Available online at www.sciencedirect.com
ScienceDirect
Journal of Financial Economics 83 (2007) 123–170
JOURNAL OF Financial ECONOMICS
www.elsevier.com/locate/jfec

Market price of risk specifications for affine models: Theory and evidence☆

Patrick Cheridito[a], Damir Filipović[b], Robert L. Kimmel[c,*]

[a]Princeton University, Department of Operations Research and Financial Engineering, Princeton, NJ 08544, USA
[b]University of Munich, Department of Mathematics, 80333 Munich, Germany
[c]Ohio State University, Fisher College of Business, Columbus, OH, 43210, USA

Received 14 December 2004; received in revised form 15 September 2005; accepted 23 September 2005
Available online 7 September 2006

Abstract

We extend the standard specification of the market price of risk for affine yield models, and apply it to U.S. Treasury data. Our specification often provides better fit, sometimes with very high statistical significance. The improved fit comes from the time-series rather than cross-sectional features of the yield curve. We derive conditions under which our specification does not admit arbitrage opportunities. The extension has extremely strong statistical significance for affine yield models with multiple square-root type variables. Although we focus on affine yield models, our specification can be used with other asset pricing models as well.

JEL classifications: C51; G12; G13

Keywords: Term structure; Market price of risk; Affine yield models; No-arbitrage pricing

☆We would like to thank Jun Liu, George Chacko, Chris Jones, an anonymous referee, seminar participants at the Canadian Mathematical Society, Princeton University, Northwestern University, the University of Illinois Urbana-Champaign, Duke University, Ohio State University, the Triangle Econometrics Workshop, the Econometric Society, the Centre for Advanced Studies in Finance at the University of Waterloo, the University of California Santa Barbara, the Western Finance Association, the Bachelier Society Third World Congress, and the Chinese Academy of Sciences Centre for Statistical Research for many helpful comments and suggestions. Any remaining errors are solely our responsibility. We would also like to thank Robert Bliss and Greg Duffee for the data set used in this study.
*Corresponding author. Fax: +1 614 292 2418.
E-mail address: kimmel.42@osu.edu (R.L. Kimmel).

0304-405X/$ - see front matter
doi:10.1016/j.jfineco.2005.09.008

1. Introduction

The square-root process of Feller (1951) has been used widely in financial economics, appearing in term structure models such as Cox et al. (1985) and stochastic volatility models of equity prices such as Heston (1993). Multivariate extensions of the square-root process have appeared in the term structure literature; see, for example, Duffie and Kan (1996), Dai and Singleton (2000), and Duffie (2002). The widespread use of this process is undoubtedly due at least in part to its relatively straightforward analytical properties: in the square-root process, a state variable follows a diffusion in which both the drift and the diffusion coefficients are affine functions of the state variable itself. Of course, a model for asset prices must specify not only the stochastic process followed by a set of underlying factors, but also the attitude of investors towards the risk of those factors. Since the pioneering work of Harrison and Kreps (1979) and Harrison and Pliska (1981), this task is often accomplished by specifying the behavior of the state variable(s) under both an objective probability measure and an equivalent martingale measure. A common practice is to have the state variables follow a Feller square-root process under both probability measures, but with different governing parameters.

This latter objective is normally met by assigning to each state variable a market price of risk process that is proportional to the square root of that state variable. Since the instantaneous volatility of each state variable is also proportional to its square root, the product of the market price of risk and volatility is proportional to the state variable itself. Subtraction of this product from the drift under the objective probability measure therefore results in a drift under the equivalent martingale measure that is also affine. If a process is within the Feller square-root class under the objective probability measure, this market price of risk specification ensures that it is within the same class under the equivalent martingale measure as well. A market price of risk that is inversely proportional to the square root of the state variable would also retain the affinity of the drift under both measures. However, such a market price of risk specification is rarely used in financial modeling.[1] Cox et al. (1985) discuss this possibility, and point out that it leads to arbitrage opportunities if the boundary value of the process can be achieved: while the instantaneous volatility of the state variable is zero at the boundary, if the market price of risk is inversely proportional to the square root of the state variable, the risk premium associated with that state variable does not go to zero as the volatility approaches zero. Ingersoll (1987) imposes the condition that the risk premium goes to zero as volatility goes to zero in a similar setting. Bates (1996) also imposes this condition in a stochastic volatility model; Chernov and Ghysels (2000), working in a similar setting, discuss the type of market price specification we propose, but leave unresolved the issue of whether it precludes arbitrage opportunities. In a recent term structure application, Duffee (2002) specifically avoids this market price of risk specification. However, whether or not a Feller square-root process can achieve the boundary value depends on the values of the governing parameters. For some parameter values, the instantaneous volatility of the state variable can approach zero arbitrarily closely but never actually achieve this value. The market price of risk can then

[1] In a model for stochastic volatility of equity prices, Eraker et al. (2003) consider a jump-diffusion model in which if the jump part of the model is ignored, the market price of risk is of this form. However, it is not explicitly identified and the absence of arbitrage is not formally demonstrated.

be arbitrarily large, but finite, when the state variable takes values near zero. It is not immediately clear whether arbitrage opportunities exist in this case; we show that they do not.

Although the reason for the avoidance of this market price of risk specification in the literature is not clear, it may be related to the difficulty of proving that it does not offer arbitrage opportunities. Specifically, it is quite difficult or impossible to determine whether this specification satisfies conventional criteria, e.g., those of Novikov or Kazamaki. However, these criteria are sufficient but not necessary to prove that the Girsanov ratio is a martingale. Using the approach of Cheridito et al. (2005), we show that this market price of risk specification does not offer arbitrage opportunities, provided certain parameter restrictions are observed. Using the extended market price of risk specification, we estimate several affine term structure models (specifically, all nine canonical families of affine models with one, two, or three factors, as Dai and Singleton, 2000, describe), and compare the results to those that obtain using more traditional market price of risk specifications. Although there are nine distinct canonical families, our extension is degenerate (i.e., is the same as the specification of Duffee, 2002) for three of the families. For the remaining six families, we find that the extended specification usually results in a statistically significant improvement in the fit of affine yield models to data on U.S. Treasury securities. The improvement is particularly strong for the three models with multiple semi-bounded state variables. To determine the cause of the statistical improvement, we explore both the fit of the cross-sectional shape of the yield curve (i.e., the difference between the shape of the yield curve predicted by an estimated model and the shape of the yield curve observed in the data) and the accuracy of the predicted time-series behavior. Our extended market price of risk specification appears to offer little improvement in the cross-sectional fit of the yield curve, and actually results in a slight degradation for some models. However, the time-series behavior predicted by models using the extended market price of risk specification is often substantially more accurate than that predicted by more traditional specifications. This improvement almost always manifests itself in reduced bias of yield forecasts, but for some models, the volatility of yield changes is also modelled much more accurately. Among three-factor models, some authors (e.g., Dai and Singleton, 2000, who introduce the model classification scheme and notation we use here) find that one particular model, the $A_1(3)$ model, captures many features of term structure behavior more accurately than other three-factor models. With the introduction of our market price of risk, the fit of the $A_1(3)$ model improves, but the fit of two other three-factor models, namely, the $A_2(3)$ and $A_3(3)$ models, improves substantially more. The relatively larger improvement in these latter two models could possibly reverse the preference ordering of three-factor models once the market price of risk is generalized.

The rest of this paper is organized as follows. In Section 2, we describe a class of multivariate term structure models driven by square-root processes, and define the admissible change of measure using our extended market price of risk specification. In Section 3, we show that this specification precludes arbitrage opportunities. In Section 4, we describe the data and estimation procedure we use to estimate and test our specification. In Section 5, we present the results and show that the extended market price of risk specification offers significantly better fit to the data than do standard specifications for most models, especially those with two or more square-root-type state variables. Finally, Section 6 concludes.

2. Models

Throughout, we focus on affine yield models of the term structure of interest rates, which we define as follows.

Definition 1. An affine yield model of the term structure of interest rates is a specification of interest rate and bond price processes such that:

1. The instantaneous interest rate r_t is an affine function of an N-vector of state variables denoted by Y_t,

$$r_t = d_0 + d^T Y_t, \tag{1}$$

where d_0 is a constant and d is an N-vector. We sometimes refer to individual elements of the vector y_t, using the notation $y_t(k)$ for $1 \leqslant k \leqslant N$.

2. The state variables Y_t follow the diffusion process:

$$\mathrm{d}Y_t = \mu^P(Y_t)\mathrm{d}t + \sigma(Y_t)\mathrm{d}W_t^P, \tag{2}$$

where $\mu^P(Y_t)$ is an N-vector, $\sigma(Y_t)$ is an $N \times N$ matrix, and W_t^P is an N-dimensional standard Brownian motion under the objective probability measure P.

3. The instantaneous drift (under the measure P) of each state variable is an affine function of Y_t,

$$\mu^P(Y_t) = a^P + b^P Y_t, \tag{3}$$

for some N-vector a^P and some $N \times N$ matrix b^P.

4. The instantaneous covariance between any pair of state variables is an affine function of Y_t,

$$\left[\sigma(Y_t)\sigma^T(Y_t)\right]_{i,j} = \alpha_{ij} + \beta_{ij}^T Y_t, \tag{4}$$

where $[\sigma(Y_t)\sigma^T(Y_t)]_{i,j}$ denotes the element in row i and column j of the product $\sigma(Y_t)\sigma^T(Y_t)$, α_{ij} is a constant, and β_{ij}^T is an N-vector for each $1 \leqslant i,j \leqslant N$.

5. There exists a probability measure Q, equivalent to P, such that Y_t is a diffusion under Q:

$$\mathrm{d}Y_t = \mu^Q(Y_t)\,\mathrm{d}t + \sigma(Y_t)\,\mathrm{d}W_t^Q, \tag{5}$$

where $\mu^Q(Y_t)$ is an N-vector, W_t^Q is an N-dimensional standard Brownian motion under Q, and the drift of each state variable is an affine function of the state vector

$$\mu^Q(Y_t) = a^Q + b^Q Y_t \tag{6}$$

for some N-vector a^Q and some $N \times N$ matrix b^Q.

6. Prices of zero-coupon bonds are conditional expectations of the discounted payoffs under the measure Q:

$$B(t,T) = \mathrm{E}_t^Q\left[\mathrm{e}^{-\int_t^T r_u \mathrm{d}u}\right]. \tag{7}$$

Feller (1951) treats existence of a process satisfying the second, third, and fourth conditions in a univariate setting, and Duffie and Kan (1996) do so in a multivariate setting. Duffie et al. (2003) provide a general mathematical characterization of affine processes, including those with jumps. The diffusions we consider here are special cases. Existence can essentially be characterized as a requirement that the state vector Y_t remain within a region in which $\sigma(Y_t)\sigma^T(Y_t)$ is positive semidefinite. More formally, it suffices that there exist constants $g_1, \ldots, g_M$ and nontrivial N-vectors $h_1, \ldots, h_M$ such that $\sigma(Y_t)\sigma^{\mathrm{T}}(Y_t)$ is positive definite[2] if and only if

$$g_i + h_i^T Y_t > 0 \tag{8}$$

for each value of $1 \leqslant i \leqslant M$. We denote the region in which this condition is satisfied (for all i) by D, and the closure of D by $\bar{D}$. Note that $\sigma(Y_t)\sigma^{\mathrm{T}}(Y_t)$ is positive definite in D, positive semidefinite in $\bar{D}$, and not positive semidefinite outside $\bar{D}$. Certain conditions must hold on the boundaries of D to ensure that the state vector cannot leave the region $\bar{D}$. For each value of $Y_t \in \bar{D}$, we must have:

$$\left(g_i + h_i^T Y_t = 0\right) \Rightarrow \left(h^T \mu^P(Y_t) \geqslant 0\right), \tag{9}$$

$$\left(g_i + h_i^T Y_t = 0\right) \Rightarrow \left(h^T \sigma(Y_t)\sigma^T(Y_t) h_i = 0\right) \tag{10}$$

for each value of i. Intuitively, these two requirements are (1) the drift must not pull the state vector Y_t out of the region $\bar{D}$, since $\sigma(Y_t)\sigma^T(Y_t)$ then fails to be positive semidefinite, and (2) the volatility must not allow Y_t to move stochastically out of $\bar{D}$. Of course, we must also have $Y_0 \in \bar{D}$. It is possible that $m = 0$, i.e., that D is the entire space $\mathbb{R}^N$, in which case the restrictions of Eqs. (9) and (10) are entirely vacuous. There are no separate uniqueness criteria; if a solution to Eq. (2) exists, it is unique.[3]

In addition to existence and uniqueness, achievement of boundary values is of particular importance when analyzing our market price of risk specification. Intuitively, within the region D, the drift of the state vector must not only satisfy the existence condition of Eq. (9), but also pull the state vector back toward the interior of D with sufficient strength to ensure that the boundary cannot be achieved. Feller (1951) and Ikeda and Watanabe (1981) treat the univariate case; Duffie and Kan (1996) treat the more complex multivariate case. However, possibly after changing the coordinate system, all the models we consider in this paper are such that the region D is of the form $(0, \infty)^M \times \mathbb{R}^{N-M}$, $M = 0, \ldots, N$, in which case it is easy to derive sufficient boundary nonattainment conditions from the one-dimensional case. We always impose boundary nonattainment conditions; we refer to the first M state variables as restricted and the last $N-M$ as unrestricted.

With respect to possible changes of the coordinate system, note that any transformation

$$X_t = A + BY_t \tag{11}$$

for some N-vector A and some regular $N \times N$ matrix B of a given affine yield model with state variables Y_t constitutes another affine yield model that can produce exactly the same short-rate processes r_t as the riginal model. To ensure identification of parameters in estimation, we impose additional restrictions; for example, we require that $\sigma(Y_t)$ be diagonal.[4]

[2]We assume the nondegeneracy condition, that the instantaneous covariance matrix of the state variables be full-rank for at least some value of the state vector.

[3]Throughout, "existence" refers to the existence of a weak solution, and "uniqueness" refers to uniqueness in distribution.

[4]This normalization is one of several that Dai and Singleton (2000) use. The question of which affine processes can be represented with a diagonal diffusion matrix by a change of variables is addressed by Cheridito, et al.

Although in Eq. (7), we characterize bond prices as conditional expectations (under the Q measure), in practice, bond prices are usually calculated as solutions to a partial differential equation, which, for the affine models we consider here, is equivalent to a system of Riccati-type ordinary differential equations (see Duffie, et al. 2003). The Feynman-Kac theorem, which establishes the equivalence of the probabilistic problem and the partial differential equation problem, is well known and frequently applied to affine term structure models. However, its applicability to bond prices under some families of affine models has been formally justified only recently; see Levendorskii (2004a) for the affine diffusion case, and Levendorskii (2004b) for the case of affine processes with jumps. For general payoff functions, the applicability of the Feynman-Kac theorem remains an open issue; for bond prices, Levendorskii (2004a) establishes sufficient conditions for the applicability of the Feynman-Kac theorem for all models we consider. Grasselli and Tebaldi (2004) establish necessary and sufficient conditions for the existence of closed-form solutions to the partial differential equation (which, as we state above, is equivalent to a system of Riccati-type ordinary differential equations) for affine yield models.

Given a specification of $\mu^P(Y_t)$ and $\sigma(Y_t)$ such that a solution to Eq. (2) exists, we may define an equivalent probability measure

$$Q = \exp\left(-\int_0^T \lambda^T(Y_u)\,\mathrm{d}W_u^P - \frac{1}{2}\int_0^T \lambda^T(Y_u)\lambda(Y_u)\,\mathrm{du}\right)P \tag{12}$$

by specifying a market price of risk process $\lambda(Y_t)$ that satisfies the condition

$$E^P\left[\exp\left(-\int_0^T \lambda^T(Y_u)\,\mathrm{d}W_u^P - \frac{1}{2}\int_0^T \lambda^T(Y_u)\lambda(Y_u)\,\mathrm{d}u\right)\right] = 1. \tag{13}$$

It follows from Girsanov's theorem that the process $W_t^Q = W_t^P + \int_0^t \lambda(Y_s)\,\mathrm{d}s$ is an N-dimensional Brownian motion under Q, and

$$\mathrm{d}Y_t = \mu^Q(Y_t)\,\mathrm{d}t + \sigma(Y_t)\,\mathrm{d}W_t^Q, \tag{14}$$

where $\mu^Q(Y_t)$ is given by

$$\mu^Q(Y_t) = \mu^P(Y_t) - \sigma(Y_t)\lambda(Y_t). \tag{15}$$

Numerous sufficiency criteria, such as those of Novikov and Kazamaki (see, for example, Revuz and Yor, 1994) have been developed to show that a given stochastic exponential satisfies Eq. (13). Dai and Singleton (2000) consider a simple market price of risk specification,

$$\lambda(Y_t) = \sigma^T(Y_t)\lambda, \tag{16}$$

where λ is a vector of constants. By construction, this specification ensures that $\mu^Q(Y_t)$ is an affine function of Y_t. When $\sigma^T(Y_t)$ does not depend on Y_t, this market price of risk

(*footnote continued*)
(2005), who find that any affine diffusion defined on a state space $(0,\infty)^M \times \mathbb{R}^{N-M}$ (after affine transformation of the state variables) with $M \leqslant 1$ or $M \geqslant N-1$ can be diagonalized, with the transformed process taking values in the same state space. They also give examples of diffusions with $M = 2$ and $N = 4$ whose diffusion matrices cannot be diagonalized by affine transformation. However, in this paper, we consider only $N \leqslant 3$, in which case at least one of the conditions $M \leqslant 1$ or $M \geqslant N-1$ is always satisfied. Thus, our assumption of a diagonal diffusion matrix does not result in loss of generality.

specification satisfies the Novikov criterion for any time interval $[s, t]$. The Novikov criterion may also be satisfied for any time interval even when $\sigma^T(Y_t)$ does depend on Y_t, depending on the values of the model parameters. However, in general, the Dai and Singleton (2000) market price of risk specification only satisfies the Novikov criterion on $[s, t]$ when $t < s + \varepsilon$ for some positive ε. The value of ε depends on the model parameters, not on s or Y_s. However, this form of local satisfaction of the Novikov criterion is sufficient for the satisfaction of Eq. (13) (see, for example, Corollary 5.14 in Karatzas and Shreve, 1991).

Duffee (2002) refers to models with the market price of risk specification of Dai and Singleton (2000) as *completely affine*, and introduces the more general class of *essentially affine* models. The only constraint on the market price of risk specification in essentially affine models can be characterized as follows: if a linear combination of state variables is restricted, then the market price of risk of that linear combination must coincide with the completely affine specification. By contrast, a linear combination of state variables that is unrestricted can have any market price of risk consistent with affine dynamics under both measures. For example, in the univariate model

$$\mathrm{d}Y_t = \left(a^P + b^P Y_t\right)\mathrm{d}t + \sigma \mathrm{d}W_t^P, \tag{17}$$

the single state variable is unrestricted, so $\lambda(Y_t)$ can be any affine function of Y_t. By contrast, in the univariate model

$$\mathrm{d}Y_t = \left(a^P + b^P Y_t\right)\mathrm{d}t + \sigma\sqrt{Y_t}\,\mathrm{d}W_t^P, \tag{18}$$

the single state variable is restricted. Consequently, the essentially affine market price of risk for this model is $\lambda(Y_t) = \lambda\sqrt{Y_t}$ for some constant λ (with $\lambda = 0$ possible). In other words, $\lambda(Y_t)$ is restricted to ensure that, if the volatility of any linear combination of state variables approaches zero, the risk premium of that linear combination also approaches zero. As with the completely affine market price of risk specification, the essentially affine specification satisfies the Novikov criterion for some finite positive time interval (the size of which depends on the model parameters, but not on the initial state vector), thereby ensuring satisfaction of Eq. (13).

Our market price of risk specification, by contrast, imposes only those restrictions necessary to ensure that the boundary nonattainment conditions are satisfied under both the P and Q measures. In Section 3, we show that this requirement is sufficient to ensure that the market price of risk specification satisfies Eq. (13). Note that the essentially affine specification nests the completely affine market price of risk, and our specification, which we refer to as the *extended affine* market price of risk, always nests both the completely affine and essentially affine specifications. The completely and essentially affine specifications coincide for some models, as do the the essentially and extended affine specifications. However, the extended affine specification is always more general than the completely affine specification.

