW_{2t}^Q \\
dW_{3t}^Q
\end{bmatrix}
+dJ_t^Q,
\end{aligned}
\tag{2.18}
$$

where $J_t^Q$ is a 3-dimensional pure jump process, with the jump arrival intensity $\lambda^Q(X_t) = \lambda_0^Q > 0$ and the jump size $\xi^Q$ having a mean $\mu^Q = (\mu_{10}^Q + \mu_{11}^Q X_{1t}, \mu_{20}^Q + \mu_{21}^Q X_{2t}, \mu_{30}^Q + \mu_{31}^Q X_{3t})'$. Without jumps in (2.17)-(2.18), we get the physical and risk-neutral dynamics of the EA0(3) model. The corresponding models with the “extended” affine specification $\Lambda_t^{EX}$ coincides with EA0(3).
• EAJ\(_1\)(3): In this model, the dynamics of \(X_t\) under the physical measure are given as

\[
\begin{bmatrix}
X_{1t} \\
X_{2t} \\
X_{3t}
\end{bmatrix} =
\begin{bmatrix}
\kappa_{11} & 0 & 0 \\
\kappa_{21} & \kappa_{22} & \kappa_{23} \\
\kappa_{31} & \kappa_{32} & \kappa_{33}
\end{bmatrix}
\begin{bmatrix}
\theta_{1} - X_{1t} \\
-X_{2t} \\
-X_{3t}
\end{bmatrix} dt
+ \begin{bmatrix}
\sqrt{X_{1t}} \\
\sqrt{1 + \beta_{12} X_{1t}} \\
\sqrt{1 + \beta_{32} X_{1t}}
\end{bmatrix}
\begin{bmatrix}
dW_{1t} \\
dW_{2t} \\
dW_{3t}
\end{bmatrix}
+ dJ_t,
\tag{2.19}
\]

where \(J_t\) is a three-dimensional pure jump process, with the jump arrival intensity \(\lambda(X_t) = \lambda_0 + \lambda_{11} X_{1t}\) for \((\lambda_0, \lambda_{11}) > 0\) and the jump size \(\xi\) having a mean vector \(\mu = (\mu_{10} + \mu_{11} X_{1t}, \mu_{20} + \mu_{21} X_{2t}, \mu_{30} + \mu_{31} X_{3t})'\) for \((\mu_{10}, \mu_{11}) > 0\).

Since \(M = 1\), the first element of the state variable \(X_t, X_{1t}\), determines the conditional volatility of all three state variables. The corresponding risk-neutral dynamics for EAJ\(_1\)(3) are

\[
\begin{bmatrix}
X_{1t} \\
X_{2t} \\
X_{3t}
\end{bmatrix} =
\begin{bmatrix}
\kappa_{11} & 0 & 0 \\
\kappa_{21} & \kappa_{22} & \kappa_{23} \\
\kappa_{31} & \kappa_{32} & \kappa_{33}
\end{bmatrix}
\begin{bmatrix}
\theta_{1} - X_{1t} \\
-X_{2t} \\
-X_{3t}
\end{bmatrix} dt - \begin{bmatrix}
X_{1t} \eta_{11} \\
(1 + \beta_{12} X_{1t}) \eta_{12} \\
(1 + \beta_{32} X_{1t}) \eta_{13}
\end{bmatrix} dt
+ \begin{bmatrix}
dW_{1t}^Q \\
dW_{2t}^Q \\
dW_{3t}^Q
\end{bmatrix}
+ dJ_t^Q,
\tag{2.20}
\]

where \(J_t^Q\) is a 3-dimensional pure jump process, with the jump arrival intensity \(\lambda^Q(X_t) = \lambda_0^Q + \lambda_{11}^Q X_{1t}\) for \((\lambda_0^Q, \lambda_{11}^Q) > 0\) and the jump size \(\xi^Q\) having a mean vector \(\mu^Q=(\mu_{10}^Q + \mu_{11}^Q X_{1t}, \mu_{20}^Q + \mu_{21}^Q X_{2t}, \mu_{30}^Q + \mu_{31}^Q X_{3t})'\) for \((\mu_{10}^Q, \mu_{11}^Q) > 0\). Without jumps in both (2.19) and (2.20), we get the physical and risk-neutral dynamics of the EAJ\(_1\)(3) model. By adding a term \(\begin{bmatrix}
\gamma_{10} \\
0 \\
0
\end{bmatrix} dt\) into the physical dynamics
of $\text{EA}_1(3)$, we obtain the corresponding “extended” affine model $\text{EXA}_1(3)$.

- $\text{EAJ}_2(3)$: In this model, the dynamics of $X_t$ under the physical measure are given as

$$
\begin{align*}
\frac{d}{dt} & \begin{bmatrix}
X_{1t} \\
X_{2t} \\
X_{3t}
\end{bmatrix} = \\
& \begin{bmatrix}
\kappa_{11} & \kappa_{12} & 0 \\
\kappa_{21} & \kappa_{22} & 0 \\
\kappa_{31} & \kappa_{32} & \kappa_{33}
\end{bmatrix} \begin{bmatrix}
\theta_1 - X_{1t} \\
\theta_2 - X_{2t} \\
-X_{3t}
\end{bmatrix} dt + \\
& \begin{bmatrix}
\sqrt{X_{1t}} \\
\sqrt{X_{2t}} \\
\sqrt{1 + \beta_{31}X_{1t} + \beta_{32}X_{2t}}
\end{bmatrix} \begin{bmatrix}
dW_{1t} \\
dW_{2t} \\
dW_{3t}
\end{bmatrix} + dJ_t, 
\end{align*}
$$

where $J_t$ is a three-dimensional pure jump process, with the jump arrival intensity $\lambda(X_t) = \lambda_0 + \lambda_{11}X_{1t} + \lambda_{12}X_{2t}$ for $(\lambda_0, \lambda_{11}, \lambda_{12}) > 0$ and the jump size $\xi$ having a mean vector $\mu = (\mu_{10} + \mu_{11}X_{1t}, \mu_{20} + \mu_{21}X_{2t}, \mu_{30} + \mu_{31}X_{3t})'$ for $(\mu_{10}, \mu_{11}, \mu_{20}, \mu_{21}) > 0$.

Since $M = 2$, the first two elements of the state variable $X_t$, $X_{1t}$ and $X_{2t}$, determine the conditional volatility of all three state variables. The corresponding risk-neutral dynamics for $\text{EAJ}_2(3)$ are

$$
\begin{align*}
\frac{d}{dt} & \begin{bmatrix}
X_{1t} \\
X_{2t} \\
X_{3t}
\end{bmatrix} = \\
& \begin{bmatrix}
\kappa_{11} & \kappa_{12} & 0 \\
\kappa_{21} & \kappa_{22} & 0 \\
\kappa_{31} & \kappa_{32} & \kappa_{33}
\end{bmatrix} \begin{bmatrix}
\theta_1 - X_{1t} \\
\theta_2 - X_{2t} \\
-X_{3t}
\end{bmatrix} dt - \\
& \begin{bmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
\eta_{2,31} & \eta_{2,32} & \eta_{2,33}
\end{bmatrix} \begin{bmatrix}
X_{1t} \\
X_{2t} \\
X_{3t}
\end{bmatrix} dt + \\
& \begin{bmatrix}
\sqrt{X_{1t}} \\
\sqrt{X_{2t}} \\
\sqrt{1 + \beta_{31}X_{1t} + \beta_{32}X_{2t}}
\end{bmatrix} \begin{bmatrix}
dW_{1t}^\infty \\
dW_{2t}^\infty \\
dW_{3t}^\infty
\end{bmatrix} + dJ_t^\infty, 
\end{align*}
$$

where $J_t^\infty$ is a three-dimensional pure jump process, with the jump arrival intensity $\lambda^\infty(X_t) = \lambda_0^\infty + \lambda_{11}^\inftyX_{1t} + \lambda_{12}^\inftyX_{2t}$ for $(\lambda_0^\infty, \lambda_{11}^\infty, \lambda_{12}^\infty) > 0$ and the jump size $\xi^\infty$ having a mean vector $\mu^\infty = (\mu_{10}^\infty + \mu_{11}^\inftyX_{1t}, \mu_{20}^\infty + \mu_{21}^\inftyX_{2t}, \mu_{30}^\infty + \mu_{31}^\inftyX_{3t})'$ for $(\mu_{10}^\infty, \mu_{11}^\infty, \mu_{20}^\infty, \mu_{21}^\infty) > 0$. Without jumps in both (2.21) and (2.22), we get
the physical and risk-neutral dynamics of the $EA_2(3)$ model. By adding a term \[
\begin{bmatrix}
\gamma_{10} + \gamma_{12}X_{2t} \\
\gamma_{20} + \gamma_{21}X_{1t} \\
0
\end{bmatrix}
dt
\]
into $EA_2(3)$, we obtain the corresponding “extended” affine model $EXA_2(3)$.

- $EAJ_3(3)$: In this model, the dynamics of $X_t$ under the physical measure are given as

\[
d\begin{bmatrix}
X_{1t} \\
X_{2t} \\
X_{3t}
\end{bmatrix} =
\begin{bmatrix}
\kappa_{11} & \kappa_{12} & \kappa_{13} \\
\kappa_{21} & \kappa_{22} & \kappa_{23} \\
\kappa_{31} & \kappa_{32} & \kappa_{33}
\end{bmatrix}
\begin{bmatrix}
\theta_1 - X_{1t} \\
\theta_2 - X_{2t} \\
\theta_3 - X_{3t}
\end{bmatrix}
dt + \begin{bmatrix}
\sqrt{X_{1t}} \\
\sqrt{X_{2t}} \\
\sqrt{X_{3t}}
\end{bmatrix}
\begin{bmatrix}
dW_{1t} \\
dW_{2t} \\
dW_{3t}
\end{bmatrix}
+ dJ_t,
\]

where $J_t$ is a three-dimensional pure jump process, with the jump arrival intensity $\lambda(X_t) = \lambda_0 + \lambda_{11}X_{1t} + \lambda_{12}X_{2t} + \lambda_{13}X_{3t}$ for $(\lambda_0, \lambda_{11}, \lambda_{12}, \lambda_{13}) > 0$ and the jump size $\xi$ having a mean vector $\mu = (\mu_{10} + \mu_{11}X_{1t}, \mu_{20} + \mu_{21}X_{2t}, \mu_{30} + \mu_{31}X_{3t})$ for $(\mu_{10}, \mu_{11}, \mu_{20}, \mu_{21}, \mu_{30}, \mu_{31}) > 0$.

Since $M = 3$, all three components of the state variable $X_t$ enter the conditional volatility of the state variables. The risk-neutral dynamics for $EAJ_3(3)$ are

\[
d\begin{bmatrix}
X_{1t} \\
X_{2t} \\
X_{3t}
\end{bmatrix} =
\begin{bmatrix}
\kappa_{11} & \kappa_{12} & \kappa_{13} \\
\kappa_{21} & \kappa_{22} & \kappa_{23} \\
\kappa_{31} & \kappa_{32} & \kappa_{33}
\end{bmatrix}
\begin{bmatrix}
\theta_1 - X_{1t} \\
\theta_2 - X_{2t} \\
\theta_3 - X_{3t}
\end{bmatrix}
dt - \begin{bmatrix}
X_{1t}\eta_{11} \\
X_{2t}\eta_{12} \\
X_{3t}\eta_{13}
\end{bmatrix}
dt
\]

\[
+ \begin{bmatrix}
\sqrt{X_{1t}} \\
\sqrt{X_{2t}} \\
\sqrt{X_{3t}}
\end{bmatrix}
\begin{bmatrix}
dW_{1t}^Q \\
dW_{2t}^Q \\
dW_{3t}^Q
\end{bmatrix}
+ dJ_t^Q,
\]

where $J_t^Q$ is a 3-dimensional pure jump process, with the jump arrival intensity $\lambda^Q(X_t) = \lambda_{0}^Q + \lambda_{11}^QX_{1t} + \lambda_{12}^QX_{2t} + \lambda_{13}^QX_{3t}$ for $(\lambda_{0}^Q, \lambda_{11}^Q, \lambda_{12}^Q, \lambda_{13}^Q) > 0$ and the jump size $\xi^Q$ having a mean vector $\mu^Q=(\mu_{10}^Q + \mu_{11}^QX_{1t}, \mu_{20}^Q + \mu_{21}^QX_{2t}, \mu_{30}^Q + \mu_{31}^QX_{3t})$ for $(\mu_{10}^Q, \mu_{11}^Q, \mu_{20}^Q, \mu_{21}^Q, \mu_{30}^Q, \mu_{31}^Q) > 0$. Without jumps in both (2.23) and (2.24), we get the physical and risk-neutral dynamics of the $EA_3(3)$ model. By
adding a term 
\[ \sum_{i=1}^{d} \left( \gamma_{10} + \gamma_{12}X_{2i} + \gamma_{13}X_{3i} \right) dt \]
into the physical dynamics of EA$_3$(3), we obtain the corresponding “extended” affine model EXA$_3$(3).

3 Estimation Methods

We estimate all the models by the infinitesimal operator-based approach proposed in Song (2010, 2011), which enjoys computational convenience as the infinitesimal operator is available in close form for all the continuous-time models. For simplicity, we only consider models without latent variables here. Similar procedures to those in Ait-Sahalia and Kimmel (2010) can be applied to deal with the latent factors. Consider a $d$-dimensional jump diffusion model defined on the probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$,

\[ dX_t = b(X_t; \theta)dt + \sigma(X_t; \theta)dW_t + dJ_t, \tag{3.1} \]

where $W_t$ is a $d \times 1$ standard Brownian motion and $\Theta \subset \mathbb{R}^p$ is a finite-dimensional parameter space. The jump process $J_t$ is of a Poisson-type with the intensity $\lambda(X_t, \theta)$ and the jump size $\xi_t$, which is independent of $\mathcal{F}_t$ and has probability density $\nu(\cdot, \theta) : \mathbb{R}^d \rightarrow \mathbb{R}$.

To characterize the model in (3.1), we employ the infinitesimal operator defined as (Revuz and Yor, 2005):

\[ \mathcal{A}f(x) = \sum_{i=1}^{d} b_i(x; \theta) f_i'(x) + \frac{1}{2} \sum_{i,j=1}^{d} a_{ij}(x; \theta) f_{ij}''(x) + \lambda(x, \theta) \int [f(x + c) - f(x)] d\nu(c, \theta), \tag{3.2} \]

where $x \in \mathbb{R}^d$, $f \in D(\mathcal{A})$, the domain of $\mathcal{A}$, and

\[ a_{ij}(x; \theta) = \sum_{k=1}^{d} \sigma_{i,k}(x; \theta) \sigma_{j,k}(x; \theta). \tag{3.3} \]

From (3.2)-(3.3), it can be seen that the infinitesimal operator is always analytic. Furthermore, it fully characterizes the dynamics of the model and is equivalent to the transition density; see Song.
To utilize the infinitesimal operator for econometric inferences, we consider a transformation based on the celebrated “martingale problem,” which is defined as follows (Karatzas and Shreve, 1991): A probability measure $\mathbb{P}$ under which

$$M_t^f = f(X_t) - f(X_0) - \int_0^t (Af)(X_s)ds$$

is a martingale for every $f \in D(A)$ (3.4) is called a solution to the martingale problem associated with the operator $A$. This “martingale problem” is a variation of the weak solution to the stochastic differential equation in (3.1) and hence can be employed as the identification condition (Revuz and Yor, 2005). The martingale property of the transformed processes in (3.4) can be written as a conditional moment restriction:

$$E[hM_t^f(X) | \mathcal{I}_t] = M_t^f(X)$$

for any $f \in D(A)$ and $t' < t$,

where $\mathcal{I}_t = \sigma\{X_{t'} \}_{t' < t}$ is the information set generated by the history of $\{X_t\}$ at time $t'$.

For econometric convenience, the following equivalent conditional moment restrictions can be derived by the martingale difference ($m.d.s.$) property of the first-order difference of the transformed process $M_t^f(\theta)$:

$$E[Z_t^f(\theta)| \mathcal{I}_{t'}] = 0 \text{ for any } t' < t,$$ (3.5)

where $\mathcal{I}_{t'} = \sigma\{X_{t'} \}_{t' < t'}$, $Z_t^f(\theta) = M_t^f(\theta_0) - M_{t-\Delta}^f(\theta)$, and $\Delta$ is the sampling interval. By the Markov property, (3.5) is further equivalent to

$$E[Z_t^f(\theta)|X_{t-\Delta}] = 0, \text{ for any } f \in D(A),$$ (3.6)

There are two types of solutions to a stochastic differential equation, the strong solution and the weak solution. Loosely speaking, the difference between strong and weak solutions, intuitively, is very similar to that between a random variable and its distribution. Since econometric inferences are concerned only with the dynamic probability laws of the process instead of specific sample paths, it is sufficient to consider a weak solution.
which further delivers the moment condition we shall depend on

$$E\left[ g(Y_t, \theta_0) \right] = 0 \text{ for some } \theta_0 \in \Theta$$

(3.7)

where \( g(Y_t, \theta) = Z_t^f(\theta) h(X_{t-\Delta}) \), \( Y_t \) contains the elements of \( X_{t-\Delta} \) and \( X_t \) and \( h(\cdot) \) is a moment choice function. Finally, we use the exponential function \( f(x) = \exp\left[-\left(x_1^2 + \cdots + x_n^2\right)/2\right] \) in (3.7).\(^{11}\)

This delivers the moment condition (3.7) with

$$Z_t^f(\theta) = e^{-\left(X_{t,s}^2 + \cdots + X_{d,s}^2\right)/2} - e^{-\left(X_{t,t-\Delta}^2 + \cdots + X_{d,t-\Delta}^2\right)/2} - \int_{t-\Delta}^t A_\theta e^{-\left(X_{t,s}^2 + \cdots + X_{d,s}^2\right)/2} ds,$$

(3.8)

where

$$A_\theta e^{-\left(X_{t,s}^2 + \cdots + X_{d,s}^2\right)/2} = e^{-\left(X_{t,s}^2 + \cdots + X_{d,s}^2\right)/2} \left\{ - \sum_{i=1}^d b_i(X_s; \theta) X_i,s + \frac{1}{2} \sum_{i,j=1}^d a_{i,j}(X_s; \theta) X_i,s X_j,s \right.\right.$$  

$$\left. - \frac{1}{2} \sum_{i,j=1}^d a_{i,j}(X_s; \theta) + \lambda(X_s, \theta) \int [e^{-c \cdot X_s} - |c|^2/2 - 1] d\nu(c, \theta) \right\}.$$  

(3.9)

It is clear that the moment restrictions are expressed explicitly by the drift, volatility, and jump terms and can be used directly. Hence, the infinitesimal operator methods are particularly convenient for multivariate models, for which the transition density methods are extremely complicated and computationally challenging. In contrast, the transition density-based methods of Lo (1988) and Aït-Sahalia (2002, 2008) have to approximate the transition density or numerically solve it. Moreover, they are not applicable to jump models but only to diffusion processes.

4 Data and Stylized Facts

4.1 Data

Though previous studies on the “expectation hypothesis,” such as Fama and Bliss (1987) and Campbell and Shiller (1991), have used monthly U.S. Treasury yields, we follow Dai and Singleton (2000, 2001) and Song (2011) discussed choices of test functions for diffusion models.

\(^{11}\)Conly et al. (1997) and Song (2011) discussed choices of test functions for diffusion models.
2003) and Piazzesi (2005) to focus on Libor swap rates for which daily data are readily available. The data contain daily Libor swap rates with 3-month, 6-month, 9-month, 2-year, 3-year, 4-year, 5-year, 7-year, and 10-year maturities, obtained from Datastream. The sample period is from August 13, 1990, to December 31, 2008, with a total of 4,757 observations.

Figure 1 provides time series plots of the Libor swap rates, and Figure 2 contains daily changes of 2-year and 10-year swap rates from 2006 to 2008. We see relatively infrequent but fairly large spikes especially in 2008 even for the long-term yield (10-year swap rate), which we interpret as jumps. As suggested by Johannes (2004), these large spikes are potentially related to the arrival of significant new information regarding the current or future state of the economy.

Table 1 reports summary statistics of the levels and changes of Libor swap rates. We can see that the average yield curve is upward sloping with long-term yields higher than short-term yields. On average, all yields exhibit negative changes, consistent with declining interest rates during the sample period. The standard deviation of yield changes increases with maturity up to nine months and then decreases with maturity. Changes of short-term yields exhibit higher kurtosis and are more negatively skewed than that of long-term yields. Yield levels are very persistent, with first-order autoregressive coefficients close to one, while yield changes are much less persistent, with first-order autoregressive coefficients ranging from -0.1134 to 0.0424. Principal component analysis shows that the first three principal components can explain more than 99.9% and 96.9% of the variations of the levels and changes of the yields, respectively. These results confirms the claim of Dai and Singleton (2000) that Treasury yields and Libor swap rates share similar distributional characteristics even though the institutional structures of the two markets are quite different.

4.2 Expectation Puzzles for Libor swap Rates

In this section, we document “expectation puzzles” for Libor swap rates. To fix notations, we use $P_t^\tau$ to denote the time-$t$ price of a zero-coupon bond with $\tau$ periods to maturity, $y_t^\tau (\equiv - \ln P_t^\tau / \tau)$ to denote its corresponding continuously compounded yield, and $r_t \equiv y_t^1$ to denote the short-term
interest rate. We define one-period expected excess holding period returns as

\[ e_t^\tau \equiv E_t \left[ \ln \left( \frac{P_{t+1}^{\tau-1}}{P_t^\tau} \right) \right] - r_t. \]  

(4.1)

Then by definition we have

\[ e_t^\tau / (\tau - 1) + E_t \left[ y_{t+1}^{\tau-1} - y_t^\tau \right] = (y_t^\tau - r_t) / (\tau - 1), \]  

(4.2)

where \( E_t \left[ y_{t+1}^{\tau-1} - y_t^\tau \right] \) is the expected one-period yield change and \( (y_t^\tau - r_t) / (\tau - 1) \) is the average yield spread (also called the slope of the term structure).

One way to characterize the expectation puzzle is the “yield regression” used in Campbell and Shiller (1991) and Backus et al. (2001):

\[ y_t^{(\tau-1)} - y_t^\tau = constant + \phi_\tau \left( \frac{y_t^\tau - r_t}{\tau - 1} \right) + residual, \]  

(4.3)

which is termed LPY (i) (linear projection of yields) in Dai and Singleton (2002). If the “expectation hypothesis” holds in the generalized sense of \( e_t^\tau \equiv constant \), we should have \( \phi_\tau = 1 \) for all \( \tau \). Any deviation from the benchmark of \( \phi_\tau = 1 \) for all \( \tau \) implies that the expected excess return, that is, the risk premium, is time-varying, and the empirical pattern of the deviation reflects the nature of the time variation in the risk premium.

Another characterization of the expectations puzzle is through “risk-adjusted yield regression”:

\[ y_{t+1}^{(\tau-1)} - y_t^\tau + \frac{e_t^\tau}{\tau - 1} = constant + \phi_\tau^R \left( \frac{y_t^\tau - r_t}{\tau - 1} \right) + residual, \]  

(4.4)

which is termed LPY (ii) in Dai and Singleton (2002). That is, if a term structure model can really capture bond risk premiums, then by adding model implied risk premium \( e_t^\tau \) to the above regression,

\[ e_t^\tau \equiv \left( y_t^\tau - \frac{1}{\tau} \sum_{i=0}^{\tau-1} E_t \left[ \ln \left( \frac{P_{t+i}^{\tau-1}}{P_t^\tau} \right) \right] \right) \]  

forward term premiums \( p_t^\tau \equiv f_t^\tau - E_t \left[ r_{t+1} \right] \), where \( f_t^\tau \equiv -\ln \left( \frac{P_{t+1}^{\tau+1}}{P_t^\tau} \right) \) is the forward rate), which lead to a range of empirical predictive regressions for testing the “expectation hypothesis”; for details, see Fama (1984a, b, 2006), Fama and Bliss (1987), Stambaugh (1988), Campbell and Shiller (1991), Backus et al. (2001), Bekaert et al. (1997, 2001), Bekaert and Hodrick (2001), and Dai and Singleton (2002). Moreover, Chernov and Mueller (2012) recently studied the term structure of inflation expectations rather than only the nominal yields.
one should get $\phi^R_\tau$ close to unity. In some sense, LPY (i) points to the historical behavior of term structure dynamics under the physical measure $P$, while LPY (ii) focuses on matching term structure dynamics under the risk-neutral measure $Q$. When evaluating model performance, both measures should be used.