While affine yield models are treated in a systematic way by Duffe and Kan (1996), many other models in the literature, such as Vasicek (1977), Cox, et al. (1985), Balduzzi, et al. (1996), and Chen (1996), are special cases of the general affine model. Dai and Singleton (2000) note that for each integer $N \geqslant 1$, there exist $N+1$ nonnested families of N-factor affine yield models, and develop the classification scheme that we use below. Each affine yield model can be placed into a family, designated $A_M(N)$, where N is the number of state variables and M is the number of linearly independent $\beta_{ij}, 1 \leqslant i, j \leqslant N$ with M

necessarily taking on values from 0 to N. The $A_M(N)$ model contains M state variables that are restricted. Each of these state variables follows a process similar to the Feller square-root process, except that the drift of one restricted state variable may depend on the value of another restricted state variable. The remaining $M-N$ state variables are unrestricted. The unrestricted state variables jointly follow a process similar to a multivariate Ornstein-Uhlenbeck process, but with two modifications: both the drift and the variance of an unrestricted state variable may depend on the values of the restricted state variables.

For now, we take as given that our market price of risk specification is free from arbitrage, and examine in detail each of the single-factor, two-factor, and three-factor affine yield models that we estimate. In addition to specifying both the dynamics of the state variables under both the P and Q measures and the definition of the interest rate process, we specify any parameter restrictions needed to ensure existence of the specified process or to ensure restricted state variables do not achieve their boundary values. We also identify any restrictions needed to make sure that a model has a unique representation.

2.1. Single-factor models

In a single-factor affine yield model, the interest rate process is specified as

$$r_t = d_0 + d_1 Y_t(1) \tag{19}$$

for some constants d_0 and d_1. However, the state variable $Y_t(1)$ can follow one of two distinct types of diffusions, namely, the $A_0(1)$ or $A_1(1)$ model (as per Dai and Singleton, 2000). In the $A_0(1)$ model, $Y_t(1)$ follows the process

$$\mathrm{d}Y_t(1) = \left[b_{11}^P Y_t(1)\right]\mathrm{d}t + \mathrm{d}W_t^P(1), \tag{20}$$

where is $W_t^P(1)$ a standard Brownian motion under the objective measure P, and b_{11}^P is an arbitrary constant. Note that this process has no constant term in the drift, and the diffusion coefficient has been normalized to one. These restrictions are not a loss of generality, but rather a normalization that ensures a unique representation of the model. Any process with an affine drift and constant diffusion can be transformed into a process of this type by an affine transformation of the state variable. An observationally equivalent interest rate model results by making an appropriate change to the d_0 and d_1 coefficients. Under the measure Q, the process $Y_t(1)$ can be written as

$$\mathrm{d}Y_t(1) = \left[a_1^Q + b_{11}^Q Y_t(1)\right]\mathrm{d}t + \mathrm{d}W_t^Q(1), \tag{21}$$

where $W_t^Q(1)$ is a standard Brownian motion under Q. The process exists for any value of b_{11}^P; furthermore, there is no finite boundary value (i.e., the process Y_t (1) an take on any real value), and the boundaries at infinity are unattainable in finite time, regardless of the parameter values. The market price of risk process is defined by

$$\Lambda_t = [\sigma(Y_t)]^{-1}\left[\mu^P(Y_t) - \mu^Q(Y_t)\right] = -a_1^Q\left(b_{11}^P - b_{11}^Q\right)Y_t(1) \equiv \lambda_{10} + \lambda_{11} Y_t(1). \tag{22}$$

In the completely affine models of Dai and Singleton (2000), the λ_{11} parameter is restricted to be zero. By contrast, in the essentially affine models of Duffee (2002), the λ_{10} and λ_{11} parameters can take any values. Existence of the measure Q with either the completely affine or essentially affine market price of risk specification follows from satisfaction of the Novikov criterion for a finite positive time interval, as we discuss above. For the $A_0(1)$ model, our market price of risk specification coincides with the essentially affine specification, offering no further generality.

The $A_1(1)$ model is based on the square-root process of Feller (1951). Under this specification, the process $Y_t(1)$ can be expressed as

$$\mathrm{d}Y_t(1) = \left[a_1^Q + b_{11}^Q Y_t(1)\right]\mathrm{d}t + \sqrt{Y_t(1)}\mathrm{d}W_t^P(1), \tag{23}$$

where $W_t^P(1)$ is a standard Brownian motion under the objective measure P. Note that the diffusion term may be taken to be Y_t itself, rather than some affine function of Y_t, by an appropriate change of variables, as we describe above. Existence of such a process requires only that $a_1^P \geqslant 0$. $Y_t(1)$ is bounded below by zero; this state variable cannot achieve its boundary value if $2a_1^P \geqslant 1$. Under the measure Q, the process $Y_t(1)$ can be written as

$$\mathrm{d}Y_t(1) = \left[a_1^Q + b_{11}^Q Y_t(1)\right]\mathrm{d}t + \sqrt{Y_t(1)}\mathrm{d}W_t^Q(1), \tag{24}$$

where $W_t^Q(1)$ is a standard Brownian motion under the measure Q. The market price of risk process is defined as

$$\begin{aligned}\Lambda_t = [\sigma(Y_t)]^{-1}\left[\mu^P(Y_t) - \mu^Q(Y_t)\right] &= \frac{a_1^P - a_1^Q}{\sqrt{Y_t(t)}} + \left(b_{11}^P - b_{11}^Q\right)\sqrt{Y_t(t)} \\ &\equiv \frac{\lambda_{10}}{\sqrt{Y_t(t)}} + \lambda_{11}\sqrt{Y_t(t)}.\end{aligned} \tag{25}$$

The completely affine and essentially affine specifications coincide for the $A_1(1)$ model; in both, the λ_{11} parameter can take any arbitrary value, but the λ_{10} parameter is restricted to be zero. For each value of λ_{11}, the Novikov criterion is satisfied for some finite positive time horizon. We permit λ_{10} to take on any value such that boundary nonattainment conditions are satisfied under Q as well as P. This requirement can be expressed as

$$\lambda_{10} \leqslant a_1^P - \frac{1}{2}. \tag{26}$$

It is unclear whether this specification satisfies the traditional Novikov and Kazamaki criteria; in Section 3, we use another method to show that it satisfies Eq. (13).

2.2. Two-factor models

Two-factor affine yield models have an interest rate process given by:

$$r_t = d_0 + d_1 Y_t(1) + d_2 Y_t(2), \tag{27}$$

where the process followed by $Y_t(1)$ and $Y_t(2)$ falls into one of three categories, the $A_0(2)$, $A_1(2)$, or $A_2(2)$ family. The P-measure dynamics for the $A_0(2)$ model are

$$d\begin{bmatrix} Y_t(1) \\ Y_t(2) \end{bmatrix} = \begin{bmatrix} b_{11}^P & b_{12}^P \\ b_{21}^P & b_{22}^P \end{bmatrix}\begin{bmatrix} Y_t(1) \\ Y_t(2) \end{bmatrix}\mathrm{d}t + d\begin{bmatrix} W_t^P(1) \\ W_t^P(2) \end{bmatrix}. \tag{28}$$

These dynamics reflect any change of variables necessary to ensure that the matrix $\sigma(Y_t)$ is an identity matrix and the constant terms in the drifts of the state variables are zero.

Even with these normalizations, however, the $A_0(2)$ representation is not unique, as a new set of state variables can be formed by taking any orthogonal rotation of the old state variables. Dai and Singleton (2000) choose the identification restriction $b_{12}^P = 0$, which guarantees a unique representation whenever the two components of Y_t are not independent, i.e., when the normalization does not also cause the b_{21}^P parameter to be zero. If the normalization causes both b_{12}^P and b_{21}^P to be zero, then a reordering of the state variable indices is also possible. This method of normalization also precludes b matrices with eigenvalues that are complex conjugate pairs.[5] Under the measure Q, the process followed by Y_t is given by

$$d\begin{bmatrix} Y_t(1) \\ Y_t(2) \end{bmatrix} = \left(\begin{bmatrix} a_1^Q \\ a_2^Q \end{bmatrix} + \begin{bmatrix} b_{11}^Q & b_{12}^Q \\ b_{21}^Q & b_{22}^Q \end{bmatrix} \begin{bmatrix} Y_t(1) \\ Y_t(2) \end{bmatrix} \right) \mathrm{d}t + d\begin{bmatrix} W_t^Q(1) \\ W_t^Q(2) \end{bmatrix}. \tag{29}$$

No parameter restrictions are needed to ensure the existence of the process, or of the Q measure. Furthermore, there are no finite boundaries and no additional boundary nonattainment conditions. The market price of risk specification is

$$\begin{aligned} \Lambda_t &= [\sigma(Y_t)]^{-1}\left[\mu^P(Y_t) - \mu^Q(Y_t)\right] \\ &= \left(-\begin{bmatrix} a_1^Q \\ a_2^Q \end{bmatrix} + \begin{bmatrix} b_{11}^P - b_{11}^Q & b_{12}^P - b_{12}^Q \\ b_{21}^P - b_{21}^Q & b_{22}^P - b_{22}^Q \end{bmatrix} \begin{bmatrix} Y_t(1) \\ Y_t(2) \end{bmatrix} \right) \\ &\equiv \begin{bmatrix} \lambda_{10} \\ \lambda_{20} \end{bmatrix} + \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{bmatrix} \begin{bmatrix} Y_t(1) \\ Y_t(2) \end{bmatrix}. \end{aligned} \tag{30}$$

The completely affine market price of risk specifications restricts λ_{11}, λ_{12}, λ_{21}, and λ_{22} to be zero. The essentially affine specification relaxes these restrictions, and allows all six market price of risk parameters to take on arbitrary values. Both of these specifications satisfy the Novikov criterion for a finite positive time interval, thereby ensuring that the specified Q measure exists and is equivalent to P. For the $A_0(2)$ model, our specification coincides with the essentially affine market price of risk, offering no further flexibility.

The P measure dynamics of the $A_1(2)$ model are given by

$$\begin{aligned} d\begin{bmatrix} Y_t(1) \\ Y_t(2) \end{bmatrix} &= \left(\begin{bmatrix} a_1^P \\ 0 \end{bmatrix} + \begin{bmatrix} b_{11}^P & 0 \\ b_{21}^P & b_{22}^P \end{bmatrix} \begin{bmatrix} Y_t(1) \\ Y_t(2) \end{bmatrix} \right) \mathrm{d}t \\ &\quad + \begin{bmatrix} \sqrt{Y_t(1)} & 0 \\ 0 & \sqrt{\alpha_2 + \beta_{21} Y_t(1)} \end{bmatrix} d\begin{bmatrix} W_t^Q(1) \\ W_t^Q(2) \end{bmatrix}, \end{aligned} \tag{31}$$

where $\alpha_1 \in \{0.1\}$. Existence of this process requires that $a_1^P \geqslant 0$ and $\beta_{21}^P \geqslant 0$. The process $Y_t(1)$ is bounded from below by zero; the additional restriction $2a_1^P \geqslant 1$ 1s needed to ensure that $Y_t(1)$ does not achieve the boundary value. The dynamics under the measure Q for the

[5]Depending on the number and the maturities of the bond yields observed, identification issues may arise when some of the eigenvalues of the slope matrix in the drift are complex. See Beaglehole and Tenney (1991).

$A_1(2)$ model are given by

$$d\begin{bmatrix} Y_t(1) \\ Y_t(2) \end{bmatrix} = \left(\begin{bmatrix} a_1^Q \\ a_2^Q \end{bmatrix} + \begin{bmatrix} b_{11}^Q & 0 \\ b_{21}^Q & b_{22}^Q \end{bmatrix} \begin{bmatrix} Y_t(1) \\ Y_t(2) \end{bmatrix} \right) \mathrm{d}t + \begin{bmatrix} \sqrt{Y_t(1)} & 0 \\ 0 & \sqrt{\alpha_2 + \beta_{21} Y_t(1)} \end{bmatrix} d\begin{bmatrix} W_t^Q(1) \\ W_t^Q(2) \end{bmatrix}. \tag{32}$$

Note that both b_{12}^P and b_{12}^Q are constrained to be zero. In the $A_0(2)$ model, the constraint on b_{12}^P is to ensure identification; for the essentially affine market price of risk specifications, there is no corresponding restriction under the Q measure. By contrast, the restriction here is for existence of the process under the P measure, and for the existence of the Q measure. Intuitively, the drift of $Y_t(1)$ cannot depend on $Y_t(2)$, since $Y_t(2)$ can take on any value, positive or negative, whereas $Y_t(1)$ must remain nonnegative for the diffusion matrix to remain positive semidefinite. A nonzero value for b_{12}^P would give the drift of $Y_t(1)$ the wrong sign sometimes, allowing the $Y_t(1)$ process to have a drift in the wrong direction when it is at the boundary. This restriction must therefore be imposed under both measures.

The market price of risk process is given by

$$\begin{aligned} \Lambda_t &= [\sigma(Y_t)]^{-1}\left[\mu^P(Y_t) - \mu^Q(Y_t)\right] \\ &= \begin{bmatrix} \frac{(a_1^P - a_1^Q)}{\sqrt{Y_t(1)}} + \left(b_{11}^P - b_{11}^Q\right)\sqrt{Y_t(1)} \\ \frac{(-a_2^Q) + \left(b_{21}^P - b_{21}^Q\right) Y_t(1) + \left(b_{22}^P - b_{22}^Q\right) Y_t(2)}{\sqrt{\alpha_2 + \beta_{21} Y_t(1)}} \end{bmatrix} \\ &\equiv \begin{bmatrix} \frac{\lambda_{10}}{\sqrt{Y_t(1)}} + \lambda_{11}\sqrt{Y_t(1)} \\ \frac{\lambda_{20} + \lambda_{21} Y_t(1) + \lambda_{22} Y_t(2)}{\sqrt{\alpha_2 + \beta_{21} Y_t(1)}} \end{bmatrix}. \end{aligned} \tag{33}$$

Previous studies of affine yield models impose some restrictions on the market price of risk parameters of the $A_1(2)$ model. The completely affine market price of risk allows λ_{11}, λ_{20} and λ_{21} to be nonzero, but requires that λ_{20} and λ_{21} satisfy $\beta_{21}\lambda_{20} = \lambda_{21}\alpha_2$, so only two parameters can be chosen independently. In essentially affine models, all parameters except λ_{10} can be nonzero.[6] Both of these specifications satisfy the Novikov criterion at least for some finite positive time interval. We permit all parameters to be nonzero, requiring only that boundary nonattainment conditions for Y_t be satisfied under the measure Q. This holds if

$$\lambda_{10} \leqslant a_1^P - \frac{1}{2}. \tag{34}$$

[6] It should be noted that neither Dai and Singleton (2000) nor Duffiee (2002) permit $\alpha_2 = 0$. However, the requirement that $\alpha_2 = 1$ appears to be of little consequence, since a diffusion with a very small value of α_2 can be converted to one with $\alpha_2 = 1$ by an affine change of the state vector. Therefore, these authors preclude models with $\alpha_2 = 0$, but not models in which the α_2 parameter is effectively arbitrarily close to zero.

When λ_{10} is nonzero, it is unclear whether this specification satisfies the Novikov or the Kazamaki criterion.

The dynamics under the measure P of the $A_2(2)$ model are given by

$$d\begin{bmatrix} Y_t(1) \\ Y_t(2) \end{bmatrix} = \left(\begin{bmatrix} a_1^P \\ a_2^P \end{bmatrix} + \begin{bmatrix} b_{11}^P & b_{12}^P \\ b_{21}^P & b_{22}^P \end{bmatrix} \begin{bmatrix} Y_t(1) \\ Y_t(2) \end{bmatrix} \right) \mathrm{d}t + \begin{bmatrix} \sqrt{Y_t(1)} & 0 \\ 0 & \sqrt{Y_t(2)} \end{bmatrix} d\begin{bmatrix} W_t^P(1) \\ W_t^P(2) \end{bmatrix}, \tag{35}$$

with existence constraints $a_1^P \geqslant 0$, $a_2^P \geqslant 0$, $b_{12}^P \geqslant 0$, and $b_{21}^P \geqslant 0$. Both state variables are bounded from below by zero; boundary nonattainment conditions are $2a_1^P \geqslant 1$ and $2a_2^P \geqslant 1$. The diagonal form of the diffusion matrix is a result of the normalization procedure; apart from a reordering of indices, each $A_2(2)$ model has a unique representation. Dynamics under the measure Q are given by

$$d\begin{bmatrix} Y_t(1) \\ Y_t(2) \end{bmatrix} = \left(\begin{bmatrix} a_1^Q \\ a_2^Q \end{bmatrix} + \begin{bmatrix} b_{11}^Q & b_{12}^P \\ b_{21}^Q & b_{22}^P \end{bmatrix} \begin{bmatrix} Y_t(1) \\ Y_t(2) \end{bmatrix} \right) \mathrm{d}t + \begin{bmatrix} \sqrt{Y_t(1)} & 0 \\ 0 & \sqrt{Y_t(2)} \end{bmatrix} d\begin{bmatrix} W_t^P(1) \\ W_t^P(2) \end{bmatrix}, \tag{36}$$

The market price of risk process is defined as

$$\begin{aligned} \Lambda_t &= [\sigma(Y_t)]^{-1} \left[\mu^P(Y_t) - \mu^Q(Y_t)\right] \\ &= \begin{bmatrix} \dfrac{\left(a_2^P - a_2^Q\right) + \left(b_{11}^P - b_{11}^Q\right) Y_t(1) + \left(b_{12}^P - b_{12}^Q\right) Y_t(2)}{\sqrt{Y_t(1)}} \\ \dfrac{\left(a_2^P - a_2^Q\right) + \left(b_{21}^P - b_{21}^Q\right) Y_t(1) + \left(b_{22}^P - b_{22}^Q\right) Y_t(2)}{\sqrt{Y_t(2)}} \end{bmatrix} \\ &\equiv \begin{bmatrix} \dfrac{\lambda_{10} + \lambda_{11} Y_t(1) + \lambda_{12} Y_t(2)}{\sqrt{Y_t(1)}} \\ \dfrac{\lambda_{10} + \lambda_{21} Y_t(1) + \lambda_{22} Y_t(2)}{\sqrt{Y_t(2)}} \end{bmatrix}. \end{aligned} \tag{37}$$

Completely affine and essentially affine market price of risk specifications coincide for the $A_2(2)$ model. In both, only the λ_{11} and λ_{22} parameters can be nonzero. This specification satisfies the Novikov criterion for a finite positive time interval (which depends on the model parameters). By contrast, our specification permits all six parameters to be nonzero, with only the boundary nonattainment conditions under the measure Q restricting their values. These conditions are more complex than in the $A_1(2)$ model:

$$\lambda_{10} \leqslant a_1^P - \frac{1}{2}, \tag{38}$$

$$\lambda_{20} \leqslant a_2^P - \frac{1}{2}, \tag{39}$$

$$\lambda_{12} \leqslant b_{12}^P, \tag{40}$$

$$\lambda_{21} \leqslant b_{21}^P. \tag{41}$$

This specification cannot easily be shown to satisfy either the Novikov and Kazamaki criteria for any finite positive time interval.

2.3. Three-factor models

There are four distinct families of three-factor models, specifically, the $A_0(3)$, $A_1(3)$, $A_2(3)$, and $A_3(3)$ models. Although many of the properties of these models are analogous to those of one- and two-factor models, the existence of three factors allows for a much richer interplay of the factors; for example, cross-terms in the drift between restricted state variables and dependence of the diffusion of unrestricted state variables on the value of restricted state variables can occur in the same model. Relative to the essentially affine specification, our extended affine market price of risk specification introduces no new parameters for the $A_0(3)$ model, one new parameter for the $A_1(3)$ model, four new parameters for the $A_2(3)$ model, and nine new parameters for the $A_3(3)$ model. Relative to the completely affine specification, the extended specification adds nine parameters for the $A_0(3)$ and $A_3(3)$ models, and seven new parameters for the $A_1(3)$ and $A_2(3)$ models. Due to the complexity of three-factor models, we provide a full characterization in Appendix B rather than here.

2.4. General comments

At this point, some general comments on the market price of risk parameter are appropriate. Both the completely and essentially affine market price of risk specifications permit only the speed of mean reversion for the restricted state variables to differ between the P and Q measures; the constant term in the drift, as well as the slope terms on other restricted state variables, remain the same. For example, if the drift of $Y_t(2)$ (assumed to be restricted) is given by $a_2^P + b_{21}^P Y_t(1) + b_{22}^P Y_t(2)$, only the b_{22}^Q parameter may differ from its P-measure counterpart. Thus, the risk premium associated with a restricted state variable is not only constrained to depend on its own current level, it must also depend on its level in a very particular way, so that the constant term in the drift does not change with the measure change. The extended affine market price of risk allows the constant term to change as well, so that the unconditional mean of the process can change independently of the speed of mean reversion. However, the extended affine specification is more general; it allows the risk premium of a restricted state variable to depend on other restricted state variables as well. A number of interesting possibilities, which are impossible with the more traditional market price of risk specification, can therefore occur. For example, consider a two-factor model with at least one restricted state variable $Y_t(1)$. If the interest rate does not depend on the second state variable (i.e., if d_2 is equal to zero), and the second state variable is unrestricted, then it can have no effect on either the shape of the yield curve or the time-series behavior of yields. The unrestricted state variable $Y_t(2)$ does not affect the interest rate directly, and it cannot affect it indirectly either, because the dynamics of a restricted state variable cannot depend on an unrestricted state variable. If $Y_t(2)$ is restricted, then it can affect (and also be affected by) $Y_t(1)$ through the cross-terms in the drift slope matrix (b_{12} and b_{21}). But with traditional market prices of risk, these parameters must be the same under both measures. With the extended affine market price of risk, these parameters can be zero under the Q measure, but nonzero under the P measure, in which case the second state variable has no effect on the shape of the yield curve because it does not affect the interest rate directly, and under the Q measure, it does not affect $Y_t(1)$ either. However, the second state variable could affect the time-series properties of the yield curve through its presence in the P-measure drift of the first state variable. We therefore have the possibility of a nondegenerate $A_2(2)$ model in which the value of one of the state variables cannot be determined from only the shape of the term structure. This situation cannot occur with the traditional market price of risk. Many other such situations can be contemplated, owing to the rich interplay between state variables not only under the Q measure, but also in the risk premium, that becomes possible with the extended affine specification.

3. Absence of arbitrage

The relation between the absence of arbitrage and existence of an equivalent martingale measure is well known, with the foundational work of Harrison and Kreps (1979) and Harrison and Pliska (1981) extended by many, such as Delbaen and Schachermayer (1994) and Delbaen and Schachermayer (1998). However, the standard techniques used to demonstrate the existence of an equivalent probability measure do not work well with our extended market price of risk specification. For example, it is not clear whether the Novikov and Kazamaki criteria are satisfied. As a restricted state variable approaches its boundary value, the extended affine specification allows the market price of risk of that state variable to grow (positively or negatively) without bound. Simply being unbounded is not necessarily a problem; for example, the standard market price of risk specification in the model of Cox, et al. (1985) grows without bound as the interest rate becomes very large. However, the market price of risk in this model, although unbounded, grows slowly enough with increasing interest rates to allow application of the Novikov and Kazamaki criteria. The extended affine market price of risk grows more quickly near the zero boundary than traditional specifications do near the infinity boundary. We must therefore use another approach, for instance, that of Cheridito et al. (2005), to demonstrate that our specification precludes arbitrage opportunities.

Theorem 1. Let $\mu^P(\cdot)$, $\mu^Q(\cdot)$, and $\sigma(\cdot)$ be functions of the form specified in Eqs. (3), (6), and (4), respectively, such that both pairs (μ^P, σ) and (μ^Q, σ) satisfy the existence conditions 8 through 10 and the boundary nonattainment conditions. Then the following three statements hold:

(a) There exists a probability space (Ω, F, P) supporting a Brownian motion $(W_t^P)t \geqslant 0$ such that for each $Y_0 \in D$, there exists a stochastic process $(Y_t)_{t \geqslant 0}$ on (Ω, F, P) satisfying

$$Y_t = Y_0 + \int_0^t \mu^P(Y_s)\,\mathrm{d}s + \int_0^t \sigma(Y_s)\,\mathrm{d}W_s^P, \quad t \geqslant 0. \tag{42}$$

(b) The distribution of $(Y_t)_{t \geqslant 0}$ under P is unique.

(c) For each $T > 0$, there exists *a* measure Q equivalent to P such that

$$Y_t = Y_0 + \int_0^t \mu^Q(Y_s)\,\mathrm{d}s + \int_0^t \sigma(Y_s)\,\mathrm{d}W_s^Q, \quad t \in [0, T], \tag{43}$$

where $(W_t^Q)_{t \in [0,T]}$ is a Brownian motion under Q.

Proof:. See Appendix A.

The term structure literature, from the first use of the square-root process in Cox et al. (1985) to recent work by Duffee (2002), quite explicitly avoids market price of risk specifications that do not go to zero as the volatility of the corresponding state variable goes to zero. Theorem 1 demonstrates that this restriction can be relaxed, provided the parameters of the model do not permit attainment of the boundary under either probability measure. In this case, the market price of risk can become arbitrarily large; however, since the boundary is not achieved, it always remains finite. If the boundary nonattainment conditions are satisfied under one of the P or Q measures, but not the other, then the two measures clearly cannot be equivalent. In this case, the measure under which the boundary cannot be achieved is absolutely continuous with respect to the measure under which the boundary can be achieved. However, absolute continuity is not sufficient to preclude arbitrage opportunities.

From Theorem 1, we can construct arbitrage-free models simply by ensuring that the existence and boundary non-attainment conditions are satisfied under both measures. This result allows for considerable flexibility, especially when there are several square-root-type state variables in a model. The dynamics of a square-root type variable (we drop the superscript notation indicating the measure for purposes of this example) in a canonical affine diffusion are given by

$$\mathrm{d}Y_t = (a_1 + b_{11} Y_t)\,\mathrm{d}t + \sqrt{Y_t}\,\mathrm{d}W_t. \tag{44}$$

Traditional market price of risk specifications permit only the slope coefficient, b_{11}, to differ under the two probability measures. Our specification allows both the slope and constant terms, a_1 and b_{11}, to differ, provided $2a_1 \geqslant 1$ under both measures. With two square-root-type variables, the dynamics are

$$\mathrm{d}Y_t = \left(\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} Y_t \right) \mathrm{d}t + \begin{bmatrix} \sqrt{Y_t(1)} & 0 \\ 0 & \sqrt{Y_t(2)} \end{bmatrix} d \begin{bmatrix} W_t(1) \\ W_t(2) \end{bmatrix}. \tag{45}$$

Traditional market price of risk specifications permit only b_{11} and b_{12} to change under the two measures; our specification permits all six drift parameters to change, provided $b_{12} \geqslant 0$ and $b_{21} \geqslant 0$ (for existence), and $2a_1 \geqslant 1$ and $2a_2 \geqslant 1$ (for boundary nonattainment). The extended affine market price of risk specification therefore provides one additional degree of freedom with one square-root-type variable, four additional degrees of freedom with two, nine additional degrees of freedom with three, etc.

4. Estimation procedure

To determine whether our extended market price of risk specification results in a better fit to U.S. data, we estimate the parameters of nine affine yield models (all Dai and Singleton, 2000, canonical families of affine yield models with three or fewer state variables) using three different market price of risk specifications: the completely affine specification of Dai and Singleton (2000), the essentially affine specification of Duffee (2002), and our extended affine specification. Although our specification always nests the corresponding essentially affine models, and essentially affine models always nest completely affine models, two of the three specifications sometimes coincide. For any $A_0(N)$ affine yield model, our specification and the essentially affine specification coincide, and for any $A_N(N)$ affine yield model, the essentially affine and completely affine models are the same. Therefore, although there are nine different families of models with three market price of risk specifications for each family, there are only 21 distinct combinations to be estimated.

For data, we use zero-coupon yields extracted from U.S. Treasury security prices by the commonly used method of McCulloch (1975) (see, e.g., Duffe, 2002). McCulloch and Kwon (1993) extend the original data set. This method of constructing a yield curve is evaluated by Bliss (1997), who also periodically produces updates to the data set. The data set we use has monthly observations of zero-coupon yields for a 31-year period, from January 1972 until December 2002.[7] This method has a number of desirable features. For example, it smooths over the effect of idiosyncratic prices for a single maturity, and it controls for tax effects that influence the prices of bonds trading at a large discount or premium. The use of a smoothing technique therefore avoids some of the problems that would be likely to occur by other methods of constructing a zero-coupon yield curve; for example, although STRIPS are zero-coupon bonds, they may suffer from liquidity issues, resulting in a lumpy zero-coupon yield curve. Nonetheless, it is possible that this smoothing technique does not fully correct for liquidity issues or tax effects, so our results could differ from those obtained using other data sets, such as swap rates or LIBOR futures prices.

Our estimation procedure is that of maximum likelihood. Apart from its statistical efficiency, use of maximum likelihood estimation makes it straightforward to calculate likelihood ratio statistics to test the significance of our extension. However, maximum likelihood estimation in a multifactor setting with a state vector that is not directly observed presents some challenges that must be overcome.

The state variables of the canonical affine diffusion are not observed directly, but must be extracted from the observed term structure of bond prices or yields. We denote by $y(Y_t, t, T)$ the time t continuously compounded annualized yield of a zero-coupon bond maturing at time T, with the value of the state vector equal to Y_t, that is, $y(Y_t, t, T) = -\ln B(t, T)/(T-t)$. As per Duffie and Kan (1996), for any set of maturities $T_1, \ldots, T_K$, the corresponding yields are affine functions of the state vector:

$$\begin{bmatrix} y(Y_t, t, T_1) \\ \vdots \\ y(Y_t, t, T_K) \end{bmatrix} = \begin{bmatrix} A(T_1 - t) \\ \vdots \\ A(T_K - t) \end{bmatrix} + \begin{bmatrix} B_1(T_1 - t) & \cdots & B_N(T_1 - t) \\ \vdots & \ddots & \vdots \\ B_1(T_K - t) & \cdots & B_N(T_K - t) \end{bmatrix} Y_t, \tag{46}$$

where $A(.)$ and $B_1(.)$ through $B_N(.)$ are deterministic functions that depend on the parameters of the Q-measure dynamics of the state variables and on the parameters of the interest rate process. One is immediately confronted with a dilemma. If we observe fewer bond prices than state variables in the model, it is not possible to determine the exact value of the state vector at any particular time. Estimation then becomes a filtering problem; the likelihood of the next observation depends not only on the currently observed bond prices, but possibly on the entire history. However, if we observe more bond prices than the number of state variables in the model, the observed prices will generally be inconsistent

[7]The data set Robert Bliss produces includes observations from January 1970; however, for the first two years, there is insufficient information to construct zero-coupon yields reliably for the longer maturities used in this study.

with any value of the state vector. The values of the state variables can normally be inferred from an equal number of bond prices, and the remaining bond prices are then predicted exactly, without any error. In practice, no data set ever conforms to a structural model this strictly.

It would seem that the ideal solution would be to use a number of bond prices that is equal to the number of state variables; in this way, for each time-series observation of the set of bond yields, the value of the state vector can be uniquely determined. However, in general, not all of the parameters of the model will be identified. To take a simple example, consider the $A_0(1)$ model, which is equivalent to the model of Vasicek (1977). If one observes only the instantaneous interest rate (which we may consider to be the yield on a zero-maturity zero-coupon bond), the interest rate follows the process

$$\mathrm{d}r_t = \left(-b_{11}^P d_0 + b_{11}^P r_t\right)\mathrm{d}t + d_1\,\mathrm{d}W_t^P(1). \tag{47}$$

The market price of risk parameters (whichever specification we choose) do not affect the observed interest rate process, and are therefore not identified. The situation does not improve if we observe instead a bond with maturity greater than zero; in this case, we may identify d_0 or a single market price of risk parameter, but not both. Similarly, even if the simplest market price of risk restriction is chosen (i.e., the completely affine market price of risk) in an $A_0(N)$ model with $N>1$, a single parameter is always unidentified.

One way to overcome this difficulty is to collect data on more bonds than state variables, but to assume that some of the bond yields are observed with error; see, for example, Pearson and Sun (1994). We take this approach, assuming that for the $A_M\,(N)$ model, N yields are observed without error, but some additional bond yields are observed with a vector of observation errors that is independent, identically distributed, and multivariate Gaussian with mean of zero. Brandt and He (2002) describe an alternate approach, in which all yields are taken to be observed with error.

We also have need of the transition density of the state vector Y_t. This density is needed not only to calculate the estimates themselves, but also to calculate standard errors of the estimates and to perform likelihood ratio tests for the different market price of risk specifications. For four of the nine models we consider (specifically, the $A_0(1)$, $A_0(2)$, $A_0(3)$ and $A_1(1)$ models), the likelihood function is known in closed-form. For the five remaining models (i.e., the $A_1(2)$, $A_2(2)$, $A_1(3)$, $A_2(3)$, and $A_3(3)$ models), the likelihood function is known in closed-form only if additional parameter restrictions are imposed. These restrictions apply under the objective probability measure (i.e., there is no need to calculate likelihoods under the equivalent martingale measure), and can be placed into three categories. First, the β parameters that correspond to the unrestricted state variables in the diffusion matrix must be zero; in other words, the volatility of an unrestricted state variable must be constant. Second, the drift of an unrestricted state variable cannot depend on the values of restricted variables. Finally, the drift of one restricted state variable cannot depend on the value of another restricted state variable. These restrictions are quite strong, and severely restrict the generality of the models. We therefore use the approximate maximum likelihood approach of Aït-Sahalia (2001), as implemented in Aït-Sahalia and Kimmel (2005), for these five models. By using the "reducible" likelihoods developed (and shown to be accurate) in these papers, we are able to relax all restrictions except the first, that the β parameters in the diffusion matrix must be zero. However, these same restrictions are imposed for all market price of risk specifications; since our purpose is to test different specifications with the data, the likelihood ratio tests are still fair comparisons.

Just as parameter restrictions are needed to ensure a closed-form likelihood function, under the P measure, similar restrictions are needed under the Q measure to ensure closed-form bond prices. With the completely affine market price of risk specification, the P-measure restrictions for a closed-form likelihood also ensure closed-form bond prices. However, for the more general market price of risk specifications we consider, this is not necessarily the case, so we can not rely on the existence of closed-form bond prices even if we were to impose restrictions needed for closed-form likelihoods. However, one of the main advantages of affine yield models is that, even when bond prices cannot be found in closed-form, they can be found numerically through very fast algorithms. Bond prices are solutions to the Feynman-Kac partial differential equation; provided a diffusion is affine under the Q measure and the interest rate is an affine function of the state variables, this partial differential equation can be decomposed into a system of ordinary differential equations, which can be solved far more rapidly than a general parabolic partial differential equation of the same dimensionality.[8] We calculate bond prices numerically, even in those cases in which the P-measure dynamics and the market price of risk specification are sufficiently constrained to allow for closed-form bond prices. Since our purpose is to compare different market price of risk specifications, use of the same method to calculate bond prices ensures that any differences that obtain are due to the specification itself, and not the computational method used in the estimation procedure.

As Duffie and Kan (1996) discuss, and as Eq. (46) shows, bond yields in affine yield models are affine functions of the state variables; this is the case for all three market price of risk specifications we consider. Our estimation procedure for an $A_M(N)$ model is then as follows. The parameter vector includes, in addition to the parameters of the $A_M(N)$ model, the standard deviations of observation errors for any extra bond yields, denoted by σ_{N+1} through σ_K (where K is the total number of maturities used in the estimation procedure), and the correlations between observation errors for each pair of the extra bond yields, denoted by ρ_{ij}, with $N+1 \leqslant i, j \leqslant K$, $i \neq j$. For a particular value of the parameter vector we numerically calculate the coefficients $A(T_1-t), \ldots, A(T_N-t)$, $B_1(T_1-t), \ldots, B_N(T_N-t)$ of the relation between bond yields and state variables (Eq. 46) for N maturities, $y(Y_t, t, T_1)$ through $y(Y_t, t, T_N)$. We use rolling maturities throughout, i.e., the value of T_i-t is held fixed, not the value of T_i itself. The bond pricing formula, being affine in Y_t, is easily inverted to find the value of the state variables for each time-series observation of the N bond yields. Holding the model parameters fixed, the state variables are given by

$$Y_t = \begin{bmatrix} B_1(T_1-t) & \cdots & B_N(T_1-t) \\ \vdots & \ddots & \vdots \\ B_1(T_N-t) & \cdots & B_N(T_N-t) \end{bmatrix}^{-1} \begin{bmatrix} y(Y_t, t, T_1) - A(T_1-t) \\ \vdots \\ y(Y_t, t, T_N) - A(T_N-t) \end{bmatrix}. \tag{48}$$

With the time-series values of Y_t (conditional on the current choice of the parameter vector) in hand, we calculate the joint likelihood of the implied time series of observations of the state vector, using the closed-form likelihood expressions. If any of the implied

[8] The numeric tractability of bond pricing depends only on affinity under the measure Q continuing to hold even if the state variable dynamics are not affine under P.

values of the restricted components of Y_t (i.e., the first M elements in the $A_M(N)$ model) are on the wrong side of the boundary, the joint likelihood of the entire time series is set to zero.[9] Using the change of variables formula, we then calculate the joint likelihood of the time series of observations of the N bond yields themselves (note that, for a given value of the parameter vector, the determinant of the Jacobian matrix does not depend on the values of the state variables). The likelihood of the vector of these N yields at some time t, conditional on the last observation, is given by

$$L_y\left(\begin{bmatrix} y(Y_t, t, T_1) \\ \vdots \\ y(Y_t, t, T_N) \end{bmatrix} \middle| \begin{bmatrix} y(Y_{t-\Delta}, t-\Delta, T_1-\Delta) \\ \vdots \\ y(Y_{t-\Delta}, t-\Delta, T_N-\Delta) \end{bmatrix}\right) = \frac{L_Y(Y_t | Y_{t-\Delta})}{\left|\begin{bmatrix} B_1(T_1-t) & \cdots & B_N(T_1-t) \\ \vdots & \ddots & \vdots \\ B_1(T_N-t) & \cdots & B_N(T_N-t) \end{bmatrix}\right|}. \tag{49}$$

where $L_y(\cdot)$ and $L_y(\cdot)$ denote the transition likelihoods for the yield vector and the vector of state variables Y_t, respectively. The joint likelihood is simply the product of the likelihoods for each individual time step. Finally, we calculate the implied observation errors for the additional bond yields $y(Y_t,t,T_{N+1}),\ldots,y(Y_t,t,T_K)$,

$$\begin{bmatrix} \varepsilon(t, T_{N+1}) \\ \vdots \\ \varepsilon(t, T_K) \end{bmatrix} = \begin{bmatrix} y(Y_t, t, T_{N+1}) \\ \vdots \\ y(Y_t, t, T_K) \end{bmatrix} - \left(\begin{bmatrix} A(T_{N+1}-t) \\ \vdots \\ A(T_K-t) \end{bmatrix} - \begin{bmatrix} B_1(T_{N+1}-t) & \cdots & B_N(T_{N+1}-t) \\ \vdots & \ddots & \vdots \\ B_1(T_K-t) & \cdots & B_N(T_K-t) \end{bmatrix} Y_t\right), \tag{50}$$

and multiply the likelihood of the time series of the first N bond yields by the likelihood function for these observation errors (which, as per the previous discussion, are assumed to be Gaussian mean zero and i.i.d.). The result is the joint likelihood of the panel of bond data, including the maturities that are assumed to be observed with error. We repeat this procedure for many values of the parameter vector until the parameter vector that maximizes the value of the likelihood function is discovered. Our search procedure is the Nelder-Mead simplex method.

Many search algorithms perform poorly when there are hard parameter constraints. Particularly troublesome in the estimation of affine yield models is the boundary nonattainment condition for the restricted state variables (which are, of course, our

[9]Use of maximum likelihood ensures that the estimated parameter values are consistent with the observed data. Duffee (2002) points out that not all estimation techniques have this property; the estimated parameter vector for such techniques may imply that the observed time series of bond yields could not have occurred.

primary interest). As Feller (1951) shows, the conditional likelihood of the square-root process (conditional on a past observation) goes to zero near the boundary when the boundary nonattainment condition is satisfied. When the boundary nonattainment inequality is not satisfied, the likelihood either goes to positive infinity near the boundary, or to a finite nonzero value. This strong sensitivity of the likelihood to small changes in model parameters confuses many search algorithms. Consequently, we employ several normalizations to the model parameters to make the likelihood depend on them more smoothly. For example, in the $A_1(1)$ model, we replace a_1^P by

$$c_1^P = \sqrt{a_1^P - 0.5}. \tag{51}$$

Maximum likelihood estimation is invariant to the particular parameterization chosen, so this change of parameters does not affect the estimated model. However, despite this convenient normalization, all parameter estimates, standard errors, etc. are reported in terms of the original model parameters.