We report the estimates $\phi_{rT}$ of $\phi_\tau$ in the “yield regression” of (4.3) using Libor swap rates\(^{13}\) in Table 2 and Figure 3, where the three-month Libor rate represents the spot rate $r_t$.\(^{14}\) We see that the estimated coefficients $\phi_{rT}$ are close to zero for all $\tau$ and decrease with maturity, changing from slightly positive values for maturities less than two years to slightly negative values for maturities longer than two years. Though the standard errors are relatively large, the estimated coefficients are still significantly different from one at conventional significance levels. This empirical pattern of violations of the “expectation hypothesis” for Libor swap yields is similar to that for U.S. Treasury yields in terms of the increasingly negative regression coefficients at longer maturities (Dai and Singleton, 2002). They differ, however, insofar as the former has much smaller regression coefficients (from -0.0051 to 0.0095) than the latter (from -0.428 to -4.173) (Dai and Singleton 2002).

4.3 Time-Varying Conditional Volatility for Libor Swap Rates

We document time-varying conditional volatility for Libor swap rates from the following two aspects: (i) the hump-shaped term structure of yield volatility and (ii) the degree of persistence of the conditional volatility of yields. Figure 4 plots the historical unconditional volatility, computed as sample standard deviations of daily log changes of the Libor swap rates, against maturity. Notably, the term structure of historical volatilities is hump shaped with a peak around the nine-month to one-year maturity range. In addition, following Dai and Singleton (2003), we estimate a GARCH (1,1) model for the Libor swap yields. The estimation results in Table 3 show that the coefficient $\beta$, which represents

\(^{13}\)The regression coefficient estimator $\phi_{rT}$ depends only on $\tau$ and not on $T$. We keep the notation as $\phi_{rT}$ to be consistent with Campbell and Shiller (1991) and Dai and Singleton (2002) for comparisons.

\(^{14}\)Following Collin-Dufresne and Goldstein (2002) and Li and Zhao (2006), we assume the quoted swap rate is equivalent to a par-bond rate for an issuer with Libor credit quality. For the “yield regression” in (4.3), we note that zero-coupon yields need to be used. However, the swap yield is equivalent to a par-bond yield with coupon payments. We are grateful to Pamela Moulton for pointing this out. To be consistent with the literature that examines term structure dynamics using swap rates directly (Dai and Singleton, 2000; Piazzesi, 2005), we treat swap rates as approximations of corresponding zero-coupon yields. We have studied the differences between swap rates and constructed zero-coupon yields and found that approximation errors are small and do not affect the testing results of the “expectation hypothesis.” These results are available upon request.
the degree of volatility persistence, is fairly big, ranging from 0.6579 to 0.7960. This confirms the
high degree of time variations and volatility persistence for Libor swap yields at all maturities. The
humped volatility term structure and high volatility persistence are strong evidence of time-varying
conditional volatilities of yields and are treated as a descriptive measure of time variations in yields
volatilities that an empirically successful DTSM should capture.

5 ATSMs vs. AJTSMs: Model Diagnostics

5.1 Parameter Estimates

In this section, we discuss parameter estimates of four groups of three-factor models with increasing
level of complexities in Tables 4 to 6 respectively: “essentially” ATSMs, “extended” ATSMs, and
“essentially” AJTSMs. The models are estimated using daily Libor swap yields with maturities of 6
months, 2, 3, 5, 7, and 10 years, which have the same maturity structure as that of the monthly U.S.
Treasury yields used in Dai and Singleton (2002). We assume that the yields of 6-month, 2-year, and
10-year maturities are measured without error, while those of 3-year, 5-year, and 7-year maturities
are observed with error. Note from (3.7) that moments conditions need be chosen through $h(X_{t-\Delta})$.

As our main objective is to study jumps in dynamics of the state variable $X_t$, we choose moments of
$X_{t-\Delta}$ up to order 4 to include skewness and kurtosis that are affected by jumps substantially.

The parameter estimates in Table 6 show that most of the jump parameters are significantly
different from zero. The estimates of jump intensity parameters $\lambda$'s show that jump arrivals are
state-dependent, suggesting a degree of predictability by current market conditions for the frequency
of future jumps in bond yields. We also find overwhelming evidence of negative jumps in the state
variables under both physical and risk-neutral measures.

Our parameter estimates suggest a positive jump risk premium and negative volatility risk pre-
mium. Take EAJ1(3) as an example. The estimate of $(\mu_{10} - \mu_{10}^Q, \mu_{11} - \mu_{11}^Q)$, risk premiums associated
with the jump size uncertainty for the volatility factor $X_{1t}$, is (0.46%, 0.05%), while the estimate of
$\lambda_0 - \lambda_0^Q$, risk premiums associated with jump arrival uncertainty, is (17.76%, 0.12%). The elements
of the parameters of diffusive risk $\eta_1 = (-7.33\%, -10.01\%, -17.50\%)$ and the non-zero elements
of $\eta_2$, $(\eta_{2,31}, \eta_{2,32}, \eta_{2,33}, \eta_{3,31}, \eta_{3,32}, \eta_{3,33}) = (0.00\%, 0.00\%, -18.66\%, -32.24\%, 0.00\%, 0.00\%)$, imply a negative volatility risk premium, consistent with Dai and Singleton (2000).\textsuperscript{15}

Our parameter estimates answer some of the questions raised by Johannes and Polson (2009). In particular, we show that multiple risk factors do jump and the market prices of diffusive and jumps risks are different: jump risk premiums are positive, while the diffusive risk premium is negative.

5.2 Model Diagnostics

We now consider model diagnostics. In particular, we compare the “extended” ATSMs with AJTSMs by how well they capture the conditional moments of yields. Table 7 reports the mean, volatility, skewness, and kurtosis of the yields conditional on an upward and downward sloping term structure, along with the model-implied conditional moments for the EXA\textsubscript{2}(3) and EAJ\textsubscript{2}(3) models.\textsuperscript{16} We choose the EAJ\textsubscript{2}(3) model as it successfully captures the time-varying risk premiums and conditional volatility of Libor swap rates simultaneously as shown Section 6. The EXA\textsubscript{2}(3) model is used for comparison because of its most general specification for the market price of risks in ATSMs without jumps. Four maturities are considered, including three months, two years, five years, and 10 years. The slope of the term structure, defined as the difference between the 10-year and 3-month yields, is upward if positive and vice versa. To compute model-implied conditional moments, we simulate 1,000 sample paths from the models using estimated parameters in Tables 5, 6, and 7, classify the time series of yields into two subgroups based on the slope of the term structure, compute the four moments of both subgroups for each sample, and finally average the moments of each subgroup across the 1,000 sample paths.

We first observe from Table 7 that the EXA\textsubscript{2}(3) model tends to overestimate the volatility but underestimate the skewness and kurtosis although it matches the mean well. This pattern holds across all the four maturities considered. Moreover, the mismatch is more serious conditional on a downward slope than conditional on an upward slope. For example, the skewness of the 5-year yield

\textsuperscript{15}Following Pan (2002), who studies the jump risk premium using the S&P 500 index and near-the-money short-dated option prices, we set insignificant parameters to zero here and in future simulation studies to reduce the noise in the analysis.

\textsuperscript{16}We choose the slope of the term structure as the conditioning variable following Dai and Singleton (2000). They examine the pricing errors of ATSMs conditional on the slope. Moreover, Pan (2002) uses volatility as the conditioning variable to study how well a jump diffusion model can capture S&P 500 option prices by pricing errors as well.
in the data and implied by the model are 0.0335 versus 0.0165 for an upward slope and 0.1227 versus 0.0334 for a downward slope. Overall, the “extended” ATSMs, without incorporating jump dynamics of the yields, fail to match the conditional high-order moments of the Libor swap rates. In contrast, we find that the EAJ$_2$(3) model performs much better in fitting the conditional moments, especially the skewness and kurtosis for all the four maturities. For example, the model-implied skewness of the 5-year yield from the EAJ$_2$(3) is 0.0222 for an upward slope and 0.0883 for a downward slope, while those of the data are 0.0335 and 0.1227.

Based on a univariate jump-diffusion model for the spot rate, Das and Sundaram (1999) show that jump effects tend to dissipate quickly. Pan (2002) also shows that stochastic volatility models with price jumps cannot capture medium and long-dated options prices. We study how jumps affect the conditional distributions of long-term yields. In particular, the last column of Table 7 reports the absolute difference between the model-implied moments of EAJ$_2$(3) and EXA$_2$(3) standardized by the corresponding moments from the data. The results in Table 7 show that though the differences do decrease as maturity increases for all conditional moments, they remain significant even at 10-year maturity. For example, the difference as a percentage of the corresponding moments in the data is 165.00%, 45.69%, and 33.54% for volatility, skewness, and kurtosis, respectively, under a downward sloping yield curve. Therefore, under our AJTSMs, jumps affect both the short and long-end of the yield curve.

In contrast to the univariate model of Das and Sundaram (1999), in which only the short-term interest rate can jump, our multi-factor AJTSMs allow jumps in all the factors. Furthermore, the affine structure allows the latent factors to be recovered from observable yields, as shown by Joslin, Singleton, and Zhu (2011). In other words, all three-factor AJTSMs can be regarded as building on the dynamics of three individual yields, which are usually chosen as short, medium, and long-term yields in the literature. Therefore, in our multi-factor AJTSMs, we allow persistent jumps in short-, medium-, and long-term yields with state-dependent jump intensity, which lead to persistent jump effects for all maturities. This is consistent with the observed jumps in both short- and long-term yields as shown in Figure 2.
6 ATSMs vs. AJTSMs: Risk Premium and Conditional Volatility

We now evaluate the empirical performance of ATSMs and AJTSMs in capturing time variations in risk premium and conditional volatility of Libor swap rates.

6.1 Evaluation Procedures of Model Performances

For time-varying risk premium, we first compare the empirical pattern of the coefficients of the “yield regression” (4.3) with the model-implied population coefficients

$$\phi_\tau = \frac{\text{cov} \left( y_{t+1}^{(\tau-1)} - y_t^\tau, \frac{(y_t^\tau - r_t)}{(\tau - 1)} \right)}{\text{var} \left( \frac{(y_t^\tau - r_t)}{(\tau - 1)} \right)}, \tag{6.1}$$

which are calculated based on estimated model parameters. The actual yields data do not enter (6.1) directly and show up only through the parameter estimates.\(^{17}\) We also compare the unity line of the “risk-adjusted yield regression” (4.4) with the model-implied coefficients

$$\phi^R_\tau = \frac{\text{cov} \left( y_{t+1}^{(\tau-1)} - y_t^\tau + e_t^\tau / (\tau - 1), \frac{(y_t^\tau - r_t)}{(\tau - 1)} \right)}{\text{var} \left( \frac{(y_t^\tau - r_t)}{(\tau - 1)} \right)}, \tag{6.2}$$

where \(y_{t+1}^{(\tau-1)}\) and \(y_t^\tau\) represent the historical yields and \(e_t^\tau\) are calculated using the estimated model parameters and fitted state variables.