One potential problem of this estimation procedure is the assumption that the vector of observation errors of the additional bond yields is i.i.d. Although an assumption of i.i.d. errors may be a misspecification, it is difficult to relax this assumption without causing other estimation problems. For example, rather than being i.i.d., observation errors could be assumed to follow a multivariate Ornstein-Uhlenbeck process. Such a model allows for autocorrelation in the observation error process, but also provides two mechanisms for bond yields to violate no-arbitrage restrictions to a significant degree. If the unconditional mean of the observation error process is nonzero and large in magnitude, then the additional bond yields can differ from their theoretical no-arbitrage values by a large amount, with no statistical penalty in the estimation procedure. Even if the observation error process is constrained to have an unconditional mean of zero, the possibility of very slow (or nonexistent) mean reversion still allows observation errors to remain large for long periods of time, again with no statistical penalty. The assumption of i.i.d. observation errors has the advantage that parameter vectors implying large and/or persistent observation errors are severely penalized in the estimation procedure.

One possible compromise is to employ a two-stage estimation procedure. In the first stage, the observation errors are assumed to be i.i.d. In the second stage, the observation error for each maturity is assumed to follow an Ornstein-Uhlenbeck process, with zero unconditional mean and a speed of mean reversion that is common to all maturities. (Note that the innovations to observation errors for different maturities may be correlated with each other.) However, in the second stage, the unconditional variances of the observation errors are held fixed at the values implied by the first-stage estimates. This procedure allows autocorrelation in the observation error process, but retains the advantage of the one-stage procedure, namely, that parameter vectors implying large and/or persistent observation errors are penalized in the estimation procedure. In the next section, we rely primarily on the one-stage estimation procedure, assuming i.i.d. observation errors, but we also use the two-stage procedure to assess the likely effect of possible model misspecification induced by the i.i.d. assumption.[10]

[10]We would like to thank an anonymous referee for suggesting the two-stage procedure.

5. Results

Tables 1–13 show the estimated parameters of the nine affine yield models we consider, based on the single-stage estimation procedure with an i.i.d. vector of observation errors. As discussed, the extended affine specification is more general than the essentially affine specification of Duffee (2002) in six of the nine models, but for completeness we show all nine. For each $A_M(N)$ model, we use $N+4$ zero-coupon bonds maturing at two-year intervals, except at the very short end of the yield curve, where a one-month maturity is used. For example, for the $A_1(2)$ model, we use zero-coupon bond yields with maturities of one month and of two, four, six, eight, and ten years. For each model, we assume the N shortest maturities are observed without error, and the remaining maturities have observation error; for example, for the $A_M(3)$, the one-month and two- and four-year maturities are considered observed without error, whereas the six-, eight-, 10-, and 12-year maturities have observation error. Each model is estimated with the completely affine, essentially affine, and extended affine market price of risk specifications. Tables 14 and 15 present likelihood ratio tests comparing the different market price of risk specifications.

In seven of the nine models we consider, the likelihood ratio statistics show that the extended affine specification (which contains both the market price of risk parameters introduced by Duffee, 2002, and by us) fits the data better than the completely affine specification at the conventional 95% confidence level, failing to be statistically significant

Table 1
$A_0(1)$ Model estimates

Parameter	Completely affine		Essentially affine		Extended affine	
	Estimate	Std. err.	Estimate	Std. err.	Estimate	Std. err.
b^P_{11}	−0.0445	0.0042	−0.4025	0.1696	−0.4025	0.1696
a^Q_1	0.1832	0.2248	0.1626	0.0346	0.1626	0.0346
b^Q_{11}	−0.0445	0.0042	−0.0444	0.0042	−0.0444	0.0042
d_0	0.0492	0.1291	0.0613	0.0204	0.0613	0.0204
d_1	0.0255	0.0006	0.0257	0.0006	0.0257	0.0006
σ_2	0.0113	0.0006	0.0113	0.0006	0.0113	0.0006
σ_3	0.0134	0.0008	0.0134	0.0008	0.0134	0.0008
σ_4	0.0144	0.0009	0.0144	0.0009	0.0144	0.0009
σ_5	0.0148	0.0009	0.0148	0.0009	0.0148	0.0009
ρ_{32}	0.9702	0.0043	0.9699	0.0044	0.9699	0.0044
ρ_{42}	0.9450	0.0083	0.9445	0.0084	0.9445	0.0084
ρ_{43}	0.9941	0.0009	0.9940	0.0010	0.9940	0.0010
ρ_{52}	0.9283	0.0108	0.9277	0.0110	0.9277	0.0110
ρ_{53}	0.9856	0.0022	0.9855	0.0023	0.9855	0.0023
ρ_{54}	0.9975	0.0004	0.9975	0.0004	0.9975	0.0004

This table shows the parameter estimates and standard errors for the $A_0(1)$ model parameters, for the completely affine, essentially affine, and extended affine market price of risk specifications. A zero-coupon bond yield with maturity of one month is assumed to be observed without error; zero-coupon bond yields with maturities of two, four, six, and eight years are assumed to be observed with error. The essentially affine and extended affine specifications coincide for this model. Note that, for the completely affine market price of risk specification, the b^P_{11} and b^Q_{11} parameters must coincide. For the other two market price of risk specifications, all parameters can vary independently.

Table 2
$A_1(1)$ Model estimates

Parameter	Completely affine		Essentially affine		Extended affine	
	Estimate	Std. err.	Estimate	Std. err.	Estimate	Std. err.
a_1^P	0.5000	0.0495	0.5000	0.0495	0.5000	0.9196
b_{11}^P	−0.0828	0.0670	−0.0828	0.0670	−0.0828	0.1172
a_1^Q	0.5000	0.0495	0.5000	0.0495	0.5000	0.0583
b_{11}^Q	−0.0137	0.0660	−0.0137	0.0660	−0.0137	0.0056
d_0	0.0110	0.0012	0.0110	0.0012	0.0110	0.0022
d_1	0.0074	0.0004	0.0074	0.0004	0.0074	0.0004
σ_2	0.0119	0.0006	0.0119	0.0006	0.0119	0.0006
σ_3	0.0144	0.0009	0.0144	0.0009	0.0144	0.0009
σ_4	0.0155	0.0010	0.0155	0.0010	0.0155	0.0010
σ_5	0.0159	0.0010	0.0159	0.0010	0.0159	0.0010
ρ_{32}	0.9727	0.0042	0.9727	0.0042	0.9727	0.0042
ρ_{42}	0.9511	0.0078	0.9511	0.0078	0.9511	0.0079
ρ_{43}	0.9950	0.0008	0.9950	0.0008	0.9950	0.0008
ρ_{52}	0.9371	0.0101	0.9371	0.0101	0.9371	0.0102
ρ_{53}	0.9877	0.0020	0.9877	0.0020	0.9877	0.0020
ρ_{54}	0.9978	0.0003	0.9978	0.0003	0.9978	0.0003

This table shows the parameter estimates and standard errors for the $A_1(1)$ model parameters, for the completely affine, essentially affine, and extended affine market price of risk specifications. A zero-coupon bond yield with maturity of one month is assumed to be observed without error; zero-coupon bond yields with maturities of two, four, six, and eight years are assumed to be observed with error. The completely affine and essentially affine specifications coincide for this model. Note that, for the completely affine and essentially affine market price of risk specifications, the a_1^P and a_1^Q parameters must coincide. For the extended affine market price of risk specification, all parameters can vary independently.

only for the $A_1(1)$ model (which is a slight generalization of the model of Cox, et al., 1985) and the $A_1(2)$ model. In several models, including all three-factor models, the likelihood ratio statistic is far above the 95% cutoff level, indicating strong rejection of the hypothesis that the extended market price of risk specification (relative to the completely affine market price of risk) is not needed. Considering only the essentially affine models of Duffee (2002), we note that five of the six models for which the essentially affine specification is not degenerate have likelihood ratio statistics (relative to the completely affine case) above, and often far above, the 95% cutoff value, excepting only the $A_1(2)$ model. This finding confirms the improved fit that Duffee (2002) finds for this specification, in a data set that is only partially overlapping with his. Turning to the extended affine specification we introduce, there are six models for which this specification is more general than the essentially affine specification. In four of these models, the hypothesis that the extended affine market price of risk is not necessary, relative to the essentially affine market price of risk, is rejected at the 90% level; in three of these four models, the rejection still holds at the 95% level. The rejection is particularly strong for those models with multiple restricted state variables, that is, the $A_2(2)$, $A_2(3)$, and $A_3(3)$ models. The extended affine specification fails to be statistically significant for the $A_1(3)$ model at the 95% level, but the likelihood ratio statistic is above the 90% cutoff value. The extended affine specification would appear to add virtually no improved fit for the $A_1(2)$ model over that of the essentially

Table 3
$A_0(2)$ Model estimates

Parameter	Completely affine		Essentially affine		Extended affine	
	Estimate	Std. err.	Estimate	Std. err.	Estimate	Std. err.
b_{11}^P	−0.0400	0.0037	−0.1570	0.1143	−0.1570	0.1143
b_{21}^P	1.0657	0.0877	−0.3279	0.3382	−0.3279	0.3382
b_{22}^P	−2.0970	0.1581	−2.2883	0.2694	−2.2883	0.2694
a_1^Q	0.2023	0.1885	0.6616	0.1797	0.6616	0.1797
a_2^Q	0.9942	0.1972	0.6880	0.2778	0.6880	0.2778
b_{11}^Q	−0.0400	0.0037	−0.1607	0.2219	−0.1607	0.2219
b_{12}^Q	0.0000	0.0000	−1.3452	0.2799	−1.3452	0.2799
b_{21}^Q	1.0657	0.0877	−0.1741	0.2656	−0.1741	0.2656
b_{22}^Q	−2.0970	0.1581	−1.9840	0.3114	−1.9840	0.3114
d_0	0.0306	0.0947	0.0562	0.0333	0.0562	0.0333
d_1	0.0067	0.0011	0.0194	0.0026	0.0194	0.0026
d_2	0.0261	0.0005	0.0178	0.0030	0.0178	0.0030
σ_3	0.0033	0.0001	0.0033	0.0002	0.0033	0.0002
σ_4	0.0049	0.0002	0.0049	0.0003	0.0049	0.0003
σ_5	0.0057	0.0003	0.0057	0.0003	0.0057	0.0003
σ_6	0.0063	0.0003	0.0063	0.0003	0.0063	0.0003
ρ_{43}	0.9742	0.0034	0.9742	0.0037	0.9742	0.0037
ρ_{53}	0.9448	0.0074	0.9447	0.0080	0.9447	0.0080
ρ_{54}	0.9891	0.0014	0.9891	0.0015	0.9891	0.0015
ρ_{63}	0.9266	0.0091	0.9265	0.0098	0.9265	0.0098
ρ_{64}	0.9720	0.0034	0.9720	0.0035	0.9720	0.0035
ρ_{65}	0.9917	0.0010	0.9917	0.0010	0.9917	0.0010

This table shows the parameter estimates and standard errors for the $A_0(2)$ model parameters, for the completely affine, essentially affine, and extended affine market price of risk specifications. Zero-coupon bond yields with maturities of one month and two years are assumed to be observed without error; zero-coupon bond yields with maturities of four, six, eight, and ten years are assumed to be observed with error. The essentially affine and extended affine specifications coincide for this model. Note that, for the completely affine market price of risk specification, the slope coefficient parameters in the drift must coincide (i.e., b_{11}^P and b_{11}^Q are the same, b_{21}^P and b_{21}^Q are the same, and b_{22}^P and b_{22}^Q are the same). Furthermore, for the completely affine market price of risk specification, the b_{12}^Q parameter is held fixed at zero. For the other two market price of risk specifications, all parameters can vary independently.

affine specification. However, for the $A_1(1)$ model, the likelihood ratio statistic is well above the needed cutoff value. In the $A_2(2)$, $A_2(3)$ and $A_3(3)$ models, the likelihood ratio statistic is much higher than the 95% cutoff level, indicating a very strong significance for the extended specification relative to either the completely or essentially affine specification (note that these latter two specifications coincide for the $A_2(2)$ and $A_3(3)$ models).

Looking at the parameter estimates, we note a few points. First, the canonical state variables in the representation of Dai and Singleton (2000) do not necessarily have simple, intuitive interpretations, as they are not linked directly to observable characteristics of the term structure, but only indirectly through the bond pricing formulae. Furthermore, in the more richly parameterized models (such as the two- and three-factor models), many individual parameters often fail to be statistically significant; the root cause appears to be

Table 4
$A_1(2)$ Model estimates

Parameter	Completely affine		Essentially affine		Extended affine	
	Estimate	Std. err.	Estimate	Std. err.	Estimate	Std. err.
a_1^P	0.9553	0.1656	0.9974	0.1788	1.7247	2.0680
b_{11}^P	0.0000	0.0400	−0.0633	0.0415	−0.0981	0.0965
b_{21}^P	0.2193	0.0308	0.1471	0.0338	0.1468	0.0344
b_{22}^P	−2.1609	0.1627	−1.8294	0.2523	−1.8289	0.2524
a_1^Q	0.9553	0.1656	0.9974	0.1788	1.0035	0.1838
a_2^Q	1.0309	0.1970	−0.0183	0.0417	0.2300	0.9169
b_{11}^Q	−0.0173	0.0399	0.2349	0.8867	−0.0183	0.0044
b_{21}^Q	0.2193	0.0308	0.2209	0.0340	0.2202	0.0340
b_{22}^Q	−2.1609	0.1627	−2.2361	0.2069	−2.2340	0.2085
D_0	−0.0200	0.0079	−0.0114	0.0122	−0.0117	0.0123
d_1	0.0015	0.0002	0.0015	0.0002	0.0015	0.0002
d_2	0.0261	0.0005	0.0256	0.0005	0.0256	0.0005
σ_3	0.0034	0.0001	0.0034	0.0002	0.0034	0.0002
σ_4	0.0051	0.0002	0.0051	0.0002	0.0051	0.0003
σ_5	0.0059	0.0003	0.0059	0.0003	0.0059	0.0003
σ_6	0.0065	0.0003	0.0065	0.0003	0.0065	0.0003
ρ_{43}	0.9775	0.0029	0.9775	0.0029	0.9775	0.0030
ρ_{53}	0.9509	0.0064	0.9508	0.0065	0.9508	0.0066
ρ_{54}	0.9900	0.0013	0.9899	0.0013	0.9899	0.0014
ρ_{63}	0.9327	0.0083	0.9326	0.0084	0.9325	0.0085
ρ_{64}	0.9736	0.0032	0.9736	0.0032	0.9736	0.0033
ρ_{65}	0.9923	0.0009	0.9923	0.0009	0.9923	0.0009

This table shows the parameter estimates and standard errors for the $A_1(2)$ model parameters, for the completely affine, essentially affine, and extended affine market price of risk specifications. Zero-coupon bond yields with maturities of one month and two years are assumed to be observed without error; zero-coupon bond yields with maturities of four, six, eight, and ten years are assumed to be observed with error. Note that, for the completely affine and essentially affine market price of risk specifications, the a_1^P and a_1^Q parameters must coincide. For the completely affine market price of risk specification, the b_{21}^Q and b_{22}^Q parameters must be the same as their P measure counterparts b_{21}^P and b_{22}^P For the extended affine market price of risk specification, all parameters can vary independently.

that many of the parameter estimates are correlated with each other. The likelihood ratio statistics tell us whether the parameters introduced by the extended affine specification are collectively statistically significant, but it is sometimes difficult to assess the influence of individual parameters. Nonetheless, from the parameter estimates in Tables 1–13, at least one consistent theme emerges, which is that, when the more restricted market price of risk specifications are used, cross-sectional fitting of the term structure is matched at the expense of matching time-series behavior. When the market price of risk specification is very constrained, the model parameters must perform two functions simultaneously. Specifically, they must generate a term structure with the approximate shape of the observed term structure, and they must also generate time-series behavior that is consistent with what is observed in the data. It may be difficult for both modelling tasks to be performed well simultaneously. This tension is reduced when the market price of risk is

Table 5
$A_2(2)$ Model estimates

Parameter	Completely affine		Essentially affine		Extended affine	
	Estimate	Std. err.	Estimate	Std. err.	Estimate	Std. err.
a_1^P	0.5702	0.3730	0.5702	0.3730	0.8474	1.3866
a_2^P	0.5000	1.5454	0.5000	1.5454	0.5000	1.6729
b_{11}^P	−0.3012	0.1491	−0.3012	0.1491	−0.7925	0.2674
b_{12}^P	0.3975	0.1339	0.3975	0.1339	1.1530	0.4114
b_{21}^P	1.1904	0.3253	1.1904	0.3253	0.7102	0.2564
b_{22}^P	−2.0768	0.2334	−2.0768	0.2334	−1.2642	0.2641
a_1^Q	0.5702	0.3730	0.5702	0.3730	0.5000	0.2695
a_2^Q	0.5000	1.5454	0.5000	1.5454	0.5000	1.2716
b_{11}^Q	−0.3024	0.0608	−0.3024	0.0608	−0.2910	0.1913
b_{12}^Q	0.3975	0.1339	0.3975	0.1339	0.3764	0.1562
b_{21}^Q	1.1904	0.3253	1.1904	0.3253	1.3187	0.3449
b_{22}^Q	−1.6632	0.0828	−1.6632	0.0828	−1.7966	0.2420
d_0	0.0000	0.0076	0.0000	0.0076	0.0045	0.0058
d_1	−0.0001	0.0006	−0.0001	0.0006	0.0000	0.0006
d_2	0.0087	0.0011	0.0087	0.0011	0.0088	0.0011
σ_3	0.0034	0.0001	0.0034	0.0001	0.0035	0.0002
σ_4	0.0050	0.0002	0.0050	0.0002	0.0051	0.0003
σ_5	0.0059	0.0003	0.0059	0.0003	0.0060	0.0003
σ_6	0.0064	0.0003	0.0064	0.0003	0.0065	0.0003
ρ_{43}	0.9771	0.0029	0.9771	0.0029	0.9779	0.0030
ρ_{53}	0.9500	0.0064	0.9500	0.0064	0.9516	0.0067
ρ_{54}	0.9898	0.0013	0.9898	0.0013	0.9901	0.0014
ρ_{63}	0.9313	0.0084	0.9313	0.0084	0.9332	0.0087
ρ_{64}	0.9732	0.0032	0.9732	0.0032	0.9738	0.0033
ρ_{65}	0.9922	0.0009	0.9922	0.0009	0.9923	0.0009

This table shows the parameter estimates and standard errors for the $A_2(2)$ model parameters, for the completely affine, essentially affine, and extended affine market price of risk specifications. Zero-coupon bond yields with maturities of one month and two years are assumed to be observed without error; zero-coupon bond yields with maturities of four, six, eight, and ten years are assumed to be observed with error. The completely affine and essentially affine specifications coincide for this model. Note that, for the completely affine and essentially affine market price of risk specifications, the a_1^P, a_2^P, b_{12}^P and b_{21}^P parameters must all be equal to their Q-measure counterparts, a_1^Q, a_2^Q, b_{12}^Q and b_{21}^Q. For the extended affine market price of risk specification, all parameters can vary independently.

extended to allow for different P and Q parameters, since the cross-sectional shape of the term structure is determined by the Q-measure parameters, whereas the time-series behavior of the yield curve is governed by P-measure parameters. When P- and Q-measure parameters are constrained to be the same, the need to fit the cross-sectional shape of the yield curve dominates at the expense of time series. Note, for example, the estimates for the $A_2(2)$, $A_2(3)$, and $A_3(3)$ models in Tables 5, 10, 11, 12 and 13. With the completely affine and essentially affine market price of risk specifications, the a_i^P and a_i^Q parameters are constrained to be equal to each other for each value of $1 \leqslant i \leqslant M$. However, with the extended affine market price of risk specification, these parameters can differ, and we find

Table 6
$A_0(3)$ Model estimates (Part I)

Parameter	Completely affine		Essentially affine		Extended affine	
	Estimate	Std. err.	Estimate	Std. error.	Estimate	Std. error.
b_{11}^P	−0.0309	0.0028	−0.1265	0.1173	−0.1265	0.1173
b_{21}^P	0.1258	0.0395	−0.3703	0.1943	−0.3703	0.1943
b_{22}^P	−0.8239	0.0565	−0.8302	0.2288	−0.8302	0.2288
b_{31}^P	0.6733	0.1798	−0.2363	0.2961	−0.2363	0.2961
b_{32}^P	2.5639	0.1610	0.9107	0.3613	0.9107	0.3613
b_{33}^P	−2.9985	0.3177	−3.7907	0.2095	−3.7907	0.2095
a_1^Q	0.1957	0.1854	1.4349	1.8761	1.4349	1.8761
a_2^Q	0.1525	0.2078	−0.2360	0.8625	−0.2360	0.8625
a_3^Q	1.1574	0.2382	2.3454	4.3456	2.3454	4.3456
b_{11}^Q	−0.0309	0.0028	0.5378	0.5910	0.5378	0.5910
b_{12}^Q	0.0000	0.0000	2.0788	1.3848	2.0788	1.3848
b_{13}^Q	0.0000	0.0000	−6.1613	3.2583	−6.1613	3.2583
b_{21}^Q	0.1258	0.0395	−0.4903	0.4593	−0.4903	0.4593
b_{22}^Q	−0.8239	0.0565	−1.2794	1.0198	−1.2794	1.0198
b_{23}^Q	0.0000	0.0000	1.8346	2.4336	1.8346	2.4336
b_{31}^Q	0.6733	0.1798	1.6907	1.5325	1.6907	1.5325
b_{32}^Q	2.5639	0.1610	5.3270	2.9392	5.3270	2.9392
b_{33}^Q	−2.9985	0.3177	−12.9222	6.4730	−12.9222	6.4730

This table shows the parameter estimates and standard errors for the $A_0(3)$ model parameters, for the completely affine, essentially affine, and extended affine market price of risk specifications. Zero-coupon bond yields with maturities of one month and two and four years are assumed to be observed without error; zero-coupon bond yields with maturities of six, eight, ten, and twelve years are assumed to be observed with error. The essentially affine and extended affine specifications coincide for this model. Note that, for the completely affine market price of risk specification, the slope coefficient parameters in the drift must coincide (i.e., b_{11}^P and b_{11}^Q are the same, b_{21}^P and b_{21}^Q are the same, b_{22}^P and b_{22}^Q are the same, b_{31}^P and b_{31}^Q are the same, b_{32}^P and b_{32}^Q are the same, and b_{33}^P and b_{33}^Q are the same). Furthermore, for the completely affine market price of risk specification, the b_{12}^Q, b_{13}^Q and b_{23}^Q parameters are held fixed at zero. For the other two market price of risk specifications, all parameters can vary independently. This table is continued in Table 7.

that the values of a_i^P change considerably more than the values of a_i^Q. When the P-measure and Q-measure parameters are constrained to be the same, the estimates are close to the Q-measure estimates for the extended affine specification. The need to match the cross-sectional shape of the yield curve generally dominates the need to match its time-series behavior when the model parameterization makes it difficult to match both.