For time-varying conditional volatility, the sample variances, computed using simulated Libor swap rates from the estimated models, are compared with the empirical pattern in Figure 4. We focus on both the humped shape and the levels of volatility (Dai and Singleton, 2000; Piazzesi, 2005). Furthermore, we follow Dai and Singleton (2003) and Buraschi et al. (2008) to estimate a GARCH (1,1) model using simulated yields from different models and check the degree of model-implied time-varying volatility relative to the estimates in Table 3 using actual data. In the simulations, 18 years of daily data are generated to be comparable with the historical data sample. The GARCH

\(^{17}\)Dai and Singleton (2002) find that the model-implied \(\phi_\tau\) computed from the sample of model-implied fitted yields (obtained by inverting the model for the fitted-state variables, computing model-implied fitted zero-coupon bond yields, and then estimating the “yield regression” in (4.3)), can give very misleading conclusions pertaining to the actual population distributions implied by the models. The reason for this is that it mixes the properties of a dynamic term structure model with those of the historical data.
(1,1) parameter estimates are obtained as follows: First, 1,000 sample paths of the yields of a specific maturity are simulated from the models; second, the GARCH (1,1) model is estimated for each of the sample paths; finally, the median of these 1,000 parameter estimates are taken as the model-implied coefficients.

6.2 Performance of Three-Factor ATSMs

6.2.1 “Essentially” ATSMs

Time-Varying Risk Premium. Based on the parameter estimates in Table 4 for three-factor “essentially” ATSMs, we report the population coefficients \( \phi_r \) in (6.1) for all four models (EA\(_0\)(3), EA\(_1\)(3), EA\(_2\)(3), and EA\(_3\)(3)) in Figure 5.A, together with the historical “yield regression” coefficient \( \phi_{r,T} \). While the coefficients of EA\(_0\)(3) closely resemble the historical pattern, the coefficients of the other three models differ from the historical pattern dramatically! In particular, the coefficients of EA\(_2\)(3) and EA\(_3\)(3), the two models with relatively flexible time-varying conditional volatility specifications, are close to one with an upward sloping pattern, in sharp contrast to the downward sloping pattern of the historical coefficients. This result confirms the findings of Dai and Singleton (2002) for U.S. Treasury yields that only EA\(_0\)(3) can match the time variation in bond risk premium.

In Figure 5.B, we compare the coefficients of the “risk-adjusted yield regression” \( \phi^R_r \) in (6.2) for all four models with the horizontal line at unity. Again the coefficients of EA\(_2\)(3) and EA\(_3\)(3) fail completely to match the unity line. In contrast, the coefficients \( \phi^R_r \) for EA\(_0\)(3) and EA\(_1\)(3) lie close to unity, with the former performing better.

In summary, only EA\(_0\)(3) is able to capture time variations in the risk premiums for the Libor swap rates. Even for EA\(_0\)(3), Figure 5.B reveals a gap between \( \phi^R_r \) and unity for maturities under two years.

Time-Varying Conditional Volatility. Table 9 reports the parameter estimates of the GARCH (1,1) model for both the historical and simulated five-year yields from the four groups of ATSMs.\(^{18}\) As expected, EA\(_0\)(3), which assumes constant volatility for the state variables, exhibits little volatility

\(^{18}\)Though we focus on five-year maturity, we obtain similar results at all other maturities.
persistence. \( EA_1(3) \), while slightly underestimates the degree of volatility persistence, performs much better than \( EA_0(3) \) due to the one volatility factor. The implied GARCH (1,1) estimates for \( EA_2(3) \) and \( EA_3(3) \) match the empirical counterparts closely.

Figure 6 plots the sample variances computed using historical and simulated time series of Libor swap rates from the “essentially” ATSMs evaluated at their estimated parameter values. Except for \( EA_3(3) \), all other models exhibit a volatility hump. The humps of both \( EA_0(3) \) and \( EA_1(3) \) coincide with the historical pattern, while that of \( EA_2(3) \) occurs at around the two-year maturity.\(^{19}\) The implied variances of \( EA_0(3) \), however, are much smaller than the sample variances of the actual data. It seems \( EA_1(3) \) has the best overall performance of matching the humped volatility term structure and the high degree of volatility persistence of swap yields.

Overall, we see that the “essentially” ATSMs have difficulties in simultaneously capturing the time-varying risk premium and conditional volatility of Libor swap rates. These results reveal limitations in ATSMs, which can be improved by AJTSMs.

### 6.2.2 “Extended” ATSMs

**Time-Varying Risk Premium.** Based on the parameter estimates in Table 5 for three-factor “extended” ATSMs, we report the population coefficients \( \phi_r \) in (6.1) for all three models (\( EXA_1(3) \), \( EXA_2(3) \), and \( EXA_3(3) \)) in Figure 7.A, together with the historical “yield regression” coefficient \( \phi_{rT} \).\(^{20}\) Compared to the corresponding “essentially” ATSMs, the coefficients of the “extended” ATSMs are much closer to the historical pattern. In particular, generalizing the market price of diffusion risk from the “essentially” to “extended” affine specification improves the performance of the models with two and three volatility factors significantly. However, they still fail to completely capture the expectation puzzles, a result that is further corroborated by the “risk-adjusted yield regression” coefficients \( \phi^R_r \) in (6.2) reported in Figure 7.B.

**Time-Varying Conditional Volatility.** From Table 9, we observe that all three “extended” ATSMs

\(^{19}\)Piazzesi (2005), modeling deterministic jumps by linking jump intensities directly to the meeting calendar of the Federal Open Market Committee, provides a structural (monetary) interpretation of the volatility hump.

\(^{20}\)Since \( EXA_0(3) \) reduces to \( EA_0(3) \), we only focus on the other three “extended” ATSMs.
capture the degree of volatility persistence well, though EXA$_1$(3) and EXA$_2$(3) slightly underestimate the persistence. Figure 8 plots the sample variances computed using historical and simulated time series of Libor swap rates from the “extended” ATSMs evaluated at their estimated parameter values. Except for EXA$_3$(3), all other models exhibit a volatility hump. While the hump of EXA$_1$(3) coincides with the historical pattern, that of EXA$_2$(3) occurs at around the two-year maturity.

Overall, generalizing the “essentially” to “extended” affine specification for the market price of diffusion risk does improve the performance of ATSMs in capturing time-varying risk premiums in bond yields. Nevertheless, “extended” ATSMs are not sufficient to solve the “expectation puzzles” though they capture the conditional volatility of yields well.

6.3 Performance of Three-Factor AJTSMs

In this section we study the performance of our AJTSMs. Comparison between “extended” ATSMs and AJTSMs will illustrate the contributions of jump risk premium in capturing the “expectation puzzle” and time-varying volatility.

**Time-Varying Risk Premium.** Based on the parameter estimates in Table 6 for three-factor “essentially” AJTSMs, we report the population coefficients $\phi_r$ in (6.1) for all three models (EAJ$_1$(3), EAJ$_2$(3), and EAJ$_3$(3), in Figure 9.A, together with the historical “yield regression” coefficient $\phi_{rT}$. In dramatic contrast to the results of ATSMs as in Figures 5.A and 7.A, we find that all three AJTSMs can match the data well, with the following relative ranking of performances: EAJ$_0$(3)$>$ EAJ$_1$(3)$>$ EAJ$_2$(3)$>$ EAJ$_3$(3).

In Figure 9.B, we report the coefficients of the “risk-adjusted yield regression” $\phi^R_r$ in (6.2) for all three AJTSMs. The coefficients $\phi^R_r$ for both EAJ$_1$(3) and EAJ$_2$(3) are close to unity, with the same relative ranking of performance as that for the “yield regression.” This again differs from the results for “essentially” ATSMs in Figure 5.B, where only EAJ$_0$(3) can match the unity line reasonably well. Moreover, Figure 9.B also shows a gap between $\phi^R_r$ and unity for maturities under two years. For example, the risk-adjusted coefficient $\phi^R_r$ is only around 0.90 when $\tau = 9$ months for EAJ$_1$(3).

---

$21$ The REAJ$_0$(3) model coincides with the EAJ$_0$(3) model and we report its empirical performance in the next section for unrestricted AJTSMs.
Time-Varying Conditional Volatility. Table 8 reports the parameter estimates of the GARCH (1,1) model for both the historical and simulated five-year yields from the AJTSMs. Except EAJ\(_0\)(3), all the other three AJTSMs, EAJ\(_1\)(3), EAJ\(_2\)(3), and EAJ\(_3\)(3), match the volatility persistence in the data very well.

Figure 10 plots the sample variances computed using historical and simulated time series of Libor swap rates from the restricted AJTSMs evaluated at their estimated parameter values. Except for EAJ\(_0\)(3) and EAJ\(_3\)(3), the other two models (EAJ\(_1\)(3) and EAJ\(_2\)(3)) all exhibit a hump at around one-year maturity. Comparing Figure 10 with Figure 8, we see that all four AJTSMs have higher simulated variances than the corresponding “extended” ATSMs. The volatilities of EAJ\(_1\)(3) and EAJ\(_2\)(3) are pretty close to the historical level. Whereas EAJ\(_1\)(3) performs better at short maturities, EAJ\(_2\)(3) performs better at long maturities.

In summary, the empirical results in this section show that the “essentially” AJTSMs capture the time-varying risk premium and conditional volatility of Libor swap rates much better than both the “essentially” and “extended” ATSMs. In fact, both EAJ\(_1\)(3) and EAJ\(_2\)(3) can simultaneously match time variations in the risk premium and conditional volatility rather well. Jump risk premium breaks the tension in simultaneously capturing the time-varying risk premium and conditional volatility in ATSMs and contributes to the improved performance of AJTSMs.

7 Conclusion

The ATSMs have been extensively studied in the finance literature since the pioneering works of Duffie and Kan (1996) and Dai and Singleton (2000). Numerous extensions have been introduced to the basic model to better capture the data. Yet, ATSMs still face two empirical challenges. First, they ignore well-documented jumps in interest rates, as the state variables follow affine diffusions. Second, they fail to simultaneously capture violations of the “expectation hypothesis” and time variations in conditional volatilities of bond yields. In this paper, we show that by incorporating jumps into ATSMs, we can simultaneously resolve the two challenges.

Our empirical results show that jumps are essential for modeling term structure dynamics, as
AJTSMs significantly outperform ATSMs in capturing the conditional moments, especially the skewness and kurtosis, of bond yields at both short and long maturities. Moreover, we show that jump risk premiums lead to flexible time-varying market prices of risks without restricting time variations in conditional volatility. Consequently, two sub-classes of three-factor AJTSMs simultaneously capture violations of the “expectation hypothesis” and time variations of the conditional volatility of Libor swap rates.
References


Joslin, S., Le, A., Singleton, K., 2011. The conditional distribution of bond yields implied by
gaussian macro-finance term structure models. working paper.
# Table 1: Summary Statistics

This table reports the summary statistics of daily Libor Swap rates with 3-month, 6-month, 9-month, 2-year, 3-year, 4-year, 5-year, 7-year, and 10-year maturities of from August 13, 1990, to December 31, 2008, with a total of 4757 observations.