Because the parameter estimates are for the canonical form of affine models developed by Dai and Singleton (2000), and because the relation between the parameters of these canonical-form models and observable characteristics of the yield curve is somewhat indirect, we examine the properties of the model parameters in several other ways that allow for a very natural interpretation. First, we examine the extent to which different models match the cross-sectional shape of the yield curve. Tables 16 and 17 show the means and standard deviations, respectively, of the observation errors associated with the

Table 7
$A_0(3)$ Model estimates (Part II)

Parameter	Completely affine		Essentially affine		Extended affine	
	Estimate	Std. err.	Estimate	Std. err.	Estimate	Std. err.
d_0	0.0302	0.0905	0.0529	0.0364	0.0529	0.0364
d_1	0.0045	0.0013	0.0179	0.0030	0.0179	0.0030
d_2	0.0056	0.0012	0.0005	0.0081	0.0005	0.0081
d_3	0.0268	0.0006	0.0279	0.0127	0.0279	0.0127
σ_4	0.0010	0.0000	0.0010	0.0000	0.0010	0.0000
σ_5	0.0018	0.0001	0.0018	0.0001	0.0018	0.0001
σ_6	0.0023	0.0001	0.0023	0.0001	0.0023	0.0001
σ_7	0.0027	0.0001	0.0027	0.0001	0.0027	0.0001
ρ_{54}	0.9153	0.0094	0.9155	0.0094	0.9155	0.0094
ρ_{64}	0.7734	0.0215	0.7775	0.0210	0.7775	0.0210
ρ_{65}	0.9327	0.0063	0.9338	0.0064	0.9338	0.0064
ρ_{74}	0.6847	0.0272	0.6937	0.0271	0.6937	0.0271
ρ_{75}	0.8203	0.0170	0.8240	0.0177	0.8240	0.0177
ρ_{76}	0.9526	0.0046	0.9533	0.0048	0.9533	0.0048

This table is a continuation of Table 6.

“extra” yields used in the estimation procedure, both in units of basis points. In a model that has perfect cross-sectional fit, both the means and standard deviations would be zero; a lower (absolute) value for each statistic indicates that the extra yields are very close to the values predicted by the model. The most obvious point we can take from Tables 16 and 17 is that two-factor models provide better cross-sectional fit than do one-factor models, and three-factor models are better still. But holding the number of factors fixed, neither the choice of model (e.g., $A_1(3)$ versus $A_2(3)$) nor the choice of market price of risk specification appears to have much of an effect, as the values for both means and standard deviations are similar across models and across market price of risk specifications. Paradoxically, the introduction of a more flexible market price of risk specification sometimes leads to a slightly weaker cross-sectional fit. We argue that this phenomenon occurs because the parameter estimates of the completely affine specification already reflect nearly the best possible cross-sectional fit, sacrificing the fit of the time-series properties in order to accomplish this. Therefore, when more flexible market price of risk specifications are introduced, there is little room for improvement cross-sectionally, and thus the additional parameters improve instead the time-series behavior of the model (we consider this point below in detail). This finding is consistent with what is often found in the parameter estimate tables; when P-measure and Q-measure parameters are constrained to be equal, the Q measure (which determines pricing and crosssectional fit) usually dominates. Consequently, there is little or no cross-sectional improvement (or even very slight degradation) when the P-measure and Q-measure parameters are different, because the new Q-measure estimates are similar to the more constrained estimates.

Table 18 shows first-order autocorrelations for the observation errors. Although these errors are assumed to be i.i.d., the table shows that in fact they are autocorrelated, particularly for the one- and two-factor models. No particularly striking differences between models (with the same number of factors) or across market price of risk

Table 8
$A_1(3)$ Model estimates (Part I)

Parameter	Completely affine		Essentially affine		Extended affine	
	Estimate	Std. err.	Estimate	Std. err.	Estimate	Std. err.
a_1^P	1.2077	0.3487	1.2584	0.3615	4.0136	3.8963
b_{11}^P	0.0000	0.0334	0.0000	0.0337	−0.1261	0.1012
b_{21}^P	0.0168	0.0081	0.0254	0.0238	0.0242	0.0228
b_{22}^P	−0.8597	0.0585	−0.6888	0.2386	−0.6889	0.2404
b_{31}^P	0.1241	0.0389	0.0937	0.0272	0.0894	0.0268
b_{32}^P	2.5485	0.1581	0.7223	0.3357	0.7280	0.3418
b_{33}^P	−2.8698	0.3132	−3.4711	0.1297	−3.4696	0.1332
a_1^Q	1.2077	0.3487	1.2584	0.3615	1.3691	0.4229
a_2^Q	0.1459	0.2082	−1.4947	8.0120	−1.4927	8.4542
a_3^Q	1.1254	0.2382	−11.9287	58.4569	−11.8794	62.3288
b_{11}^Q	−0.0150	0.0332	0.0166	0.0335	−0.0175	0.0041
b_{21}^Q	0.0168	0.0081	0.1958	0.7712	0.1769	0.7303
b_{22}^Q	−0.8597	0.0585	1.2287	9.7083	1.1037	9.6855
b_{23}^Q	0.0000	0.0000	−4.3956	20.9166	−4.1078	20.7864
b_{31}^Q	0.1241	0.0389	1.1679	5.6125	1.0467	5.3634
b_{32}^Q	2.5485	0.1581	14.2491	70.8946	13.4227	71.3138
b_{33}^Q	−2.8698	0.3132	−31.4520	151.6836	−29.5496	152.1154

This table shows the parameter estimates and standard errors for the $A_1(3)$ model parameters, for the completely affine, essentially affine, and extended affine market price of risk specifications. Zero-coupon bond yields with maturities of one month and two and four years are assumed to be observed without error; zero-coupon bond yields with maturities of six, eight, ten, and twelve years are assumed to be observed with error. The b_{23}^P is held at zero to ensure model identification. Note that, for the completely affine and essentially affine market price of risk specifications, the a_1^P parameter must coincide with its Q-measure counterpart a_1^Q. Furthermore, for the completely affine market price of risk specification, the b_{21}^P, b_{22}^P, b_{23}^P, b_{31}^P, b_{31}^P, and b_{32}^P parameters must all be equal to their Q-measure counterparts. For the extended affine market price of risk specification, all parameters can vary independently. This table is continued in Table 9.

specifications are evident; the only clear trend is that autocorrelations are smaller in three-factor models. Although this table does show some evidence of misspecification, the simple model we use does have the advantage of penalizing observation errors that are large. As we discuss in Section 4, a more complicated model (for example, a multivariate Ornstein-Uhlenbeck process) may also allow observation errors to have a large nonzero mean (if the mean parameter is not constrained to be small) or very slow (or nonexistent) mean reversion, with the result that observation errors could be very large through the entire data sample with no statistical penalty. Table 15 reports likelihood ratios from the two-stage estimation procedure, with the restricted model of observation error evolution as described in Section 4. The results are broadly similar to those that obtain under the one-stage procedure: for the one- and two-factor models, the significance of both the essentially affine and extended affine market price of risk (relative to the completely affine price of risk) tends to be similar to or higher than the significance based on the one-stage procedure. For most three-factor models, there is a decrease in statistical significance of the

Table 9
$A_1(3)$ Model estimates (Part II)

Parameter	Completely affine		Essentially affine		Extended affine	
	Estimate	Std. err.	Estimate	Std. err.	Estimate	Std. err.
d_0	−0.0187	0.0144	0.0098	0.1028	0.0068	0.1092
d_1	0.0008	0.0002	−0.0005	0.0096	−0.0004	0.0092
d_2	0.0051	0.0013	0.0015	0.1222	0.0026	0.1235
d_3	0.0268	0.0006	0.0527	0.2619	0.0501	0.2638
σ_4	0.0010	0.0000	0.0010	0.0000	0.0010	0.0000
σ_5	0.0018	0.0001	0.0017	0.0001	0.0018	0.0001
σ_6	0.0023	0.0001	0.0023	0.0001	0.0023	0.0001
σ_7	0.0027	0.0001	0.0027	0.0001	0.0027	0.0001
ρ_{54}	0.9119	0.0096	0.9130	0.0096	0.9133	0.0097
ρ_{64}	0.7687	0.0214	0.7755	0.0208	0.7760	0.0209
ρ_{65}	0.9323	0.0062	0.9342	0.0061	0.9343	0.0061
ρ_{74}	0.6847	0.0269	0.6975	0.0255	0.6981	0.0255
ρ_{75}	0.8223	0.0167	0.8276	0.0165	0.8278	0.0165
ρ_{76}	0.9531	0.0044	0.9542	0.0044	0.9543	0.0044

This table is a continuation of Table 8.

essentially affine market price of risk relative to the completely affine specification, although all of the statistics remain far above the 95% cutoff values. However, the significance of the extended affine specification for the three-factor models, relative to the essentially affine specification, is similar for the $A_3(3)$ model, and higher for the other models, when the two-stage procedure is used; in particular, the extended specification for the $A_1(3)$ model is now significant at the 95% level. The parameter estimates and implied time-series behavior of the yields under the two-stage procedure (not shown) are similar to those we derive from the one-stage procedure. Consequently, although there is evidence that the i.i.d. assumption results in misspecification, attempts to introduce time-dependency in the model for observation errors do not have a large effect on the estimated parameters, and also do not tend to reduce the statistical significance of the extended affine market price of risk specification.

If the improved fit indicated by the likelihood ratio statistics does not take the form of improved crosssectional fit, then the only remaining scope for improvement is in the time-series behavior. We therefore construct two types of measures of the time-series behavior of the different models, which are the first and second unconditional moments of yield forecast errors. These two measures summarize only some of the information in the time-series behavior of the model, for two reasons. First, these two moments capture only some of the information in the unconditional distribution of yield forecast errors; skewness and kurtosis, for example, are ignored. Second, there may be considerable state-dependency in the distribution of yield forecast errors; for example, a model may capture this distribution much better when interest rates are high than when they are low. The unconditional distribution of yield forecast errors does not reflect this variation. Nonetheless, examination of the first two unconditional moments can provide a good indication as to how the improved fit shown by likelihood ratio statistics appears as characteristics of directly observed quantities.

Table 10
$A_2(3)$ Model estimates (Part I)

Parameter	Completely affine		Essentially affine		Extended affine	
	Estimate	Std. err.	Estimate	Std. err.	Estimate	Std. err.
a_1^P	1.2439	0.3906	1.2770	0.6325	0.5000	3.3104
a_2^P	2.6003	0.5196	2.3657	0.6730	0.5000	1.7780
b_{11}^P	0.0000	0.0071	−0.0000	0.0250	−0.0000	0.1293
b_{12}^P	0.0000	0.0006	0.0000	0.1405	0.0000	0.3546
b_{21}^P	0.0706	0.0238	0.0990	0.0400	0.0520	0.0778
b_{22}^P	−0.8474	0.0991	−0.8748	0.0937	−0.3409	0.1897
b_{31}^P	0.1060	0.0384	0.0822	0.0247	0.0747	0.0249
b_{32}^P	1.0133	0.1064	0.3873	0.0644	0.3751	0.0659
b_{33}^P	−3.3299	0.3173	−3.3293	0.0919	−3.3292	0.0911
a_1^Q	1.2439	0.3906	1.2770	0.6325	1.0382	0.6070
a_2^Q	2.6003	0.5196	2.3657	0.6730	1.9323	0.8063
a_3^Q	1.2211	0.2400	−49.7533	176.5	−55.8606	1374.2
b_{11}^Q	−0.0160	0.0086	−0.0172	0.0276	−0.0356	0.0271
b_{12}^Q	0.0000	0.0006	0.0000	0.1405	0.1086	0.1395
b_{21}^Q	0.0706	0.0238	0.0990	0.0400	0.1246	0.0451
b_{22}^Q	−0.7691	0.0828	−0.7526	0.0804	−0.7732	0.0466
b_{31}^Q	0.1060	0.0384	2.6206	9.3513	2.6527	64.3740
b_{32}^Q	1.0133	0.1064	16.6021	60.4348	20.2528	501.2
b_{33}^Q	−3.3299	0.3173	−78.3950	279.2	−95.1782	2332.3

This table shows the parameter estimates and standard errors for the $A_2(3)$ model parameters, for the completely affine, essentially affine, and extended affine market price of risk specifications. Zero coupon bond yields with maturities of one month and two and four years are assumed to be observed without error; zero-coupon bond yields with maturities of six, eight, ten, and twelve years are assumed to be observed with error. Note that, for the completely affine and essentially affine market price of risk specifications, the a_1^P and a_1^Q parameters must coincide, as must the a_2^P and a_2^Q parameters; furthermore, the b_{12}^Q and b_{21}^Q parameters must be equal to their counterparts under the P measure. For the completely affine market price of risk specification, the b_{31}^Q, b_{32}^Q, and b_{33}^Q parameters must be equal to their counterparts under the P measure, b_{31}^P, b_{32}^P, and b_{33}^P. For the extended affine market price of risk specification, all parameters can vary independently. This table is continued in Table 11.

Table 19 shows the bias of yield forecast errors. Specifically, it shows the unconditional mean of the difference (in basis points) between the observed yield change and that predicted by the model, for all models and all market price of risk specifications, and for all maturities used in estimation (including the "extra" yields). Positive values indicate that the model tends to underestimate future yields, and negative values indicate over-estimation. Table 20 shows the difference between the second moment (in percentage points squared; note that these units are different than those used in Table 19) of observed yield changes and the predicted second moment of yield change, averaged across all observations. In both tables, all observations receive equal weight, i.e., there is no attempt to overweight observations occurring at times of low variance and underweight those at times of high variance. Positive values indicate that the model tends to underestimate the second moment of yield changes, and negative values indicate overestimation.

Table 11
$A_2(3)$ Model estimates (Part II)

Parameter	Completely affine		Essentially affine		Extended affine	
	Estimate	Std. err.	Estimate	Std. err.	Estimate	Std. err.
d_0	−0.0506	0.0101	0.0454	0.2902	0.0612	2.2307
d_1	0.0007	0.0002	−0.0030	0.0151	−0.0032	0.1042
d_2	0.0024	0.0004	−0.0173	0.0976	−0.0234	0.8113
d_3	0.0272	0.0007	0.1271	0.4526	0.1541	3.7770
σ_4	0.0010	0.0000	0.0010	0.0000	0.0010	0.0000
σ_5	0.0018	0.0001	0.0018	0.0001	0.0018	0.0001
σ_6	0.0023	0.0001	0.0023	0.0001	0.0023	0.0001
σ_7	0.0027	0.0001	0.0027	0.0001	0.0027	0.0001
ρ_{54}	0.9128	0.0092	0.9133	0.0093	0.9133	0.0095
ρ_{64}	0.7710	0.0206	0.7760	0.0203	0.7760	0.0207
ρ_{65}	0.9329	0.0061	0.9343	0.0061	0.9343	0.0061
ρ_{74}	0.6861	0.0265	0.6962	0.0256	0.6975	0.0256
ρ_{75}	0.8221	0.0169	0.8264	0.0169	0.8274	0.0166
ρ_{76}	0.9526	0.0045	0.9535	0.0045	0.9541	0.0044

This table is a continuation of Table 10.

We focus on the six models for which the extended affine specification is more general than the essentially affine specification, i.e., all models except the $A_0(1)$, $A_0(2)$, and $A_0(3)$ models. It is difficult to make general statements that hold across all models and market price of risk specifications. Beginning with the $A_1(1)$ model, we note that there is little improvement in the yield forecast errors under the extended affine specification. Only the four-year yield has a smaller forecast error, and some of the other maturities actually have larger forecast errors under the extended specification. The estimates of the $A_1(1)$ model are almost the same under all three market prices of risk. (Note that they are constrained to be the same in the completely and essentially affine cases.) The extended price of risk for the $A_1(2)$ model hardly does any better, with very low statistical significance relative to the essentially affine case and little improvement (and sometimes even some degradation) in performance in the biases and standard deviations of the observation errors or yield forecast errors relative to the essentially affine specification. However, the extended affine specification does much better in the remaining three models for which this specification is not degenerate.

The extended affine specification for the $A_2(2)$ model has a likelihood ratio statistic that is substantially larger than the 95% cutoff value, indicating strong significance. (Note that the completely affine and essentially affine specification coincide for this model.) The improved fit of this model manifests itself mainly in smaller biases of yield forecast errors, which are smaller in magnitude at every maturity. Observation error bias is also reduced for every maturity, although to a lesser extent, and the second moments of yield forecast errors are also more accurate at each maturity. Only the standard deviations of the observation errors are larger than in the completely and essentially affine cases (which coincide for this model), and this increase is small.