**Panel A: Summary statistics of the levels of Libor Swap yields.**

<table>
<thead>
<tr>
<th>Maturity</th>
<th>3-month</th>
<th>6-month</th>
<th>9-month</th>
<th>2-year</th>
<th>3-year</th>
<th>4-year</th>
<th>5-year</th>
<th>7-year</th>
<th>10-year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (%)</td>
<td>4.4139</td>
<td>4.5059</td>
<td>4.5954</td>
<td>5.0224</td>
<td>5.3047</td>
<td>5.5210</td>
<td>5.6928</td>
<td>5.9369</td>
<td>6.1693</td>
</tr>
<tr>
<td>Std. Dev. (%)</td>
<td>1.7518</td>
<td>1.7429</td>
<td>1.7398</td>
<td>1.6318</td>
<td>1.5364</td>
<td>1.4702</td>
<td>1.4250</td>
<td>1.3668</td>
<td>1.3114</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.3672</td>
<td>-0.3913</td>
<td>-0.3850</td>
<td>-0.2393</td>
<td>-0.1134</td>
<td>0.0039</td>
<td>0.1042</td>
<td>0.2634</td>
<td>0.3572</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>2.2109</td>
<td>2.2427</td>
<td>2.2672</td>
<td>2.4076</td>
<td>2.4713</td>
<td>2.4857</td>
<td>2.4730</td>
<td>2.4961</td>
<td>2.4745</td>
</tr>
<tr>
<td>First-order Partial Autocorrelation</td>
<td>0.9989</td>
<td>0.9989</td>
<td>0.9988</td>
<td>0.9984</td>
<td>0.9982</td>
<td>0.9980</td>
<td>0.9979</td>
<td>0.9977</td>
<td>0.9976</td>
</tr>
</tbody>
</table>

**Panel B: Summary statistics of the changes of Libor Swap yields.**

<table>
<thead>
<tr>
<th>Maturity</th>
<th>3-month</th>
<th>6-month</th>
<th>9-month</th>
<th>2-year</th>
<th>3-year</th>
<th>4-year</th>
<th>5-year</th>
<th>7-year</th>
<th>10-year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (%)</td>
<td>-0.0014</td>
<td>-0.0013</td>
<td>-0.0013</td>
<td>-0.0014</td>
<td>-0.0014</td>
<td>-0.0014</td>
<td>-0.0014</td>
<td>-0.0014</td>
<td>-0.0014</td>
</tr>
<tr>
<td>Std. Dev. (%)</td>
<td>0.0276</td>
<td>0.0318</td>
<td>0.0340</td>
<td>0.0170</td>
<td>0.0159</td>
<td>0.0146</td>
<td>0.0139</td>
<td>0.0124</td>
<td>0.0112</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.2676</td>
<td>-0.4306</td>
<td>-0.1718</td>
<td>0.0895</td>
<td>0.1373</td>
<td>0.1814</td>
<td>0.1859</td>
<td>0.1602</td>
<td>0.1063</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>163.3652</td>
<td>247.1552</td>
<td>237.7388</td>
<td>6.2827</td>
<td>6.5372</td>
<td>5.7005</td>
<td>5.6979</td>
<td>5.6864</td>
<td>5.7938</td>
</tr>
<tr>
<td>First-order Partial Autocorrelation</td>
<td>0.0031</td>
<td>-0.1049</td>
<td>-0.1134</td>
<td>0.0424</td>
<td>0.0197</td>
<td>0.0250</td>
<td>0.0199</td>
<td>0.0215</td>
<td>0.0289</td>
</tr>
</tbody>
</table>

**Panel C. Principal component analysis of the levels and changes of Libor Swap yields.** The entries represent the percentages of the variations of the levels and changes of Libor Swap yields explained by each of their first six principal components.

<table>
<thead>
<tr>
<th>Principal Component</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td>92.50%</td>
<td>7.11%</td>
<td>0.34%</td>
<td>0.03%</td>
<td>0.01%</td>
<td>0.01%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Change</td>
<td>73.28%</td>
<td>20.99%</td>
<td>2.61%</td>
<td>1.44%</td>
<td>0.62%</td>
<td>0.38%</td>
<td>0.30%</td>
<td>0.20%</td>
<td>0.18%</td>
</tr>
</tbody>
</table>
Table 2: Yield Regression Using Libor Swap Rates.

This table reports the results of “yield regression” in (4.3) using daily Libor Swap rates with maturities as indicated in the table from August 13, 1990 to December 31, 2008. The 3-month Libor rate is used as the spot rate $r_t$. In the “s.e.” row are the Newey-West standard errors of $\phi_{rT}$.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>6-month $\phi_{rT}$</th>
<th>6-month s.e.</th>
<th>9-month $\phi_{rT}$</th>
<th>9-month s.e.</th>
<th>2-year $\phi_{rT}$</th>
<th>2-year s.e.</th>
<th>3-year $\phi_{rT}$</th>
<th>3-year s.e.</th>
<th>4-year $\phi_{rT}$</th>
<th>4-year s.e.</th>
<th>5-year $\phi_{rT}$</th>
<th>5-year s.e.</th>
<th>7-year $\phi_{rT}$</th>
<th>7-year s.e.</th>
<th>10-year $\phi_{rT}$</th>
<th>10-year s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.0095</td>
<td>0.0027</td>
<td>0.0088</td>
<td>0.0026</td>
<td>0.0025</td>
<td>0.0031</td>
<td>-0.0002</td>
<td>-0.0025</td>
<td>-0.0015</td>
<td>-0.0008</td>
<td>-0.0025</td>
<td>-0.0040</td>
<td>-0.0025</td>
<td>-0.0051</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: GARCH (1,1) Parameters for the Libor Swap Yields.

This table reports the maximum likelihood estimates of a GARCH(1,1) model: $\sigma_t^2 = \sigma^2 + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2$, where $\varepsilon_t$ is the innovation from the AR(1) representation of the Libor Swap yields with maturities as indicated in the table from August 13, 1990, to December 31, 2008. Standard errors are given in the “SE” columns.

<table>
<thead>
<tr>
<th>GARCH(1,1)</th>
<th>$\sigma$</th>
<th>$\alpha$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maturity</td>
<td>Est. SE</td>
<td>Est. SE</td>
<td>Est. SE</td>
</tr>
<tr>
<td>3-month</td>
<td>0.0001</td>
<td>3.8.10^{-5}</td>
<td>0.2486</td>
</tr>
<tr>
<td>6-month</td>
<td>0.0002</td>
<td>3.9.10^{-5}</td>
<td>0.2112</td>
</tr>
<tr>
<td>9-month</td>
<td>0.0002</td>
<td>3.4.10^{-5}</td>
<td>0.1989</td>
</tr>
<tr>
<td>2-year</td>
<td>0.0003</td>
<td>6.5.10^{-5}</td>
<td>0.1723</td>
</tr>
<tr>
<td>3-year</td>
<td>0.0001</td>
<td>1.2.10^{-5}</td>
<td>0.1706</td>
</tr>
<tr>
<td>4-year</td>
<td>0.0002</td>
<td>1.9.10^{-5}</td>
<td>0.1350</td>
</tr>
<tr>
<td>5-year</td>
<td>0.0001</td>
<td>2.6.10^{-5}</td>
<td>0.1679</td>
</tr>
<tr>
<td>7-year</td>
<td>0.0005</td>
<td>5.1.10^{-5}</td>
<td>0.1464</td>
</tr>
<tr>
<td>10-year</td>
<td>0.0002</td>
<td>4.7.10^{-5}</td>
<td>0.1564</td>
</tr>
</tbody>
</table>
Table 4 Parameter Estimates for Three-Factor “Essentially” ATSMs

This table reports parameter estimates for three-factor “essentially” ATSMs using Libor Swap rates with 6-month, 2-year, 10-year, 3-year, 5-year, and 7-year maturities and sampled daily from August 13, 1990, to December 31, 2008. The rates with the first three maturities are assumed to be observed without error and the remaining three observed with error. Estimation follows the infinitesimal operator-based procedure described in Section 3. Specifications of models, $EA_0(3)$, $EA_1(3)$, $EA_2(3)$, and $EA_3(3)$, are presented in Section 2.3. Reported in the parentheses of the “SE” column are the standard errors of the coefficients estimates. The blanks in the “SE” columns refer to those parameters pre-specified by the model structures.

<table>
<thead>
<tr>
<th></th>
<th>$EA_0(3)$</th>
<th>$EA_1(3)$</th>
<th>$EA_2(3)$</th>
<th>$EA_3(3)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\kappa_{11}$</td>
<td>0.0420 (0.0281)</td>
<td>0.0967 (0.0052)</td>
<td>0.1154 (0.1332)</td>
<td>0.3166 (0.0667)</td>
</tr>
<tr>
<td>$\kappa_{12}$</td>
<td>0.0</td>
<td>-</td>
<td>1.0458 (0.0007)</td>
<td>0.0023 (0.3009)</td>
</tr>
<tr>
<td>$\kappa_{13}$</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>$\kappa_{21}$</td>
<td>-0.1513 (0.2444)</td>
<td>-0.0684 (0.0152)</td>
<td>-0.1150 (0.0726)</td>
<td>0.0440 (0.0213)</td>
</tr>
<tr>
<td>$\kappa_{22}$</td>
<td>0.0371 (0.0026)</td>
<td>0.1273 (0.0904)</td>
<td>0.7451 (0.1111)</td>
<td>0.0775 (0.0074)</td>
</tr>
<tr>
<td>$\kappa_{23}$</td>
<td>0.0</td>
<td>-</td>
<td>-3.1449 (1.4334)</td>
<td>0.0702 (0.0667)</td>
</tr>
<tr>
<td>$\kappa_{31}$</td>
<td>3.3572 (2.4403)</td>
<td>0.0128 (0.0206)</td>
<td>0.1485 (0.2234)</td>
<td>-0.0463 (0.0061)</td>
</tr>
<tr>
<td>$\kappa_{32}$</td>
<td>-0.0193 (0.0804)</td>
<td>0.3532 (0.0925)</td>
<td>0.0933 (0.0458)</td>
<td>0.0208 (0.0092)</td>
</tr>
<tr>
<td>$\kappa_{33}$</td>
<td>0.0415 (0.0140)</td>
<td>0.1014 (0.1307)</td>
<td>0.1652 (0.0149)</td>
<td>0.2771 (0.1709)</td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>0.0</td>
<td>-</td>
<td>0.1067 (0.0885)</td>
<td>0.0955 (0.0210)</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>$\theta_3$</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>1.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>1.0</td>
<td>-</td>
<td>1.0</td>
<td>-</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>1.0</td>
<td>-</td>
<td>1.0</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_{11}$</td>
<td>0.0</td>
<td>-</td>
<td>1.0</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_{12}$</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_{13}$</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_{21}$</td>
<td>0.0</td>
<td>-</td>
<td>0.1635 (0.0291)</td>
<td>0.0</td>
</tr>
<tr>
<td>$\beta_{22}$</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_{23}$</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_{31}$</td>
<td>0.0</td>
<td>-</td>
<td>0.2991 (0.2507)</td>
<td>0.1212 (0.1404)</td>
</tr>
<tr>
<td>$\beta_{32}$</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_{33}$</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
</tbody>
</table>

36
<table>
<thead>
<tr>
<th></th>
<th>EA0(3)</th>
<th></th>
<th>EA1(3)</th>
<th></th>
<th>EA2(3)</th>
<th></th>
<th>EA3(3)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est.</td>
<td>SE</td>
<td>Est.</td>
<td>SE</td>
<td>Est.</td>
<td>SE</td>
<td>Est.</td>
<td>SE</td>
</tr>
<tr>
<td>(\eta_{11})</td>
<td>-0.3008</td>
<td>(0.0042)</td>
<td>-0.1544</td>
<td>(0.2033)</td>
<td>-0.2247</td>
<td>(0.2481)</td>
<td>-0.0190</td>
<td>(0.0655)</td>
</tr>
<tr>
<td>(\eta_{12})</td>
<td>-0.0213</td>
<td>(0.0114)</td>
<td>-0.2007</td>
<td>(0.2011)</td>
<td>-0.0913</td>
<td>(0.0449)</td>
<td>-0.0418</td>
<td>(0.0248)</td>
</tr>
<tr>
<td>(\eta_{13})</td>
<td>-0.2822</td>
<td>(0.0211)</td>
<td>-0.0688</td>
<td>(0.0443)</td>
<td>-0.0701</td>
<td>(0.0545)</td>
<td>-0.1406</td>
<td>(0.0221)</td>
</tr>
<tr>
<td>(\delta_0)</td>
<td>-0.0226</td>
<td>(0.0040)</td>
<td>0.2410</td>
<td>(0.0748)</td>
<td>-0.0416</td>
<td>(0.0903)</td>
<td>0.6004</td>
<td>(0.0811)</td>
</tr>
<tr>
<td>(\delta_{11})</td>
<td>0.0690</td>
<td>(0.0433)</td>
<td>0.1152</td>
<td>(0.1067)</td>
<td>0.0707</td>
<td>(0.0242)</td>
<td>0.7107</td>
<td>(2.1633)</td>
</tr>
<tr>
<td>(\delta_{12})</td>
<td>0.0172</td>
<td>(0.0089)</td>
<td>0.0934</td>
<td>(0.0805)</td>
<td>0.8805</td>
<td>(0.1924)</td>
<td>0.2022</td>
<td>(0.2139)</td>
</tr>
<tr>
<td>(\delta_{13})</td>
<td>0.1006</td>
<td>(0.0088)</td>
<td>0.3444</td>
<td>(0.0455)</td>
<td>-0.1430</td>
<td>(0.0782)</td>
<td>0.0467</td>
<td>(0.0141)</td>
</tr>
<tr>
<td>(\eta_{2,11})</td>
<td>-0.0122</td>
<td>(0.0227)</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>(\eta_{2,12})</td>
<td>-0.1709</td>
<td>(0.1066)</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>(\eta_{2,13})</td>
<td>-0.5282</td>
<td>(0.1544)</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>(\eta_{2,21})</td>
<td>-0.4111</td>
<td>(0.2554)</td>
<td>-0.0157</td>
<td>(0.0266)</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>(\eta_{2,22})</td>
<td>-0.2333</td>
<td>(0.2607)</td>
<td>-0.0333</td>
<td>(0.0182)</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>(\eta_{2,23})</td>
<td>-0.2900</td>
<td>(0.0177)</td>
<td>-0.0910</td>
<td>(0.0187)</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>(\eta_{2,31})</td>
<td>-0.3171</td>
<td>(0.2403)</td>
<td>-0.0229</td>
<td>(0.0316)</td>
<td>-0.0114</td>
<td>(0.2548)</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>(\eta_{2,32})</td>
<td>-0.1155</td>
<td>(0.0778)</td>
<td>-0.0142</td>
<td>(0.0040)</td>
<td>-0.0502</td>
<td>(0.0133)</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>(\eta_{2,33})</td>
<td>-0.1582</td>
<td>(0.1267)</td>
<td>-0.0143</td>
<td>(0.0122)</td>
<td>-0.0704</td>
<td>(0.0166)</td>
<td>0.0</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 5 Parameter Estimates for Three-Factor “Extended” ATSMs

This table reports parameter estimates for three-factor “extended” ATSMs using Libor Swap rates with 6-month, 2-year, 10-year, 3-year, 5-year, and 7-year maturities and sampled daily from August 13, 1990, to December 31, 2008. The rates with the first three maturities are assumed to be observed without error and the remaining three observed with error. Estimation follows the infinitesimal operator-based procedure described in Section 3. Specifications of models, EXA₁(3), EXA₂(3), and EXA₃(3), are presented in Section 2.3. Reported in the parentheses of the “SE” column are the standard errors of the coefficients estimates. The blanks in the “SE” columns refer to those parameters pre-specified by the model structures.

<table>
<thead>
<tr>
<th></th>
<th>EXA₁(3)</th>
<th></th>
<th>EXA₂(3)</th>
<th></th>
<th>EXA₃(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est.</td>
<td>SE</td>
<td>Est.</td>
<td>SE</td>
<td>Est.</td>
</tr>
<tr>
<td>κ₁₁</td>
<td>0.0952</td>
<td>(0.0052)</td>
<td>0.1108</td>
<td>(0.1066)</td>
<td>0.3104</td>
</tr>
<tr>
<td>κ₁₂</td>
<td>0.0</td>
<td>-</td>
<td>0.9225</td>
<td>(0.3024)</td>
<td>0.0017</td>
</tr>
<tr>
<td>κ₁₃</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>-0.0208</td>
</tr>
<tr>
<td>κ₂₁</td>
<td>-0.0684</td>
<td>(0.0152)</td>
<td>-0.1154</td>
<td>(0.0730)</td>
<td>0.0438</td>
</tr>
<tr>
<td>κ₂₂</td>
<td>0.1273</td>
<td>(0.0904)</td>
<td>0.7451</td>
<td>(0.1111)</td>
<td>0.0722</td>
</tr>
<tr>
<td>κ₂₃</td>
<td>-3.1449</td>
<td>(1.4334)</td>
<td>0.0</td>
<td>-</td>
<td>0.0701</td>
</tr>
<tr>
<td>κ₃₁</td>
<td>0.0128</td>
<td>(0.0206)</td>
<td>0.1492</td>
<td>(0.2242)</td>
<td>-0.0458</td>
</tr>
<tr>
<td>κ₃₂</td>
<td>0.3532</td>
<td>(0.0925)</td>
<td>0.0909</td>
<td>(0.0447)</td>
<td>0.0211</td>
</tr>
<tr>
<td>κ₃₃</td>
<td>0.1014</td>
<td>(0.1307)</td>
<td>0.1651</td>
<td>(0.0149)</td>
<td>0.2553</td>
</tr>
<tr>
<td>θ₁</td>
<td>0.1023</td>
<td>(0.0836)</td>
<td>0.0948</td>
<td>(0.0223)</td>
<td>0.0844</td>
</tr>
<tr>
<td>θ₂</td>
<td>0.0</td>
<td>-</td>
<td>0.2210</td>
<td>(0.2706)</td>
<td>0.0569</td>
</tr>
<tr>
<td>θ₃</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0546</td>
</tr>
<tr>
<td>α₁</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
</tr>
<tr>
<td>α₂</td>
<td>1.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
</tr>
<tr>
<td>α₃</td>
<td>1.0</td>
<td>-</td>
<td>1.0</td>
<td>-</td>
<td>0.0</td>
</tr>
<tr>
<td>β₁₁</td>
<td>1.0</td>
<td>-</td>
<td>1.0</td>
<td>-</td>
<td>1.0</td>
</tr>
<tr>
<td>β₁₂</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
</tr>
<tr>
<td>β₁₃</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
</tr>
<tr>
<td>β₂₁</td>
<td>0.1635</td>
<td>(0.0291)</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
</tr>
<tr>
<td>β₂₂</td>
<td>0.0</td>
<td>-</td>
<td>1.0</td>
<td>-</td>
<td>1.0</td>
</tr>
<tr>
<td>β₂₃</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
</tr>
<tr>
<td>β₃₁</td>
<td>0.2991</td>
<td>(0.2507)</td>
<td>0.1213</td>
<td>(0.1404)</td>
<td>0.0</td>
</tr>
<tr>
<td>β₃₂</td>
<td>0.0</td>
<td>-</td>
<td>0.0353</td>
<td>(0.0371)</td>
<td>0.0</td>
</tr>
<tr>
<td>β₃₃</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>1.0</td>
</tr>
</tbody>
</table>
Table 5. Continued.

<table>
<thead>
<tr>
<th></th>
<th>EXA₁(3)</th>
<th></th>
<th>EXA₂(3)</th>
<th></th>
<th>EXA₃(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est.</td>
<td>SE</td>
<td>Est.</td>
<td>SE</td>
<td>Est.</td>
</tr>
<tr>
<td>( \eta_{11} )</td>
<td>-0.1544</td>
<td>(0.2033)</td>
<td>-0.2250</td>
<td>(0.2493)</td>
<td>-0.0214</td>
</tr>
<tr>
<td>( \eta_{12} )</td>
<td>-0.2007</td>
<td>(0.2011)</td>
<td>-0.0906</td>
<td>(0.0451)</td>
<td>-0.0420</td>
</tr>
<tr>
<td>( \eta_{13} )</td>
<td>-0.0688</td>
<td>(0.0443)</td>
<td>-0.0792</td>
<td>(0.0550)</td>
<td>-0.1414</td>
</tr>
<tr>
<td>( \gamma_{10} )</td>
<td>1.6604</td>
<td>(0.9117)</td>
<td>-0.3111</td>
<td>(0.2499)</td>
<td>-0.1103</td>
</tr>
<tr>
<td>( \gamma_{12} )</td>
<td>0.0</td>
<td>-</td>
<td>-0.0967</td>
<td>(0.0772)</td>
<td>0.0001</td>
</tr>
<tr>
<td>( \gamma_{13} )</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>-0.0667</td>
</tr>
<tr>
<td>( \gamma_{20} )</td>
<td>0.0</td>
<td>-</td>
<td>-1.0020</td>
<td>(0.8843)</td>
<td>3.8004</td>
</tr>
<tr>
<td>( \gamma_{21} )</td>
<td>0.0</td>
<td>-</td>
<td>-0.0321</td>
<td>(0.0066)</td>
<td>-6.3012</td>
</tr>
<tr>
<td>( \gamma_{23} )</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>-7.1121</td>
</tr>
<tr>
<td>( \gamma_{30} )</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>-5.008</td>
</tr>
<tr>
<td>( \gamma_{31} )</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>-0.0363</td>
</tr>
<tr>
<td>( \gamma_{32} )</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0006</td>
</tr>
<tr>
<td>( \delta_0 )</td>
<td>0.2410</td>
<td>(0.0748)</td>
<td>-0.0416</td>
<td>(0.0903)</td>
<td>0.5772</td>
</tr>
<tr>
<td>( \delta_{11} )</td>
<td>0.1152</td>
<td>(0.1067)</td>
<td>0.0707</td>
<td>(0.0242)</td>
<td>0.7881</td>
</tr>
<tr>
<td>( \delta_{12} )</td>
<td>0.0934</td>
<td>(0.0805)</td>
<td>0.8006</td>
<td>(0.3665)</td>
<td>0.1443</td>
</tr>
<tr>
<td>( \delta_{13} )</td>
<td>0.3444</td>
<td>(0.0455)</td>
<td>-0.1433</td>
<td>(0.0771)</td>
<td>0.0484</td>
</tr>
<tr>
<td>( \eta_{2,11} )</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
</tr>
<tr>
<td>( \eta_{2,12} )</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
</tr>
<tr>
<td>( \eta_{2,13} )</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
</tr>
<tr>
<td>( \eta_{2,21} )</td>
<td>-0.0157</td>
<td>(0.0266)</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
</tr>
<tr>
<td>( \eta_{2,22} )</td>
<td>-0.0333</td>
<td>(0.0182)</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
</tr>
<tr>
<td>( \eta_{2,23} )</td>
<td>-0.0910</td>
<td>(0.0187)</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
</tr>
<tr>
<td>( \eta_{2,31} )</td>
<td>-0.0229</td>
<td>(0.0316)</td>
<td>-0.0115</td>
<td>(0.2548)</td>
<td>0.0</td>
</tr>
<tr>
<td>( \eta_{2,32} )</td>
<td>-0.0142</td>
<td>(0.0040)</td>
<td>-0.0476</td>
<td>(0.0152)</td>
<td>0.0</td>
</tr>
<tr>
<td>( \eta_{2,33} )</td>
<td>-0.0143</td>
<td>(0.0122)</td>
<td>-0.0707</td>
<td>(0.0166)</td>
<td>0.0</td>
</tr>
</tbody>
</table>
Table 6: Parameter Estimates for Three-Factor “Essentially” AJTSMs

This table reports parameter estimates for three-factor restricted essentially AJSTMs using Libor Swap rates with 6-month, 2-year, 10-year, 3-year, 5-year, and 7-year maturities and sampled daily from August 13, 1990, to December 31, 2008. The rates with the first three maturities are assumed to be observed without error and the remaining three observed with error. Estimation follows the infinitesimal operator-based procedure described in Section 3. The specifications of the models, EAJ0(3), EAJ1(3), EAJ2(3), and EAJ3(3), are presented in Section 2.3. Reported in the parentheses of the “SE” column are the standard errors of the coefficients estimates. The blanks in the “SE” columns refer to those parameters pre-specified by the model structures.