Relative to the essentially affine specification, the extended specification is statistically significant for the $A_1(3)$ model at the 90% level. The statistics on the means and standard

Table 12
$A_3(3)$ Model estimates (Part I)

Parameter	Completely affine		Essentially affine		Extended affine	
	Estimate	Std. err.	Estimate	Std. err.	Estimate	Std. err.
a_1^P	0.7739	0.8281	0.7739	0.8281	0.5000	189.1
a_2^P	30614	63165	30614	63165	11544	29457
a_3^P	2.9081	8.4479	2.9081	8.4479	0.5000	63.3889
b_{11}^P	−0.0000	0.0225	−0.0000	0.0225	0.0182	0.1644
b_{12}^P	0.0000	0.0002	0.0000	0.0002	0.0000	0.0542
b_{13}^P	0.0001	0.0898	0.0001	0.0898	0.0223	0.4930
b_{21}^P	27.8084	29.9284	27.8084	29.9284	4.5075	7.1155
b_{22}^P	−8.5039	0.6301	−8.5039	0.6301	−3.3183	0.2526
b_{23}^P	0.0001	6.0464	0.0001	6.0464	28.1601	32.3955
b_{31}^P	0.0333	0.0217	0.0333	0.0217	0.0684	0.0846
b_{32}^P	0.0000	0.0023	0.0000	0.0023	0.0000	0.0181
b_{33}^P	−0.4901	0.1272	−0.4901	0.1272	−0.4064	0.2235
a_1^Q	0.7739	0.8281	0.7739	0.8281	0.7752	285.6
a_2^Q	30614	63165	30614	63165	27870	74539
a_3^Q	2.9081	8.4479	−2.9081	8.4479	2.2788	65.5446
b_{11}^Q	−0.0098	0.0223	0.0098	0.0223	−0.0335	0.1586
b_{12}^Q	0.0000	0.0002	0.0000	0.0002	0.0000	0.0830
b_{13}^Q	0.0001	0.0898	0.0001	0.0898	0.1079	1.1929
b_{21}^Q	27.8084	29.9284	27.8084	29.9284	14.8848	22.2814
b_{22}^Q	−8.4748	0.0303	−8.4748	0.0303	−8.0938	0.9588
b_{23}^Q	0.0001	6.0464	0.0001	6.0464	116.1	149.1
b_{31}^Q	0.0333	0.0217	0.0333	0.0217	0.1181	0.0660
b_{32}^Q	0.0000	0.0023	0.0000	0.0023	0.0000	0.0191
b_{33}^Q	−0.7075	0.1232	−0.7075	0.1232	−0.7851	0.3042

This table shows the parameter estimates and standard errors for the $A_3(3)$ model parameters, for the completely affine, essentially affine, and extended affine market price of risk specifications. Zero-coupon bond yields with maturities of one month, and of two and four years are assumed to be observed without error; zero-coupon bond yields with maturities of six, eight, ten, and twelve years are assumed to be observed with error. The completely affine and essentially affine specifications coincide for this model. Note that, for the completely affine and essentially affine market price of risk specifications, the a_1^Q, a_2^Q, and a_3^Q parameters must be equal to their P-measure counterparts, a_1^P, a_3^P, and a_3^P. Furthermore, the b_{12}^Q, b_{13}^Q, b_{21}^Q, b_{23}^Q, b_{31}^Q and b_{32}^Q parameters must be equal to their counterparts under the P measure. For the extended affine market price of risk specification, all parameters can vary independently. This table is continued in Table 13.

deviations of the observation errors are almost identical to the essentially affine statistics. However, there is a large reduction in the bias of the yield forecast errors at every maturity considered. The second moments of yield forecast errors present a more mixed picture, with improvement at some maturities and degradation at others. Thus, the improved fit of the model would appear to come largely from improved bias of yield forecast errors.

For the $A_2(3)$ model, the extended affine specification has a strong likelihood ratio statistic relative to either the completely affine or essentially affine specification, and this is

Table 13
$A_3(3)$ Model estimates (Part II)

Parameter	Completely affine		Essentially affine		Extended affine	
	Estimate	Std. err.	Estimate	Std. err.	Estimate	Std. err.
d_0	–3.5350	3.5991	–3.5350	3.5991	–1.6513	2.1996
d_1	–0.0004	0.0009	–0.0004	0.0009	0.0003	0.0005
d_2	0.0010	0.0010	0.0010	0.0010	0.0005	0.0006
d_3	0.0130	0.0016	0.0130	0.0016	0.0029	0.0007
σ_4	0.0010	0.0000	0.0010	0.0000	0.0010	0.0000
σ_5	0.0018	0.0001	0.0018	0.0001	0.0017	0.0001
σ_6	0.0023	0.0001	0.0023	0.0001	0.0023	0.0001
σ_7	0.0027	0.0001	0.0027	0.0001	0.0027	0.0001
ρ_{54}	0.9126	0.0083	0.9126	0.0083	0.9127	0.0100
ρ_{64}	0.7742	0.0182	0.7742	0.0182	0.7737	0.0216
ρ_{65}	0.9345	0.0058	0.9345	0.0058	0.9337	0.0062
ρ_{74}	0.6928	0.0244	0.6928	0.0244	0.6940	0.0263
ρ_{75}	0.8255	0.0169	0.8255	0.0169	0.8259	0.0170
ρ_{76}	0.9529	0.0046	0.9529	0.0046	0.9538	0.0045

This table is a continuation of Table 12.

Table 14
Likelihood ratio statistics

Model	Ess. Aff. vs. Comp. Aff.			Ext. Aff. vs. Comp. Aff.			Ext. Aff. vs. Ess. Aff.		
	DF	95% Cutoff	LR	DF	95% Cutoff	LR	DF	95% Cutoff	LR
$A_0(1)$	1	3.84	4.49	1	3.84	4.49	0	—	—
$A_1(1)$	0	—	—	1	3.84	0.00	1	3.84	0.00
$A_0(2)$	4	9.49	13.56	4	9.49	13.56	0	—	—
$A_1(2)$	2	5.99	5.56	3	7.82	5.65	1	3.84	0.09
$A_2(2)$	0	—	—	4	9.49	15.21	4	9.49	15.21
$A_0(3)$	9	16.92	52.35	9	16.92	52.35	0	—	—
$A_1(3)$	6	12.59	42.55	7	14.07	45.66	1	3.84	3.11
$A_2(3)$	3	7.82	58.68	7	14.07	73.61	4	9.49	14.93
$A_3(3)$	0	—	—	9	16.92	342.58	9	16.92	342.58

This table shows likelihood ratio statistics for the different nested market price of risk specifications within each of the nine affine yield models considered. The first column lists the model under consideration. The next three columns contain information on the likelihood ratio of the completely affine yield market price of risk specification, relative to the essentially affine specification, which nests the completely affine specification. The following three columns contain analogous information for the completely affine specification relative to the extended affine specification, which nests both the other specifications. The last three columns compare the essentially affine specification to the nesting extended affine specification. For each comparison, the column labeled DF lists the additional degrees of freedom contained in the nesting model. The column labeled Cutoff contains the 95% chi-squared cutoff value for a likelihood ratio statistic with degrees of freedom corresponding to the number in the DF column. The column labeled LR contains the actual likelihood ratio statistic. The hypothesis that the restrictions included in the less flexible model are valid is rejected if the quantity in the LR column is greater than the quantity in the Cutoff column. Six of the 27 comparisons we consider are degenerate, in that the restricted and nesting models coincide. In these six cases, the DF column contains the value 0, and the Cutoff and LR columns are not filled in.

Table 15
Likelihood ratio statistics—two stage estimation procedure

Model	Ess. Aff. vs. Comp. Aff.			Ess. Aff. vs. Comp. Aff.			Ess. Aff. vs. Comp. Aff.		
	DF	95% Cutoff	LR	DF	95% Cutoff	LR	DF	95% Cutoff	LR
$A_0(1)$	1	3.84	4.46	1	3.84	4.46	0	—	—
$A_1(1)$	0	—	—	1	3.84	0.00	1	3.84	0.00
$A_0(2)$	4	9.49	16.89	4	9.49	16.89	0	—	—
$A_1(2)$	2	5.99	16.10	3	7.82	16.11	1	3.84	0.01
$A_2(2)$	0	—	—	4	9.49	16.14	4	9.49	16.14
$A_0(3)$	9	16.92	38.82	9	16.92	38.82	0	—	—
$A_1(3)$	6	12.59	28.50	7	14.07	32.53	1	3.84	4.03
$A_2(3)$	3	7.82	48.14	7	14.07	74.35	4	9.49	26.21
$A_3(3)$	0	—	—	9	16.92	342.37	9	16.92	342.37

This table shows likelihood ratio statistics for the different nested market price of risk specifications within each of the nine affine yield models considered, analogous to the likelihood ratio statistics in Table 14, but based on the two-stage estimation procedure. The first column lists the model under consideration. The next three columns contain information on the likelihood ratio of the completely affine yield market price of risk specification, relative to the essentially affine specification, which nests the completely affine specification. The following three columns contain analogous information for the completely affine specification relative to the extended affine specification, which nests both the other specifications. The last three columns compare the essentially affine specification to the nesting extended affine specification. For each comparison, the column labeled DF lists the additional degrees of freedom contained in the nesting model. The column labeled Cutoff contains the 95% chi-squared cutoff value for a likelihood ratio statistic with degrees of freedom corresponding to the number in the DF column. The column labeled LR contains the actual likelihood ratio statistic. The hypothesis that the restrictions included in the less flexible model are valid is rejected if the quantity in the LR column is greater than the quantity in the Cutoff column. Six of the 27 comparisons we consider are degenerate, in that the restricted and nesting models coincide. In these six cases, the DF column contains the value 0, and the Cutoff and LR columns are not filled in.

reflected in an improvement in both the first and second moments of yield forecast errors for all maturities. The improvement is sometimes substantial. The bias of observation errors is also reduced in magnitude (relative to the essentially affine values) for every single maturity, often substantially, although it should be noted that in some cases there is a slight degradation relative to the completely affine specification. Only the standard deviations of the observation errors fail to show improvement; in these cases, there is often a very small degradation. The improved fit of the $A_2(3)$ model therefore appears to derive largely from its improved time-series properties.

The huge statistical significance of the extended affine specification for the $A_3(3)$ model appears to manifest itself in often large improvements in every performance measurement at every maturity, with the exception of standard deviations of the observation errors (which often show a very slight degradation). As with the $A_2(3)$ model, the change appears to be largely driven by the model's time-series properties, with statistics that often show large improvement (perhaps not surprisingly, given the large likelihood ratio statistic) relative to the completely affine and essentially affine specifications (which coincide for this model).

For the three-factor models, we also consider some comparisons between models. The $A_I(3)$ and $A_J(3)$ models are not nested whenever $I \neq J$. Consequently, we cannot calculate

Table 16
Observation error means

Model	MPR	Maturity					
		2 yr	4 yr	6 yr	8 yr	10 yr	12 yr
$A_0(1)$	Comp. Aff.	−68.4	−73.3	−74.5	−74.0	—	—
	Ess. Aff.	−68.0	−72.7	−73.8	−73.3	—	—
	Ext. Aff.	−68.0	−72.7	−73.8	−73.3	—	—
	Comp. Aff.	−76.4	−86.1	−89.3	−88.8	—	—
$A_1(1)$	Ess. Aff.	−76.4	−86.1	−89.3	−88.8	—	—
	Ext. Aff.	−76.4	−86.1	−89.3	−88.8	—	—
	Comp.Aff.	—	2.2	2.2	3.9	3.8	—
$A_0(2)$	Ess. Aff.	—	2.3	2.4	4.1	4.0	—
	Ext. Aff.	—	2.3	2.4	4.1	4.0	—
	Comp. Aff.	—	1.0	0.4	1.8	1.6	—
$A_1(2)$	Ess. Aff.	—	0.4	−0.4	0.9	0.7	—
	Ext. Aff.	—	0.4	−0.4	0.9	0.8	—
	Comp. Aff.	—	2.5	2.5	4.1	4.0	—
$A_2(2)$	Ess. Aff.	—	2.5	2.5	4.1	4.0	—
	Ext. Aff.	—	1.7	1.5	3.0	2.9	—
	Comp. Aff.	—	—	−0.7	1.3	2.5	1.2
$A_0(3)$	Ess. Aff.	—	—	−0.1	2.1	3.4	1.8
	Ext. Aff.	—	—	−0.1	2.1	3.4	1.8
	Comp.Aff.	—	—	−0.8	1.1	2.4	1.0
$A_1(3)$	Ess. Aff.	—	—	−0.1	2.0	3.3	1.7
	Ext. Aff.	—	—	−0.1	2.0	3.3	1.7
	Comp.Aff.	—	—	−0.5	1.5	2.8	1.4
$A_2(3)$	Ess. Aff.	—	—	0.1	2.4	3.6	2.1
	Ext. Aff.	—	—	−0.1	2.1	3.3	1.8
	Comp. Aff.	—	—	0.3	2.8	4.3	2.7
$A_3(3)$	Ess. Aff.	—	—	0.3	2.8	4.3	2.7
	Ext. Aff.	—	—	−0.2	1.9	3.2	1.6

This table shows the mean (in basis points) of the observation error for those yields observed with error. These values are calculated as the mean of the difference between the predicted yield (where the current value of the state vector is extracted from those yields assumed observed without error) and the observed yield. The result is shown for each model and for each market price of risk specification. For the single-factor models, the two, four, six, and eight year yields are observed with error; for the two-factor models, the four, six, eight, and ten year yields are observed with error, and for the three-factor models, the six, eight, ten, and twelve year yields are observed with error. Of the 27 combinations of base model and market price of risk specification, only 21 are distinct, due to the degeneracy of the essentially affine specification in three cases, and of the extended affine specification in three others.

likelihood ratio statistics for between-model comparisons. However, we can qualitatively examine the first two moments of yield forecasts. Of particular note are the three-factor models. The $A_1(3)$ model is preferred by Dai and Singleton (2000), who use the completely affine market price of risk. The extended affine specification offers improvement over the essentially affine specification used by Duffee (2002) for the $A_1(3)$ model, as indicated by

Table 17
Observation error standard deviations

Model	MPR	Maturity					
		2 yr	4 yr	6 yr	8 yr	10 yr	12 yr
$A_0(1)$	Comp. Aff.	90.2	112.5	123.7	128.4	—	—
	Ess. Aff.	90.2	112.5	123.8	128.4	—	—
	Ext. Aff.	90.2	112.5	123.8	128.4	—	—
	Comp. Aff.	91.4	115.9	127.6	131.9	—	—
$A_1(1)$	Ess. Aff.	91.4	115.9	127.6	131.9	—	—
	Ext. Aff.	91.4	115.9	127.6	131.9	—	—
	Comp. Aff.	—	32.9	48.5	57.0	62.5	—
$A_0(2)$	Ess. Aff.	—	32.9	48.5	57.0	62.5	—
	Ext. Aff.	—	32.9	48.5	57.0	62.5	—
	Comp. Aff.	—	34.4	50.5	59.2	64.7	—
$A_1(2)$	Ess. Aff.	—	34.4	50.6	59.3	64.7	—
	Ext. Aff.	—	34.4	50.6	59.3	64.7	—
	Comp. Aff.	—	34.2	50.2	58.8	64.2	—
$A_2(2)$	Ess. Aff.	—	34.2	50.2	58.8	64.2	—
	Ext. Aff.	—	34.6	50.8	59.5	64.9	—
	Comp. Aff.	—	—	10.4	17.8	22.6	27.0
$A_0(3)$	Ess. Aff.	—	—	10.2	17.6	22.5	27.2
	Ext. Aff.	—	—	10.2	17.6	22.5	27.2
	Comp. Aff.	—	—	10.2	17.5	22.5	27.0
$A_1(3)$	Ess. Aff.	—	—	10.1	17.4	22.5	27.3
	Ext. Aff.	—	—	10.1	17.4	22.5	27.3
	Comp. Aff.	—	—	10.2	17.6	22.5	26.9
$A_2(3)$	Ess. Aff.	—	—	10.1	17.4	22.4	27.2
	Ext. Aff.	—	—	10.1	17.4	22.5	27.3
	Comp. Aff.	—	—	10.1	17.4	22.4	27.0
$A_3(3)$	Ess. Aff.	—	—	10.1	17.4	22.4	27.0
	Ext. Aff.	—	—	10.1	17.4	22.5	27.2

This table shows the standard deviation (in basis points) of the observation error for those yields observed with error. These values are calculated as the standard deviation of the difference between the predicted yield (where the current value of the state vector is extracted from those yields assumed observed without error) and the observed yield. The result is shown for each model and for each market price of risk specification. For the single-factor models, the two, four, six, and eight year yields are observed with error; for the two-factor models, the four, six, eight, and ten year yields are observed with error, and for the three-factor models, the six, eight, ten, and twelve year yields are observed with error. Of the 27 combinations of base model and market price of risk specification, only 21 are distinct, due to the degeneracy of the essentially affine specification in three cases, and of the extended affine specification in three others.

the likelihood ratio statistic (significant at over the 90% level) and the statistics on observation errors and yield forecast errors. However, the extended affine specification results in a large improvement in the fit of the $A_2(3)$ and $A_3(3)$ models. Since the $A_1(3)$ model is preferred by other authors, the essentially affine specification may well be worthwhile based only on the evidence pertaining to this model. However, the large

Table 18
Observation error autocorrelations

Model	MPR	Maturity					
		2 yr	4 yrs	6yr	8 yr	10 yr	12 yr
$A_0(1)$	Comp. Aff.	0.83	0.85	0.87	0.88	—	—
	Ess. Aff.	0.83	0.85	0.87	0.88	—	—
	Ext. Aff.	0.83	0.85	0.87	0.88	—	—
	Comp. Aff.	0.82	0.85	0.87	0.88	—	—
$A_1(1)$	Ess. Aff.	0.82	0.85	0.87	0.88	—	—
	Ext. Aff.	0.82	0.85	0.87	0.88	—	—
	Comp. Aff.	—	0.90	0.91	0.91	0.91	—
$A_0(2)$	Ess. Aff.	—	0.90	0.91	0.91	0.91	—
	Ext. Aff.	—	0.90	0.91	0.91	0.91	—
	Comp. Aff.	—	0.90	0.91	0.91	0.91	—
$A_1(2)$	Ess. Aff.	—	0.91	0.91	0.91	0.92	—
	Ext. Aff.	—	0.91	0.91	0.91	0.92	—
	Comp. Aff.	—	0.90	0.90	0.91	0.91	—
$A_2(2)$	Ess. Aff.	—	0.90	0.90	0.91	0.91	—
	Ext. Aff.	—	0.90	0.91	0.91	0.91	—
	Comp. Aff.	—	—	0.68	0.70	0.69	0.67
$A_0(3)$	Ess. Aff.	—	—	0.68	0.70	0.69	0.68
	Ext. Aff.	—	—	0.68	0.70	0.69	0.68
	Comp. Aff.	—	—	0.67	0.69	0.68	0.67
$A_1(3)$	Ess. Aff.	—	—	0.67	0.69	0.68	0.67
	Ext. Aff.	—	—	0.67	0.69	0.68	0.67
	Comp. Aff.	—	—	0.67	0.69	0.68	0.66
$A_2(3)$	Ess. Aff.	—	—	0.66	0.68	0.67	0.67
	Ext. Aff.	—	—	0.66	0.69	0.68	0.67
	Comp. Aff.	—	—	0.65	0.68	0.66	0.66
$A_3(3)$	Ess. Aff.	—	—	0.65	0.68	0.66	0.66
	Ext. Aff.	—	—	0.67	0.69	0.68	0.67

This table shows the first order autocorrelation of the observation errors for those yields observed with error, i.e., the sample correlation between the observation errors for a particular maturity, and lagged values of the observation errors of the same maturity. Correlations between observation errors and lagged values of observation errors associated with different maturities are not shown. The result is shown for each model and for each market price of risk specification. For the single-factor models, the two, four, six, and eight year yields are observed with error; for the two-factor models, the four, six, eight, and ten year yields are observed with error, and for the three-factor models, the six, eight, ten, and twelve year yields are observed with error. Of the 27 combinations of base model and market price of risk specification, only 21 are distinct, due to the degeneracy of the essentially affine specification in three cases, and of the extended affine specification in three others.

improvement for other models such as the $A_2(3)$ model may change the preference ordering of models. Thus, even if an econometrician prefers the $A_1(3)$ model when restricted to the essentially affine specification, he might prefer the $A_2(3)$ or $A_3(3)$ model with the extended affine market price of risk specification.