<table>
<thead>
<tr>
<th></th>
<th>EAJ0(3)</th>
<th></th>
<th>EAJ1(3)</th>
<th></th>
<th>EAJ2(3)</th>
<th></th>
<th>EAJ3(3)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est.</td>
<td>SE</td>
<td>Est.</td>
<td>SE</td>
<td>Est.</td>
<td>SE</td>
<td>Est.</td>
<td>SE</td>
</tr>
<tr>
<td>$\kappa_{11}$</td>
<td>0.3122 (0.2063)</td>
<td>0.0127 (0.0026)</td>
<td>0.1820 (0.0966)</td>
<td>0.2725 (0.1233)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\kappa_{12}$</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>-0.1345 (0.0072)</td>
<td>-0.0051 (0.0014)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\kappa_{13}$</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>-0.0012 (0.0005)</td>
<td></td>
</tr>
<tr>
<td>$\kappa_{21}$</td>
<td>0.0908 (0.0321)</td>
<td>0.2152 (0.1375)</td>
<td>-0.0823 (0.0184)</td>
<td>0.0057 (0.0146)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\kappa_{22}$</td>
<td>0.1615 (0.0212)</td>
<td>1.2233 (0.0716)</td>
<td>0.4417 (0.0762)</td>
<td>0.0874 (0.0081)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\kappa_{23}$</td>
<td>0.0</td>
<td>-3.9424 (0.7009)</td>
<td>0.0</td>
<td>-</td>
<td>0.0001 (0.0013)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\kappa_{31}$</td>
<td>0.4276 (0.1119)</td>
<td>-0.2006 (0.0851)</td>
<td>0.1113 (0.2553)</td>
<td>-0.0007 (0.0004)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\kappa_{32}$</td>
<td>0.3007 (0.0422)</td>
<td>1.8667 (0.5094)</td>
<td>0.2004 (0.8771)</td>
<td>-0.0314 (0.0141)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\kappa_{33}$</td>
<td>0.0933 (0.0374)</td>
<td>0.0722 (0.0248)</td>
<td>0.2409 (0.0882)</td>
<td>0.7772 (0.4019)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>0.0</td>
<td>-</td>
<td>0.1503 (0.0747)</td>
<td>0.0915 (0.0374)</td>
<td>0.0059 (0.0033)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.1366 (0.0677)</td>
<td>0.5115 (0.2283)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta_3$</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0048 (0.0027)</td>
<td></td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>1.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>1.0</td>
<td>-</td>
<td>1.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>1.0</td>
<td>-</td>
<td>1.0</td>
<td>-</td>
<td>1.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_11$</td>
<td>0.0</td>
<td>-</td>
<td>1.0</td>
<td>-</td>
<td>1.0</td>
<td>-</td>
<td>1.0</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_12$</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_13$</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_{21}$</td>
<td>0.0</td>
<td>-</td>
<td>0.0507 (0.0030)</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>$\beta_{22}$</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>1.0</td>
<td>-</td>
<td>1.0</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_{23}$</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_{31}$</td>
<td>0.0</td>
<td>-</td>
<td>0.3009 (0.0118)</td>
<td>0.1224 (0.5541)</td>
<td>0.0</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{32}$</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0240 (0.0069)</td>
<td>0.0</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>$\beta_{33}$</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>1.0</td>
<td>-</td>
<td>1.0</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>EAJ$_0(3)$</td>
<td></td>
<td>EAJ$_1(3)$</td>
<td></td>
<td>EAJ$_2(3)$</td>
<td></td>
<td>EAJ$_3(3)$</td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>-----------</td>
<td>--------------------------</td>
<td>-----------</td>
<td>--------------------------</td>
<td>-----------</td>
<td>--------------------------</td>
<td>-----------</td>
<td>--------------------------</td>
</tr>
<tr>
<td></td>
<td>Est.</td>
<td>SE</td>
<td>Est.</td>
<td>SE</td>
<td>Est.</td>
<td>SE</td>
<td>Est.</td>
<td>SE</td>
</tr>
<tr>
<td>$\lambda_0/\lambda_0^Q$</td>
<td>0.2205/0.3912</td>
<td>(0.0814)/(0.3449)</td>
<td>0.1784/0.3326</td>
<td>(0.0350)/(0.1233)</td>
<td>0.1899/0.3004</td>
<td>(0.1224)/(0.0975)</td>
<td>0.0078/0.0136</td>
<td>(0.0013)/(0.0044)</td>
</tr>
<tr>
<td>$\mu_{10}/\mu_{10}^Q$</td>
<td>-0.0021/-0.0035</td>
<td>(0.0022)/(0.0014)</td>
<td>0.0087/0.0042</td>
<td>(0.0031)/(0.0020)</td>
<td>0.0053/0.0036</td>
<td>(0.0014)/(0.0018)</td>
<td>0.0033/0.0020</td>
<td>(0.0009)/(0.0014)</td>
</tr>
<tr>
<td>$\mu_{11}/\mu_{11}^Q$</td>
<td>-0.0007/-0.0009</td>
<td>(0.0006)/(0.0005)</td>
<td>0.0011/0.0003</td>
<td>(0.0006)/(0.0001)</td>
<td>0.0029/0.0022</td>
<td>(0.0015)/(0.0009)</td>
<td>0.0022/0.0015</td>
<td>(0.0004)/(0.0008)</td>
</tr>
<tr>
<td>$\mu_{20}/\mu_{20}^Q$</td>
<td>-0.0016/-0.0053</td>
<td>(0.0003)/(0.0027)</td>
<td>-0.0032/-0.0075</td>
<td>(0.0012)/(0.0035)</td>
<td>0.0042/0.0023</td>
<td>(0.0016)/(0.0011)</td>
<td>0.0043/0.0019</td>
<td>(0.0011)/(0.0007)</td>
</tr>
<tr>
<td>$\mu_{21}/\mu_{21}^Q$</td>
<td>0.0019/0.0024</td>
<td>(0.0006)/(0.0019)</td>
<td>0.0021/0.0014</td>
<td>(0.0009)/(0.0005)</td>
<td>0.0008/0.0006</td>
<td>(0.0005)/(0.0001)</td>
<td>0.0008/0.0005</td>
<td>(0.0003)/(0.0003)</td>
</tr>
<tr>
<td>$\mu_{30}/\mu_{30}^Q$</td>
<td>-0.0028/-0.0031</td>
<td>(0.0009)/(0.0018)</td>
<td>-0.0038/-0.0027</td>
<td>(0.0017)/(0.0011)</td>
<td>-0.0014/-0.0018</td>
<td>(0.0051)/(0.0022)</td>
<td>0.0021/0.0012</td>
<td>(0.0006)/(0.0007)</td>
</tr>
<tr>
<td>$\mu_{31}/\mu_{31}^Q$</td>
<td>0.0014/0.0063</td>
<td>(0.0002)/(0.0044)</td>
<td>0.0020/0.0016</td>
<td>(0.0008)/(0.0007)</td>
<td>0.0030/0.0014</td>
<td>(0.0007)/(0.0005)</td>
<td>0.0059/0.0026</td>
<td>(0.0025)/(0.0012)</td>
</tr>
<tr>
<td>$\sigma_1^2$</td>
<td>0.0408</td>
<td>(0.0233)</td>
<td>0.4462</td>
<td>(0.1396)</td>
<td>0.0223</td>
<td>(0.0217)</td>
<td>0.0260</td>
<td>(0.0154)</td>
</tr>
<tr>
<td>$\sigma_2^2$</td>
<td>0.0327</td>
<td>(0.0051)</td>
<td>0.1527</td>
<td>(0.1341)</td>
<td>0.0122</td>
<td>(0.0070)</td>
<td>0.0177</td>
<td>(0.0067)</td>
</tr>
<tr>
<td>$\sigma_3^2$</td>
<td>0.0719</td>
<td>(0.0246)</td>
<td>0.1173</td>
<td>(0.0335)</td>
<td>0.0195</td>
<td>(0.0064)</td>
<td>0.0226</td>
<td>(0.0072)</td>
</tr>
<tr>
<td>$\eta_{11}$</td>
<td>-0.2552</td>
<td>(0.0167)</td>
<td>-0.0721</td>
<td>(0.0148)</td>
<td>-0.0538</td>
<td>(0.0822)</td>
<td>-0.0825</td>
<td>(0.0778)</td>
</tr>
<tr>
<td>$\eta_{12}$</td>
<td>-0.1355</td>
<td>(0.0788)</td>
<td>-0.1052</td>
<td>(0.0309)</td>
<td>-0.0516</td>
<td>(0.0120)</td>
<td>-0.1114</td>
<td>(0.0283)</td>
</tr>
<tr>
<td>$\eta_{13}$</td>
<td>-0.3134</td>
<td>(0.0192)</td>
<td>-0.1644</td>
<td>(0.0352)</td>
<td>-0.0339</td>
<td>(0.0088)</td>
<td>-0.3034</td>
<td>(0.0826)</td>
</tr>
<tr>
<td>$\delta_0$</td>
<td>-0.0321</td>
<td>(0.0030)</td>
<td>0.0413</td>
<td>(0.0098)</td>
<td>0.0264</td>
<td>(0.0077)</td>
<td>0.0811</td>
<td>(0.0392)</td>
</tr>
<tr>
<td>$\delta_{11}$</td>
<td>0.0622</td>
<td>(0.0111)</td>
<td>0.0451</td>
<td>(0.0202)</td>
<td>0.0720</td>
<td>(0.0355)</td>
<td>0.0655</td>
<td>(0.0262)</td>
</tr>
<tr>
<td>$\delta_{12}$</td>
<td>0.1217</td>
<td>(0.0483)</td>
<td>0.0267</td>
<td>(0.0216)</td>
<td>-0.0208</td>
<td>(0.0091)</td>
<td>-0.0134</td>
<td>(0.0079)</td>
</tr>
<tr>
<td>$\delta_{13}$</td>
<td>0.0555</td>
<td>(0.0162)</td>
<td>-0.2485</td>
<td>(0.1116)</td>
<td>-0.0663</td>
<td>(0.0348)</td>
<td>-0.0388</td>
<td>(0.0084)</td>
</tr>
<tr>
<td>$\gamma_{2,11}$</td>
<td>-0.0112</td>
<td>(0.0213)</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>$\gamma_{2,12}$</td>
<td>-0.0200</td>
<td>(0.0048)</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>$\gamma_{2,13}$</td>
<td>-0.0354</td>
<td>(0.0024)</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>$\gamma_{2,21}$</td>
<td>-0.0408</td>
<td>(0.0327)</td>
<td>-0.3537</td>
<td>(0.8066)</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>$\gamma_{2,22}$</td>
<td>-0.0726</td>
<td>(0.2002)</td>
<td>-0.7225</td>
<td>(0.6392)</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>$\gamma_{2,23}$</td>
<td>-0.0616</td>
<td>(0.0135)</td>
<td>-0.1422</td>
<td>(0.0541)</td>
<td>0.0</td>
<td>-</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>$\gamma_{2,31}$</td>
<td>-0.0681</td>
<td>(0.0974)</td>
<td>-0.2886</td>
<td>(0.2033)</td>
<td>-0.0347</td>
<td>(0.2269)</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>$\gamma_{2,32}$</td>
<td>-0.0900</td>
<td>(0.0333)</td>
<td>-0.0444</td>
<td>(0.2528)</td>
<td>-0.0836</td>
<td>(0.1850)</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>$\gamma_{2,33}$</td>
<td>-0.0223</td>
<td>(0.0034)</td>
<td>-0.2177</td>
<td>(0.1694)</td>
<td>-0.0584</td>
<td>(0.0661)</td>
<td>0.0</td>
<td>-</td>
</tr>
</tbody>
</table>

41
Table 7: Model Diagnostics by Conditional Moments

This table reports model diagnostics of the EXA$_2(3)$ and EAJ$_2(3)$ models based on conditional moments of yields. The conditional mean, volatility (measured by standard deviation), skewness and kurtosis of both the data and models are reported for 3-month, 2-year, 5-year, and 10-year maturities. The conditioning variable is the slope of the term structure, defined as the difference between the 10-year and 3-month yields. We consider two states of the term structure, upward when the slope is positive and downward vice versa. To compute model-implied conditional moments, we simulate 1000 sample paths from the two models using estimated parameters in Tables 5 and 6, categorize the time series of yields according to the slope, compute the moments for each generated sample, and finally average them across the 1000 sample paths. The column "EAJ$_2(3)$ vs EXA$_2(3)$" reports the difference of the moments from the EAJ$_2(3)$ and EXA$_2(3)$ models standardized by moments of the data.

<table>
<thead>
<tr>
<th>Conditional Moment</th>
<th>Data</th>
<th>EXA$_2(3)$</th>
<th>EAJ$_2(3)$</th>
<th>EAJ$_2(3)$ vs EXA$_2(3)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Upward</td>
<td>Downward</td>
<td>Upward</td>
<td>Downward</td>
</tr>
<tr>
<td>3-month</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>0.0435</td>
<td>0.0544</td>
<td>0.0376</td>
<td>0.0442</td>
</tr>
<tr>
<td>volatility</td>
<td>0.0179</td>
<td>0.0039</td>
<td>0.0281</td>
<td>0.0147</td>
</tr>
<tr>
<td>skewness</td>
<td>-0.2754</td>
<td>1.5649</td>
<td>-0.1235</td>
<td>0.3228</td>
</tr>
<tr>
<td>kurtosis</td>
<td>2.1179</td>
<td>6.9070</td>
<td>1.1623</td>
<td>2.8506</td>
</tr>
<tr>
<td>2-year</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>0.0502</td>
<td>0.0503</td>
<td>0.0482</td>
<td>0.0487</td>
</tr>
<tr>
<td>volatility</td>
<td>0.0168</td>
<td>0.0056</td>
<td>0.0274</td>
<td>0.0223</td>
</tr>
<tr>
<td>skewness</td>
<td>-0.2301</td>
<td>-1.1837</td>
<td>-0.1278</td>
<td>-0.3001</td>
</tr>
<tr>
<td>kurtosis</td>
<td>2.0856</td>
<td>7.3680</td>
<td>0.9336</td>
<td>2.6670</td>
</tr>
<tr>
<td>5-year</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>0.0573</td>
<td>0.0507</td>
<td>0.0532</td>
<td>0.0512</td>
</tr>
<tr>
<td>volatility</td>
<td>0.0146</td>
<td>0.0046</td>
<td>0.0228</td>
<td>0.0168</td>
</tr>
<tr>
<td>skewness</td>
<td>0.0335</td>
<td>0.1227</td>
<td>0.0165</td>
<td>0.0334</td>
</tr>
<tr>
<td>kurtosis</td>
<td>2.3891</td>
<td>6.2547</td>
<td>1.2556</td>
<td>2.9009</td>
</tr>
<tr>
<td>10-year</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>0.0623</td>
<td>0.0524</td>
<td>0.0613</td>
<td>0.0530</td>
</tr>
<tr>
<td>volatility</td>
<td>0.0133</td>
<td>0.0040</td>
<td>0.0202</td>
<td>0.0144</td>
</tr>
<tr>
<td>skewness</td>
<td>0.2626</td>
<td>0.9860</td>
<td>0.1321</td>
<td>0.2992</td>
</tr>
<tr>
<td>kurtosis</td>
<td>2.4191</td>
<td>6.7735</td>
<td>1.3008</td>
<td>2.9356</td>
</tr>
</tbody>
</table>
Table 8: GARCH (1,1) Parameters for the Model-Implied Libor Swap Rates

This table reports the maximum likelihood estimates of a GARCH (1,1) model \( \sigma_t^2 = \sigma + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \), where \( \varepsilon_t \) is the innovation from the AR(1) representation of the level of the yield) using 5-year yields simulated from ATSMs and AJTSMs. The “data” row presents the parameter estimates in Table 3 using sample data of the 5-year swap yield. The GARCH (1,1) parameter estimates for ATSMs and AJTSMs are computed as follows: First, 1000 sample paths of the 5-year yields are simulated from the models; second, the GARCH (1,1) model is estimated for each of the sample path; finally, the median of these 1000 parameter estimates are taken as the estimates of the model.

<table>
<thead>
<tr>
<th></th>
<th>0.0001</th>
<th>0.1679</th>
<th>0.7780</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Essentially ATSMs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EA0(3)</td>
<td>0.0008</td>
<td>0.4223</td>
<td>0.0382</td>
</tr>
<tr>
<td>EA1(3)</td>
<td>0.0003</td>
<td>0.1755</td>
<td>0.6906</td>
</tr>
<tr>
<td>EA2(3)</td>
<td>0.0002</td>
<td>0.1702</td>
<td>0.7311</td>
</tr>
<tr>
<td>EA3(3)</td>
<td>0.0001</td>
<td>0.1660</td>
<td>0.8033</td>
</tr>
<tr>
<td>Extended ATSMs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EXA1(3)</td>
<td>0.0003</td>
<td>0.1742</td>
<td>0.7004</td>
</tr>
<tr>
<td>EXA2(3)</td>
<td>0.0002</td>
<td>0.1706</td>
<td>0.7313</td>
</tr>
<tr>
<td>EXA3(3)</td>
<td>0.0001</td>
<td>0.1665</td>
<td>0.8010</td>
</tr>
<tr>
<td>Essentially AJTSMs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EAJ0(3)</td>
<td>0.0006</td>
<td>0.4073</td>
<td>0.0351</td>
</tr>
<tr>
<td>EAJ1(3)</td>
<td>0.0003</td>
<td>0.1719</td>
<td>0.7136</td>
</tr>
<tr>
<td>EAJ2(3)</td>
<td>0.0001</td>
<td>0.1690</td>
<td>0.7492</td>
</tr>
<tr>
<td>EAJ3(3)</td>
<td>0.0001</td>
<td>0.1675</td>
<td>0.7653</td>
</tr>
</tbody>
</table>
**Figure 1. Time Series of Libor Swap Rates**

This figure plots daily Libor Swap rates from August 13, 1990, to December 31, 2008, with a total of 4757 observations. The rates plotted have, from the lowest to the highest line with occasional cross overs when the yield curves are inverted, 3-month, 6-month, 9-month, 2-year, 3-year, 4-year, 5-year, 7-year, and 10-year maturities respectively.

**Figure 2. Time Series of Daily Changes of the 2-year and 10-year Swap Rates**

This figure plots daily changes in basis points of 2-year and 10-year Libor Swap rates from January 2006 to December 2008.
Figure 3. Estimated Coefficients in the “Yield Regression”

This figure plots coefficients estimates $\phi_{\tau-T}$ of $\phi_{\tau}$ in the “yield regression” $y_{t+1} = \text{constant} + \phi_{\tau} (y_t - r_t) / (\tau - 1) + \text{residual}$ using daily Libor Swap rates with maturities as indicated in Table 2 from August 13, 1990, to December 31, 2008. The three-month Libor rate is used as the spot rate $r_t$.

Figure 4. The Term Structure of Historical Volatilities

This figure reports the term structure of historical unconditional volatilities for daily Libor Swap rates with 3-month, 6-month, 9-month, 2-year, 3-year, 4-year, 5-year, 7-year, and 10-year maturities from August 13, 1990, to December 31, 2008. The volatility here is defined as the sample standard deviation of daily changes in the logarithm of yields.
Figure 5. Model-Implied Patterns for Time-Varying Risk Premiums: “Essentially” ATSMs

Panel A reports model-implied slope coefficients ($\phi_\tau$) of the population “yield regression” of $y_{t+1} - y_t$ onto $(y_t - r_t) / (\tau - 1)$ for the four three-factor “essentially” ATSMs, i.e., EA0(3), EA1(3), EA2(3), and EA3(3), defined in Section 2.3. These population coefficients are obtained by treating the parameter estimates as the true values and then applying analytic formulas to compute the involved moments in (6.1). Panel B reports model-implied slope coefficients $\phi^R_\tau$ from the “risk-adjusted yield regression” of $y_{t+1} - y_t + e_t / (\tau - 1)$ onto $(y_t - r_t) / (\tau - 1)$, where $e_t$ denotes the time-t expected excess holding period return on the zero-coupon bond with $\tau$ periods to mature. The coefficients $\phi^R_\tau$ are supposed to be one if the model can capture risk premiums. “Sample $\phi_{\tau,T}$” displays the estimated coefficients reported in Table 2 by the sample “yield regression” using daily Libor Swap rates from August 13, 1990, to December 31, 2008.

Figure 6. Model-Implied Term Structure of Volatilities: “Essentially” ATSMs

This figure plots the term structure of the sample variances, computed using simulated time series of daily changes in the logarithms of Libor Swap rates from the four three-factor “essentially” ATSMs (EA0(3), EA1(3), EA2(3), and EA3(3)) evaluated at their estimated parameter values. The “observed” line refers to those computed by the historical rates.
Figure 7. Model-Implied Patterns for Time-Varying Risk Premiums: “Extended” ATSMs

Panel A reports model-implied slope coefficients ($\phi_\tau$) of the population “yield regression” of $y_{t+1}^{(\tau-1)} - y_t^{(\tau-1)}$ onto $(y_t^{(\tau-1)} - r_t) / (\tau - 1)$ for the four three-factor “extended” ATSMs, i.e., EXA1(3), EXA2(3), and EXA3(3), defined in Section 2.3. These population coefficients are obtained by treating the parameter estimates as the true values and then applying analytic formulas to compute the involved moments in (6.1). Panel B reports model-implied slope coefficients $R_{\tau}$ from the “risk-adjusted yield regression” of $y_{t+1}^{(\tau-1)} - y_t^{(\tau-1)} + e_t^\tau / (\tau - 1)$ onto $(y_t^{(\tau-1)} - r_t) / (\tau - 1)$, where $e_t^\tau$ denotes the time-$t$ expected excess holding period return on the zero-coupon bond with $\tau$ periods to mature. The coefficients $R_{\tau}$ are supposed to be one if the model can capture risk premiums. “Sample $\phi_{\tau T}$” displays the estimated coefficients reported in Table 10 by the sample “yield regression” using daily Libor Swap rates from August 13, 1990, to December 31, 2008.

A: Model-Implied “Yield Regression”

B: Model-Implied “Risk-Adjusted Yield Regression”

Figure 8. Model-Implied Term Structure of Volatilities: “Extended” ATSMs

This figure plots the term structure of the sample variances, computed using simulated time series of daily changes in the logarithms of Libor Swap rates from the four three-factor “extended” ATSMs (EXA1(3), EXA2(3), and EXA3(3)) evaluated at their estimated parameter values. The “observed” line refers to those computed by the historical rates.
Figure 9. Model-Implied Patterns for Time-Varying Risk Premiums: “Essentially” AJTSMs

Panel A reports model-implied slope coefficients (\( \phi_{\tau} \)) from the population “yield regression” of \( (y_{t+1} - y_t) / (\tau - 1) \) onto \( (y_t - r_t) / (\tau - 1) \) for the three-factor “essentially” AJTSMs, i.e., EAJ0(3), EAJ1(3), EAJ2(3), and EAJ3(3). These population coefficients are obtained by treating the parameter estimates as the true values and then applying analytic formulas to compute the involved moments in (6.1). Panel B reports model-implied slope coefficients \( R_{\tau} \) from the “risk-adjusted yield regression” of \( (y_{t+1} - y_t) + e_{\tau} / (\tau - 1) \) onto \( (y_t - r_t) / (\tau - 1) \), where \( e_{\tau} \) denotes the time-\( t \) expected excess holding period return on the zero-coupon bond with \( \tau \) periods to mature. The coefficients \( R_{\tau} \) are supposed to be one if the model can capture risk premiums. “Sample \( \phi_{\tau T} \)” displays the estimated coefficients reported in Table 2 by the sample “yield regression” using daily Libor Swap rates from August 13, 1990, to December 31, 2008.

Figure 10. Model-Implied Term Structure of Volatilities: “Essentially” AJTSMs

This figure plots the term structure of the sample variances, computed using simulated time series of daily changes in the logarithms of Libor Swap rates from the three-factor “essentially” AJTSMs (EAJ0(3), EAJ1(3), EAJ2(3), and EAJ3(3)) evaluated at their estimated parameter values. The “observed” line refers to those computed by the historical rates.