Table 19
Mean yield forecast errors

Model	MPR	Maturity						
		1month	2 yr	4 yr	6 yr	8 yr	10 yr	12 yr
$A_0(1)$	Comp. Aff.	−0.00	68.22	73.06	74.28	73.79	—	—
	Ess. Aff.	−0.00	67.87	72.47	73.57	73.07	—	—
	Ext. Aff.	−0.00	67.87	72.47	73.57	73.07	—	—
	Comp. Aff.	0.00	76.25	85.78	89.03	88.57	—	—
$A_1(1)$	Ess. Aff.	0.00	76.25	85.78	89.03	88.57	—	—
	Ext. Aff.	0.00	76.25	85.78	89.03	88.57	—	—
	Comp. Aff.	−0.00	−0.00	−2.29	−2.29	−3.92	−3.83	—
$A_0(2)$	Ess. Aff.	−0.00	−0.00	−2.38	−2.46	−4.11	−4.01	—
	Ext. Aff.	−0.00	−0.00	−2.38	−2.46	−4.11	−4.01	—
	Comp. Aff.	−1.99	−3.66	−4.89	−4.28	−5.58	−5.33	—
$A_1(2)$	Ess. Aff.	−0.00	−0.00	−0.49	0.33	−0.95	−0.80	—
	Ext. Aff.	−0.00	−0.00	−0.52	0.28	−1.00	−0.85	—
	Comp. Aff.	−0.01	−0.01	−2.56	−2.54	−4.15	−4.11	—
$A_2(2)$	Ess. Aff.	−0.01	−0.01	−2.56	−2.54	−4.15	−4.11	—
	Ext. Aff.	−0.00	−0.00	−1.80	−1.56	−3.08	−2.94	—
	Comp. Aff.	0.00	0.00	0.00	0.74	−1.22	−2.47	−1.07
$A_0(3)$	Ess. Aff.	0.10	0.11	0.09	0.20	−2.00	−3.23	−1.69
	Ext. Aff.	0.10	0.11	0.09	0.20	−2.00	−3.23	−1.69
	Comp. Aff.	−1.42	−2.65	−2.87	−2.10	−4.00	−5.18	−3.70
$A_1(3)$	Ess. Aff.	−1.23	−2.72	−2.94	−2.84	−4.98	−6.13	−4.49
	Ext. Aff.	−0.00	−0.00	−0.00	0.16	−2.00	−3.19	−1.63
	Comp. Aff.	−1.20	−2.43	−2.74	−2.28	−4.35	−5.58	−4.08
$A_2(3)$	Ess. Aff.	−0.97	−2.51	−2.81	−2.95	−5.25	−6.45	−4.81
	Ext. Aff.	−0.39	−1.21	−1.41	−1.38	−3.56	−4.75	−3.16
	Comp. Aff.	−2.77	−9.59	−7.14	−6.06	−7.81	−8.70	−6.78
$A_3(3)$	Ess. Aff.	−2.77	−9.59	−7.14	−6.06	−7.81	−8.70	−6.78
	Ext. Aff.	−0.78	−2.25	−2.62	−2.50	−4.65	−5.84	−4.26

This table shows the mean monthly forecast error (in basis points) for all maturities used to estimate a particular model. These values are calculated as the mean of the difference between predicted and observed yield changes, with equal weight given to all observations. The result is shown for each model and for each market price of risk specification. Of the 27 combinations of base model and market price of risk specification, only 21 are distinct, due to the degeneracy of the essentially affine specification in three cases, and of the extended affine specification in three others.

To summarize, the extended affine specification is statistically significant at the 90% confidence level for four of the six models in which it is more general than the essentially affine specification. For three of the six models (those with multiple restricted state variables), the statistical improvement is very large, with likelihood ratio tests very far above the 95% cutoff value. The parameter estimates suggest that usually, the need to match the cross-sectional shape of the term structure dominates the need to match the time series behavior of yields in the more restricted market price of risk specifications.

Table 20
Second moments of yield forecast errors

Model	MPR	Maturity						
		1month	2 yr	4yr	6 yr	8 yr	10 yr	12 yr
$A_0(1)$	Comp. Aff.	−0.44	9.18	9.91	10.55	10.82	—	—
	Ess. Aff.	−0.02	9.52	10.18	10.78	11.01	—	—
	Ext. Aff.	−0.02	9.52	10.18	10.78	11.01	—	—
	Comp. Aff.	0.14	10.44	11.62	12.51	12.83	—	—
$A_1(1)$	Ess. Aff.	0.14	10.44	11.62	12.51	12.83	—	—
	Ext. Aff.	0.14	10.44	11.62	12.51	12.83	—	—
	Comp. Aff.	0.04	−0.23	−1.16	−1.07	−1.22	−1.18	—
$A_0(2)$	Ess. Aff.	−0.09	0.07	−0.84	−0.77	−0.94	−0.91	—
	Ext. Aff.	−0.09	0.07	−0.84	−0.77	−0.94	−0.91	—
	Comp. Aff.	0.02	−0.50	−1.46	−1.41	−1.57	−1.51	—
$A_1(2)$	Ess. Aff.	0.47	0.16	−0.68	−0.58	−0.72	−0.68	—
	Ext. Aff.	0.49	0.20	−0.65	−0.55	−0.70	−0.66	—
	Comp. Aff.	0.71	0.21	−1.01	−1.08	−1.30	−1.28	—
$A_2(2)$	Ess. Aff.	0.71	0.21	−1.01	−1.08	−1.30	−1.28	—
	Ext. Aff.	0.49	0.16	−0.94	−0.97	−1.17	−1.13	—
	Comp. Aff.	−0.37	−0.14	−0.14	0.16	−0.08	−0.33	−0.12
$A_0(3)$	Ess. Aff.	−0.17	0.08	0.12	0.31	0.02	−0.24	−0.02
	Ext. Aff.	−0.17	0.08	0.12	0.31	0.02	−0.24	−0.02
	Comp. Aff.	−0.42	−0.28	−0.37	−0.13	−0.41	−0.68	−0.45
$A_1(3)$	Ess. Aff.	0.33	−0.18	−0.33	−0.24	−0.57	−0.84	−0.59
	Ext. Aff.	0.52	0.31	0.22	0.33	0.00	−0.27	−0.03
	Comp. Aff.	−0.38	−0.22	−0.33	−0.15	−0.45	−0.73	−0.50
$A_2(3)$	Ess. Aff.	0.45	−0.10	−0.29	−0.24	−0.60	−0.88	−0.63
	Ext. Aff.	0.33	−0.03	−0.12	−0.02	−0.35	−0.62	−0.37
	Comp. Aff.	−1.35	−1.37	−1.05	−0.79	−1.06	−1.29	−0.99
$A_3(3)$	Ess. Aff.	−1.35	−1.37	−1.05	−0.79	−1.06	−1.29	−0.99
	Ext. Aff.	0.31	0.20	0.33	0.21	0.54	0.81	0.56

This table shows the difference between the observed second moments of monthly yield changes (in percent—note that the units are different than those of Table 19) and the second moments predicted by the model parameters, for all maturities used to estimate a particular model. These values are calculated as the difference between the empirical second moment of monthly yield changes and the mean value of the second conditional moment predicted by the model, across all observations, with equal weight given to all observations. The result is shown for each model and for each market price of risk specification. Of the 27 combinations of base model and market price of risk specification, only 21 are distinct, due to the degeneracy of the essentially affine specification in three cases, and of the extended affine specification in three others.

This finding is confirmed by the examination of the observation errors, which improve little or not at all when introducing the extended specification. By contrast, yield forecasts are generally more accurate under the extended specification, with the improvement largely in line with the statistical strength of the results. (The notable exception is the $A_1(3)$ model, which shows huge improvements in the bias of yield forecasts for short maturities, despite its lack of statistical significance.) For some models, the improvement is in the bias of the yield forecasts, for others, it is in the improved accuracy of the second moments of yield changes, and for still others, there is substantial improvement in both. Finally, the extended affine specification appears to have the greatest effect on models that are less preferred by previous authors; this large improvement may change an econometrician's choice of preferred model. In particular, the performance of the $A_2(3)$ and $A_3(3)$ models appears to improve substantially relative to the $A_1(3)$ model with the introduction of the extended specification.

6. Conclusion

We introduce a new market price of risk specification for affine diffusions, show that this specification does not offer arbitrage opportunities, and demonstrate that the new specification provides a better fit to U.S. term structure data than do standard specifications for most affine yield models. Our specification is particularly important for models with two or more restricted state variables, for which likelihood ratio statistics for the extended specification are always higher than the 95% cutoff values, often substantially so. Although each model is different, it seems that the additional flexibility that our specification offers helps relieve the tension between matching the time-series behavior of the interest rate process and matching the cross-sectional shape of the yield curve. The former is determined by the parameters of the interest rate process under the objective probability measure; the latter is determined by the parameters under an equivalent martingale measure. Traditional market price of risk specifications for affine diffusions constrain many of the parameters to be the same under both measures, so that the same parameters must capture both aspects of interest rate and term structure behavior. By contrast, our specification allows the parameters under the two measures to differ essentially arbitrarily, subject only to existence and boundary nonattainment considerations. That is, rather than having one set of parameters do two jobs, we have a separate set of parameters for each task. The increased flexibility seems to result in a dramatically better fit for some models. Note that our formal statistical results compare only different market price of risk specifications for the same model (e.g., completely, essentially, and extended affine for the $A_2(3)$ model), but make no comparisons between families of affine yield models (e.g., $A_1(2)$ versus $A_2(3)$ model). Such comparisons cannot be made using traditional statistical measures such as likelihood ratios, because the models are not nested. However, it is possible to make comparisons between nonnested models using ad hoc measures of fit, such as the moments of yield forecast errors or means and standard deviations of observation errors. Such measures suggest that the introduction of the extended affine market price of risk improves the quality of models with many restricted state variables (e.g., the $A_2(3)$ or $A_3(3)$ models) relative to those with fewer restricted state variables (e.g., the $A_0(3)$ or $A_1(3)$ models). Models in this latter category benefit from the introduction of the essentially affine market price of risk of Duffee (2002); the extended affine market price of risk allows models in the former category to catch up, and perhaps surpass, those in the latter category.

If the two models are nested, the likelihood ratio tests we apply could also be applied in this manner, provided the data set used is the same for both models (i.e., the same zero coupon bond maturities are used). Such a comparison is necessarily a test of both the underlying models and the observation error specification.

Our technique is limited neither to term structure applications nor to affine models. Stochastic volatility models of equity prices, such as Heston (1993), often have a volatility state variable that follows a square-root type process. Our specification can be readily applied to such models, allowing for a more flexible treatment of volatility risk. Similarly, international models of interest rates and exchange rates, such as Brandt and Santa-Clara (2002), use square-root type processes, and may also benefit from a more flexible market price of risk. Furthermore, the proof of absence of arbitrage does not depend in any essential way on the affinity of the drifts, variances, and covariances of the state variables. Rather, what is needed is the existence and uniqueness in distribution of a process with risk-neutral dynamics implied by the market price of risk specification, and the requirement that the state variables not achieve their boundary values under either measure. Our technique might therefore be applied to some nonaffine models as well.

Appendix A. Proof of theorem 1

Theorem 1 is a consequence of Theorem 2.7 in Duffie et al. (2003) and Theorem 2.4 in Cheridito et al. (2005). We present a version of the proof adapted to the affine diffusions considered in this paper.

Parts (a) and (b) follow from Theorem 2.7 in Duffie, Filipović, and Schachermayer (2003). To show (c), we fix $Y_0 \in D$ and $T > 0$ Since the pair $(\mu^P(\cdot), \sigma(\cdot))$ satisfies the existence and boundary nonattainment conditions, the market price of risk

$$\lambda(Y_t) = \sigma(Y_t)^{-1}\left[\mu^P(Y_t) - \mu^Q(Y_t)\right], \quad t \geqslant 0 \tag{52}$$

is a well-defined continuous process. Therefore,

$$Z_t = \exp\left(-\int_0^t \lambda(Y_s)^T \, \mathrm{d}W_s^P - \frac{1}{2}\int_0^t \lambda(Y_s)^T \lambda(Y_s) \, \mathrm{d}s\right), \quad t \in [0, T] \tag{53}$$

is a well-defined, positive local martingale with respect to P, and thus also a P-supermartingale. Hence, if we can show that

$$\mathrm{E}^P[Z_T] = 1, \tag{54}$$

then $(Z_t)_{t \in [0,T]}$ is a P-martingale, $Q = Z_T \cdot P$ is a probability measure equivalent to P, and by Girsanov's theorem, the process

$$W_t^Q = W_t^P + \int_0^t \lambda(Y_s) \, \mathrm{d}s, \quad t \in [0, T] \tag{55}$$

is a Brownian motion under Q. Moreover,

$$Y_t = Y_0 + \int_0^t \mu^Q(Y_s) \, \mathrm{d}s + \int_0^t \sigma(Y_s) \, \mathrm{d}W_s^Q, \quad t \in [0, T] \tag{56}$$

and (c) is proved.

It remains to show (54). By (a), there exists a stochastic process $(\tilde{Y}_t)_{t \geqslant 0}$ on (Ω, F, P) that satisfies

$$\tilde{Y}_t = Y_0 + \int_0^t \mu^Q(\tilde{Y}_s) \, \mathrm{d}s + \int_0^t \sigma(\tilde{Y}_s) \, \mathrm{d}W_s^P, \quad t \geqslant 0 \tag{57}$$

and by (b), the distribution of $(\tilde{Y}_t)_{t\geqslant 0}$ is unique. Since the pair $(\mu^P(\cdot), \sigma(\cdot))$ also satisfies the existence and boundary nonattainment conditions,

$$\lambda(\tilde{Y}_t) = \sigma(\tilde{Y}_t)^{-1}\left[\mu^P(\tilde{Y}_t) - \mu^Q(\tilde{Y}_t)\right], \quad \geqslant 0 \tag{58}$$

is a well-defined continuous process. For each $n\geqslant 1$, we define the stopping times

$$\tau_n = \inf\left\{t>0 \middle| \left\|\lambda(Y_t)\right\|_2 \geqslant n\right\}\Lambda T \tag{59}$$

and

$$\tilde{\tau}_n = \inf\left\{t>0 \middle| \left\|\lambda\left(\tilde{Y}_t\right)\right\|_2 \geqslant n\right\}\Lambda T, \tag{60}$$

where $\|\lambda(Y_t)\|_2$ denotes the Euclidean norm of the vector $\lambda(Y_t)$ These stopping times satisfy

$$\lim_{n\to\infty} P[\tau_n = T] = \lim_{n\to\infty} P[\tilde{\tau}_n = T] = 1. \tag{61}$$

For each $n\geqslant 1$, we define the process

$$\lambda_t^n = \lambda(Y_t)1_{\{t\leqslant\tau_n\}}, \quad t\in[0,T]. \tag{62}$$

Note that, by construction, $\int_0^t (\lambda_s^n)^T \lambda_s^n \,\mathrm{d}s$ is bounded by $n^2 t$. For each n, the process satisfies the Novikov criterion (under the P measure):

$$\mathrm{E}^P\left[\exp\left(\frac{1}{2}\int_0^T (\lambda_s^n)^T \lambda_s^n \,\mathrm{d}s\right)\right] \leqslant \exp\left(\frac{n^2 T}{2}\right) < \infty. \tag{63}$$

It follows that, for each $n\geqslant 1$, the process defined by

$$Z_t^n = \exp\left(-\int_0^t (\lambda_s^n)^T \,\mathrm{d}W_s^P - \frac{1}{2}\int_0^t (\lambda_s^n)^T \lambda_s^n \,\mathrm{d}s\right), \quad t\in[0,T] \tag{64}$$

is a P-martingale, and by (61), $Z_T^n 1_{\{\tau_n=T\}} = Z_T 1_{\{\tau_n=T\}} \to Z_T$, P-almost surely, as $n\to\infty$. For all $n\geqslant 1$, $Q^n = Z_T^n \cdot P$, is a probability measure equivalent to P, and it follows from Girsanov's theorem that

$$W_t^n = W_t^P + \int_0^t \lambda_s^n \,\mathrm{d}s, \quad t\geqslant 0 \tag{65}$$

is a Brownian motion under Q^n. It is easy to see that

$$Y_{t\Lambda\tau_n} = Y_0 \int_0^{t\Lambda\tau_n} \mu^Q(Y_s)\,\mathrm{d}s + \int_0^{t\Lambda\tau_n} \sigma(Y_s)\,\mathrm{d}W_s^n, \quad t\in[0,T] \tag{66}$$

and it can be deduced from (a), (b), (57) and (66) that under Q^n, the stopped process $(Y_{t\Lambda\tau_n})t\geqslant 0$ has the same distribution as the stopped process $(\tilde{Y}t\Lambda\tilde{\tau}_n)_{t\geqslant 0}$ under P. Therefore

$$\mathrm{E}^P[Z_T] = \lim_{n\to\infty} E^P[Z_T^n 1_{\{\tau_n=T\}}] = \lim_{n\to\infty} Q^n[\tau_n = T] = \lim_{n\to\infty} P[\tilde{\tau}_n = T] = 1. \tag{67}$$

The first step in this chain of equalities follows from Beppo-Levi's monotone convergence theorem. The second step holds by applying the definition of the measures Q^n; note that Z_T^n is the Radon-Nikodym derivative of Q^n with respect to P. The third step follows because the distribution of $(Y_{t\Lambda\tau_n})_{t\geqslant 0}$ under Q^n is the same as the distribution of $(\tilde{Y}_{t\Lambda\tilde{\tau}_n})_{t\geqslant 0}$ under P. The last step follows from (61).

Appendix B. Three-factor affine yield models

There are four distinct families of three-factor affine yield models, namely, the $A_0(3)$, $A_1(3)$, $A_2(3)$, and $A_3(3)$ models. In all four, the interest rate process is given by

$$r_t = d_0 + d_1 Y_t(1) + d_2 Y_t(2) + d_3 Y_t(3). \tag{68}$$

Under the $A_0(3)$ model, the state variables follow the process

$$d\begin{bmatrix} Y_t(1) \\ Y_t(2) \\ Y_t(3) \end{bmatrix} = \begin{bmatrix} b_{11}^P & b_{12}^P & b_{13}^P \\ b_{21}^P & b_{22}^P & b_{23}^P \\ b_{31}^P & b_{32}^P & b_{33}^P \end{bmatrix} \begin{bmatrix} Y_t(1) \\ Y_t(2) \\ Y_t(3) \end{bmatrix} \mathrm{d}t + d\begin{bmatrix} W_t^P(1) \\ W_t^P(2) \\ W_t^P(3) \end{bmatrix}. \tag{69}$$

An $A_0(3)$ model does not have a unique representation unless additional constraints are imposed, since the state variables can be changed through orthogonal rotation. Dai and Singleton (2000) use the identifying restrictions $b_{12}^P = 0$, $b_{13}^P = 0$, and $b_{23}^P = 0$; however, this approach precludes a b matrix with complex eigenvalues. The dynamics of the state variables under the measure Q can be expressed as:

$$d\begin{bmatrix} Y_t(1) \\ Y_t(2) \\ Y_t(3) \end{bmatrix} = \left(\begin{bmatrix} a_1^Q \\ a_2^Q \\ a_3^Q \end{bmatrix} + \begin{bmatrix} b_{11}^Q & b_{12}^Q & b_{13}^Q \\ b_{21}^Q & b_{22}^Q & b_{23}^Q \\ b_{31}^Q & b_{32}^Q & b_{33}^Q \end{bmatrix} \begin{bmatrix} Y_t(1) \\ Y_t(2) \\ Y_t(3) \end{bmatrix} \right) \mathrm{d}t + d\begin{bmatrix} W_t^Q(1) \\ W_t^Q(2) \\ W_t^Q(3) \end{bmatrix}. \tag{70}$$

The market price of risk process is defined as

$$\begin{aligned} \Lambda_t &= [\sigma(Y_t)]^{-1}\left[\mu^P(Y_t) - \mu^Q(Y_t)\right] \\ &= \left(-\begin{bmatrix} a_1^Q \\ a_2^Q \\ a_3^Q \end{bmatrix} + \begin{bmatrix} b_{11}^P - b_{11}^Q & b_{12}^P - b_{12}^Q & b_{13}^P - b_{13}^Q \\ b_{21}^P - b_{21}^Q & b_{22}^P - b_{22}^Q & b_{23}^P - b_{23}^Q \\ b_{31}^P - b_{31}^Q & b_{32}^P - b_{32}^Q & b_{33}^P - b_{33}^Q \end{bmatrix} \begin{bmatrix} Y_t(1) \\ Y_t(2) \\ Y_t(3) \end{bmatrix} \right) \\ &\equiv \begin{bmatrix} \lambda_{10} \\ \lambda_{20} \\ \lambda_{30} \end{bmatrix} + \begin{bmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{bmatrix} \begin{bmatrix} Y_t(1) \\ Y_t(2) \\ Y_t(3) \end{bmatrix}. \end{aligned} \tag{71}$$

As with the $A_0(1)$ and $A_0(2)$ models, the completely affine market price of risk specification restricts the slope coefficients to be zero; only λ_{10}, λ_{20}, and λ_{30} can take on nonzero values. By contrast, the essentially affine specification allows all twelve market price of risk parameters to be nonzero. Both specifications satisfy the Novikov and Kazamaki criteria for some positive finite time interval. Our specification coincides with the essentially affine specification, offering no further generality for the $A_0(3)$ model.

In the $A_1(3)$ model, the state variables follow the process

$$d\begin{bmatrix} Y_t(1) \\ Y_t(2) \\ Y_t(3) \end{bmatrix} = \left(\begin{bmatrix} a_1^P \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} b_{11}^P & 0 & 0 \\ b_{21}^P & b_{22}^P & b_{23}^P \\ b_{31}^P & b_{32}^P & b_{33}^P \end{bmatrix} \begin{bmatrix} Y_t(1) \\ Y_t(2) \\ Y_t(3) \end{bmatrix} \right) \mathrm{d}t$$

$$+\begin{bmatrix} \sqrt{Y_t(1)} & 0 & 0 \\ 0 & \sqrt{\alpha_2 + \beta_{21} Y_t(1)} & 0 \\ 0 & 0 & \sqrt{\alpha_3 + \beta_{31} Y_t(1)} \end{bmatrix} d \begin{bmatrix} W_t^P(1) \\ W_t^P(2) \\ W_t^P(3) \end{bmatrix}, \quad (72)$$

with $\alpha_2, \alpha_3 \in \{0,1\}$. Existence imposes the restrictions $a_1^P \geqslant 0$, $\beta_{21} \geqslant 0$, and $\beta_{31} \geqslant 0$. The first state variable is bounded from below by zero, and nonattainment of the boundary requires $2a_1^P \geqslant 1$. The dynamics under the measure Q are

$$d\begin{bmatrix} Y_t(1) \\ Y_t(2) \\ Y_t(3) \end{bmatrix} = \left(\begin{bmatrix} a_1^Q \\ a_2^Q \\ a_3^Q \end{bmatrix} + \begin{bmatrix} b_{11}^Q & 0 & 0 \\ b_{21}^Q & b_{22}^Q & b_{23}^Q \\ b_{31}^Q & b_{32}^Q & b_{33}^Q \end{bmatrix} \begin{bmatrix} Y_t(1) \\ Y_t(2) \\ Y_t(3) \end{bmatrix} \right) \mathrm{d}t$$

$$+\begin{bmatrix} \sqrt{Y_t(1)} & 0 & 0 \\ 0 & \sqrt{\alpha_2 + \beta_{21} Y_t(1)} & 0 \\ 0 & 0 & \sqrt{\alpha_3 + \beta_{31} Y_t(1)} \end{bmatrix} d \begin{bmatrix} W_t^Q(1) \\ W_t^Q(2) \\ W_t^Q(3) \end{bmatrix}. \quad (73)$$

The market price of risk process is given by

$$\begin{aligned} \Lambda_t &= [\alpha(Y_t)]^{-1} \left[\mu^P(Y_t) - \mu^Q(Y_t)\right] \\ &= \begin{bmatrix} \frac{(a_1^P - a_1^Q)}{\sqrt{Y_t(1)}} + \left(b_{11}^P - b_{11}^Q\right)\sqrt{Y_t(1)} \\ \frac{(-a_2^Q) + (b_{21}^P - b_{21}^Q) Y_t(1) + (b_{22}^P - b_{22}^Q) Y_t(2) + (b_{23}^P - b_{23}^Q) Y_t(3)}{\sqrt{\alpha_2 + \beta_{21} Y_t(1)}} \\ \frac{(-a_3^Q) + (b_{31}^P - b_{31}^Q) Y_t(1) + (b_{32}^P - b_{32}^Q) Y_t(2) + (b_{33}^P - b_{33}^Q) Y_t(3)}{\sqrt{\alpha_3 + \beta_{31} Y_t(1)}} \end{bmatrix} \\ &\equiv \begin{bmatrix} \frac{\lambda_{10}}{\sqrt{Y_t(1)}} + \lambda_{11}\sqrt{Y_t(1)} \\ \frac{\lambda_{20} + \lambda_{21} Y_t(1) + \lambda_{22} Y_t(2) + \lambda_{23} Y_t(3)}{\sqrt{\alpha_2 + \beta_{21} Y_t(1)}} \\ \frac{\lambda_{30} + \lambda_{31} Y_t(1) + \lambda_{32} Y_t(2) + \lambda_{33} Y_t(3)}{\sqrt{\alpha_3 + \beta_{31} Y_t(1)}} \end{bmatrix}. \end{aligned} \quad (74)$$

Although the λ_{11}, λ_{20}, λ_{21}, λ_{30}, and λ_{31} parameters can be nonzero in the completely affine specification, these parameters must also satisfy the constraints $\alpha_2\lambda_{21} = \beta_{21}\lambda_{20}$ and $\alpha_3\lambda_{31} = \beta_{31}\lambda_{30}$. The essentially affine specification relaxes these restrictions, but still requires that the λ_{10} parameter be zero. We relax this constraint also, requiring only that λ_{10} be such that the boundary nonattainment condition is satisfied under the measure Q as well. This condition is satisfied if

$$\lambda_{10} \leqslant a_1^P - \frac{1}{2}. \quad (75)$$

When λ_{10} is not zero, it is unclear whether the Novikov and Kazamaki criteria are satisfied.

The $A_2(3)$ model has dynamics as follows:

$$d\begin{bmatrix} Y_t(1) \\ Y_t(2) \\ Y_t(3) \end{bmatrix} = \left(\begin{bmatrix} a_1^P \\ a_2^P \\ 0 \end{bmatrix} + \begin{bmatrix} b_{11}^P & b_{12}^P & 0 \\ b_{21}^P & b_{22}^P & 0 \\ b_{31}^P & b_{32}^P & b_{33}^P \end{bmatrix} \begin{bmatrix} Y_t(1) \\ Y_t(2) \\ Y_t(3) \end{bmatrix} \right) \mathrm{d}t + \begin{bmatrix} \sqrt{Y_t(1)} & 0 & 0 \\ 0 & \sqrt{Y_t(1)} & 0 \\ 0 & 0 & \sqrt{\alpha_3 + \beta_{31} Y_t(1) + \beta_{32} Y_t(1)} \end{bmatrix} d \begin{bmatrix} W_t^P(1) \\ W_t^P(2) \\ W_t^P(3) \end{bmatrix}, \tag{76}$$

with $\alpha_3 \in \{0,1\}$. Existence considerations require $a_1^P \geqslant 0$, $a_2^P \geqslant 0$, $b_{12}^P \geqslant 0$, $b_{21}^P \geqslant 0$, $\beta_{31}^P \geqslant 0$, and $\beta_{32}^P \geqslant 0$.

The boundary is not attained if $2a_1^P \geqslant 1$ and $2a_2^P \geqslant 1$.

The dynamics under the measure Q are given by

$$d\begin{bmatrix} Y_t(1) \\ Y_t(2) \\ Y_t(3) \end{bmatrix} = \left(\begin{bmatrix} a_1^Q \\ a_2^Q \\ a_3^Q \end{bmatrix} + \begin{bmatrix} b_{11}^Q & b_{12}^Q & 0 \\ b_{21}^Q & b_{22}^Q & 0 \\ b_{31}^Q & b_{32}^Q & b_{33}^Q \end{bmatrix} \begin{bmatrix} Y_t(1) \\ Y_t(2) \\ Y_t(3) \end{bmatrix} \right) \mathrm{d}t + \begin{bmatrix} \sqrt{Y_t(1)} & 0 & 0 \\ 0 & \sqrt{Y_t(2)} & 0 \\ 0 & 0 & \sqrt{\alpha_3 + \beta_{31} Y_t(1) + \beta_{32} Y_t(2)} \end{bmatrix} d \begin{bmatrix} W_t^Q(1) \\ W_t^Q(2) \\ W_t^Q(3) \end{bmatrix}. \tag{77}$$

The market price of risk process is given by

$$\begin{aligned} \Lambda_t &= [\sigma(Y_t)]^{-1} \left[\mu^P(Y_t) - \mu^Q(Y_t) \right] \\ &= \begin{bmatrix} \frac{(a_1^P - a_1^Q) + (b_{11}^P - b_{11}^Q) Y_t(1) + (b_{12}^P - b_{12}^Q) Y_t(2)}{\sqrt{Y_t(1)}} \\ \frac{(a_2^P - a_2^Q) + (b_{21}^P - b_{21}^Q) Y_t(1) + (b_{22}^P - b_{22}^Q) Y_t(2)}{\sqrt{Y_t(2)}} \\ \frac{(-a_3^P) + (b_{31}^P - b_{31}^Q) Y_t(1) + (b_{32}^P - b_{32}^Q) Y_t(2) + (b_{33}^P - b_{33}^Q) Y_t(3)}{\sqrt{\alpha_3 + \beta_{31} Y_t(1) + \beta_{32} Y_t(2)}} \end{bmatrix} \\ &\equiv \begin{bmatrix} \frac{\lambda_{10} + \lambda_{11} Y_t(1) + \lambda_{12} Y_t(2)}{\sqrt{Y_t(1)}} \\ \frac{\lambda_{20} + \lambda_{21} Y_t(1) + \lambda_{22} Y_t(2)}{\sqrt{Y_t(2)}} \\ \frac{\lambda_{30} + \lambda_{31} Y_t(1) + \lambda_{32} Y_t(2) + \lambda_{33} Y_t(3)}{\sqrt{\alpha_3 + \beta_{31} Y_t(1) + \beta_{32} Y_t(2)}} \end{bmatrix}. \end{aligned} \tag{78}$$

In the completely affine market price of risk specification, five of the parameters (λ_{11}, λ_{22}, λ_{30}, λ_{31}, and λ_{32}) can be nonzero; however, there are only three degrees of freedom, since the restrictions $\beta_{31}\beta_{32}\beta_{30} = \alpha_3\beta_{32}\lambda_{31} = \alpha_3\beta_{31}\lambda_{32}$ are also imposed. The essentially affine specification relaxes these restrictions, but still requires that λ_{10}, λ_{12}, λ_{20} and λ_{21} be zero. We further relax these restrictions, and allow all parameters to take any values such that boundary nonattainment conditions are satisfied under both Q as well as P:

$$\lambda_{10} \leqslant a_1^P - \frac{1}{2} \tag{79}$$

$$\lambda_{20} \leqslant a_2^P - \frac{1}{2} \tag{80}$$

$$\lambda_{12} \leqslant b_{12}^P \tag{81}$$

$$\lambda_{21} \leqslant b_{21}^P. \tag{82}$$

The $A_3(3)$ model has dynamics

$$d\begin{bmatrix} Y_t(1) \\ Y_t(2) \\ Y_t(3) \end{bmatrix} = \left(\begin{bmatrix} a_1^P \\ a_2^P \\ a_3^P \end{bmatrix} + \begin{bmatrix} b_{11}^P & b_{12}^P & b_{13}^P \\ b_{21}^P & b_{22}^P & b_{23}^P \\ b_{31}^P & b_{32}^P & b_{33}^P \end{bmatrix} \begin{bmatrix} Y_t(1) \\ Y_t(2) \\ Y_t(3) \end{bmatrix} \right) \mathrm{d}t + \begin{bmatrix} \sqrt{Y_t(1)} & 0 & 0 \\ 0 & \sqrt{Y_t(2)} & 0 \\ 0 & 0 & \sqrt{Y_t(3)} \end{bmatrix} d \begin{bmatrix} W_t^P(1) \\ W_t^P(2) \\ W_t^P(3) \end{bmatrix}. \tag{83}$$

Existence considerations require $a_1^P \geqslant 0$, $a_2^P \geqslant 0$, $a_3^P \geqslant 0$, $b_{12}^P \geqslant 0$, $b_{13}^P \geqslant 0$, $b_{21}^P \geqslant 0$, $b_{23}^P \geqslant 0$, $b_{31}^P \geqslant 0$, and $b_{32}^P \geqslant 0$. All three state variables are bounded from below by zero, with boundary nonattainment conditions $2a_1^P \geqslant 1$, $2a_2^P \geqslant 1$, and $2a_3^P \geqslant 1$. Under the measure Q, the state variables follow the dynamics:

$$d\begin{bmatrix} Y_t(1) \\ Y_t(2) \\ Y_t(3) \end{bmatrix} = \left(\begin{bmatrix} a_1^Q \\ a_2^Q \\ a_3^Q \end{bmatrix} + \begin{bmatrix} b_{11}^Q & b_{12}^Q & b_{13}^Q \\ b_{21}^Q & b_{22}^Q & b_{23}^Q \\ b_{31}^Q & b_{32}^Q & b_{33}^Q \end{bmatrix} \begin{bmatrix} Y_t(1) \\ Y_t(2) \\ Y_t(3) \end{bmatrix} \right) \mathrm{d}t + \begin{bmatrix} \sqrt{Y_t(1)} & 0 & 0 \\ 0 & \sqrt{Y_t(2)} & 0 \\ 0 & 0 & \sqrt{Y_t(3)} \end{bmatrix} d \begin{bmatrix} W_t^Q(1) \\ W_t^Q(2) \\ W_t^Q(3) \end{bmatrix}. \tag{84}$$

The market price of risk process is given by

$$\Lambda_t = [\sigma(Y_t)]^{-1}\left[\mu^P(Y_t) - \mu^Q(Y_t)\right]$$

$$= \begin{bmatrix} \frac{(a_1^P - a_1^Q) + (b_{11}^P - b_{11}^Q)Y_t(1) + (b_{12}^P - b_{12}^Q)Y_t(2) + (b_{13}^P - b_{13}^Q)Y_t(3)}{\sqrt{Y_t(1)}} \\ \frac{(a_2^P - a_2^Q) + (b_{21}^P - b_{21}^Q)Y_t(1) + (b_{22}^P - b_{22}^Q)Y_t(2) + (b_{23}^P - b_{23}^Q)Y_t(3)}{\sqrt{Y_t(2)}} \\ \frac{(a_3^P - a_3^Q) + (b_{31}^P - b_{31}^Q)Y_t(1) + (b_{32}^P - b_{32}^Q)Y_t(2) + (b_{33}^P - b_{33}^Q)Y_t(3)}{\sqrt{Y_t(3)}} \end{bmatrix}$$

$$\equiv \begin{bmatrix} \frac{\lambda_{10} + \lambda_{11} Y_t(1) + \lambda_{12} Y_t(2) + \lambda_{13} Y_t(3)}{\sqrt{Y_t(1)}} \\ \frac{\lambda_{20} + \lambda_{21} Y_t(1) + \lambda_{22} Y_t(2) + \lambda_{23} Y_t(3)}{\sqrt{Y_t(2)}} \\ \frac{\lambda_{30} + \lambda_{31} Y_t(1) + \lambda_{32} Y_t(2) + \lambda_{33} Y_t(3)}{\sqrt{Y_t(3)}} \end{bmatrix}. \tag{85}$$

Both the completely affine and essentially affine market price of risk specifications allow only the λ_{11}, λ_{22}, and λ_{33} parameters to be nonzero. By contrast, we allow all twelve market price of risk parameters to be nonzero, requiring only that, as usual, the boundary nonattainment condition is satisfied under the measure Q:

$$\lambda_{10} \leqslant a_1^P - \frac{1}{2}, \tag{86}$$

$$\lambda_{20} \leqslant a_2^P - \frac{1}{2}, \tag{87}$$

$$\lambda_{30} \leqslant a_3^P - \frac{1}{2}, \tag{88}$$

$$\lambda_{12} \leqslant b_{12}^P, \tag{89}$$

$$\lambda_{13} \leqslant b_{13}^P, \tag{90}$$

$$\lambda_{21} \leqslant b_{21}^P, \tag{91}$$

$$\lambda_{23} \leqslant b_{23}^P, \tag{92}$$

$$\lambda_{31} \leqslant b_{31}^P, \tag{93}$$

$$\lambda_{32} \leqslant b_{32}^P. \tag{94}$$

As with the other models in which our specification is more general than traditional specifications, it is unclear whether the Novikov and Kazamaki criteria are satisfied.

References

Aït-Sahalia, Y., 2001. Closed-form likelihood expansions for multivariate discussions. Discussion paper, Princeton University.

Aït-Sahalia, Y., Kimmel, R.L., 2005. Estimating Affine multifactor term structure models using closed-form likelihoods. NBER Working Paper.

Balduzzi, P., Das, S., Foresi, S., Sundaram, R., 1996. A simple approach to three factor affine term structure models. Journal of fixed income 6, 43–53.

Bates, D.S., 1996. Jumps and stochastic volatility: exchange rate processes implicit in Deutsche mark options. Review of financial studies 9, 69–107.
Beaglehole, D., Tenney, M., 1991. General solutions of some interest rate contingent claim pricing equations. Journal of fixed income 1, 69–83.
Bliss, R.R., 1997. Testing term structure estimation methods, vol. 9, advances in futures and options research. JAI Press, Greenwich, Connecticut.
Brandt, M., He, P., 2002. Simulated likelihood estimation of affine term structure models from panel data. Working paper, university of pennsylvania.
Brandt, M., Santa-Clara, P., 2002. Simulated likelihood estimation of diffusions with an application to exchange rate dynamics in incomplete markets. Journal of financial economics 63, 161–210.
Chen, L., 1996. Stochastic mean and stochastic volatility—A three-factor model of the term structure of interest rates and its application to the pricing of interest rate derivatives. Financial markets, institutions, and instruments 5, 1–88.
Cheridito, P., Filipović, D., Kimmel, R.L., 2005. A note on the canonical representation of affine diffusion processes. working paper, Princeton university and university of munich.
Cheridito, P., Filipović, D., Yor, M., 2005. Equivalent and absolutely continuous measure changes for jump-diffusion processes. Annals of Applied Probability 15, 1713–1732.
Chernov, M., Ghysels, E., 2000. A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options evaluation. Journal of financial economics 56, 407–458.
Cox, J.C., Ingersoll, J.E., Ross, S.A., 1985. A theory of the term structure of interest rates. Econometrica 53, 385–408.
Dai, Q., Singleton, K.J., 2000. Specification analysis of affine term structure models. Journal of finance 55, 1943–1978.
Delbaen, F., Schachermayer, W., 1994. A general version of the fundamental theorem of asset pricing. Mathematische annalen 300, 463–520.
Delbaen, F., Schachermayer, W., 1998. The fundamental theorem of asset pricing for unbounded stochastic processes. Mathematische annalen 312, 215–250.
Duffee, G.R., 2002. Term premia and interest rate forecasts in affine models. Journal of finance 57, 405–443.
Duffie, D., Filipović, D., Schachermayer, W., 2003. Affine processes and applications in finance. Annals of applied probability 13, 984–1053.
Duffie, D., Kan, R., 1996. A yield-factor model of interest rates. Mathematical finance 6, 379–406.
Eraker, B., Johannes, M., Polson, N., 2003. The impact of jumps in volatility and returns. Journal of finance 58, 1268–1300.
Feller, W., 1951. Two singular diffusion problems. Annals of mathematics 54, 173–182.
Grasselli, M., Tebaldi, C., 2004. Solvable affine term structure models. working paper, Verona university.
Harrison, M., Kreps, D., 1979. Martingales and arbitrage in multiperiod securities markets. Journal of economic theory 20, 381–408.
Harrison, M., Pliska, S., 1981. Martingales and stochastic integrals in the theory of continuous trading. Stochastic processes and their applications 11, 215–260.
Heston, S., 1993. A closed-form solution of options with stochastic volatility with applications to bonds and currency options. Review of financial studies 6, 327–343.
Ikeda, N., Watanabe, S., 1981. Stochastic differential equations and diffusion processes. North-Holland, Amsterdam.
Ingersoll, J.E., 1987. Theory of financial decision making. Rowman and Littlefield, Savage, Maryland.
Karatzas, I., Shreve, S.E., 1991. Brownian motion and stochastic calculus. Springer, New York.
Levendorskii, S., 2004a. Consistency conditions for affine term structure models. Stochastic processes and their applications 109, 225–261.
Levendorskii, S., 2004b. Consistency conditions for affine term structure models II: option pricing under diffusions with embedded jumps. working paper, University of texas.
McCulloch, J.H., 1975. The tax-adjusted yield curve. Journal of finance 30, 811–830.
McCulloch, J.H., Kwon, H., 1993. US term structure data, 1947–1991. Working paper 93-6, Ohio State University.
Pearson, N., Sun, T.-S., 1994. Exploiting the conditional density in estimating the term structure: an application to the cox, ingersoll, and ross model. Journal of finance 49, 1279–1304.
Revuz, D., Yor, M., 1994. Continuous martingales and brownian motion, Second Edn. Springer, Berlin.
Vasicek, O., 1977. An equilibrium characterization of the term structure. Journal of financial economics 5, 177–188.