Quadratic Term Structure Models: Theory and Evidence

Dong-Hyun Ahn
Korea University and University of North Carolina

Robert F. Dittmar
Indiana University

A. Ronald Gallant
University of North Carolina

This article theoretically explores the characteristics underpinning quadratic term structure models (QTSMs), which designate the yield on a bond as a quadratic function of underlying state variables. We develop a comprehensive QTSM, which is maximally flexible and thus encompasses the features of several diverse models including the double square-root model of Longstaff (1989), the univariate quadratic model of Beaglehole and Tenney (1992), and the squared-autoregressive-independent-variable nominal term structure (SAINTS) model of Constantinides (1992). We document a complete classification of admissibility and empirical identification for the QTSM, and demonstrate that the QTSM can overcome limitations inherent in affine term structure models (ATSMs). Using the efficient method of moments of Gallant and Tauchen (1996), we test the empirical performance of the model in determining bond prices and compare the performance to the ATSMs. The results of the goodness-of-fit tests suggest that the QTSMs outperform the ATSMs in explaining historical bond price behavior in the United States.

Arguably the most popular state-of-the-art term structure models are affine term structure models (ATSMs), which designate the yield or log bond price as an affine function of the underlying state variables. A sequence of ATSMs including the ground breaking studies of Vasicek (1977) and Cox, Ingersoll,

---

This article has benefited considerably from the comments and suggestions of the editor, Bernard Dumas, and two anonymous referees. Special thanks to Martin Gruber, Yacov Amihud, Steve Brown, Jacob Boudoukh, Matthew Richardson, and Robert Whitelaw for their encouragement and helpful comments. We are also indebted to Torben Andersen, Robert Bliss, Jeffrey Busse, Amy Dittmar, Silverio Foresi, Bin Gao, Robert Jarrow, Jun Liu, Chang-Hyun Yun, and seminar participants at Aju University, Columbia University, Korea University, Korean Finance Institution, NYU, Princeton University, Rice University, University of North Carolina, Yonsei University, the 2001 American Finance Association Meetings, the 2000 Korea Futures Association Meetings, and the Hotelling Triangle Econometrics Conference for their helpful comments and discussions. Any remaining errors are ours alone. A previous version of this article appeared under the title "Extended Squared-Autoregressive-Independent-Variable Nominal Term Structure Model: Theory and Evidence." This article is, particularly in its theoretical parts, a significant extension of Chapter 3, "Generalized SAINTS Model" (1995) of the first author's dissertation and its precedent working paper at NYU. The first author greatly acknowledges financial support from Korea University, UNC, and the Com2Mac Center at POSTECH. Address correspondence to Prof. Dong-Hyun Ahn, Kenan-Flagler Business School, University of North Carolina, McColl Building, Chapel Hill, NC 27599-3490, or e-mail: ahnd@bschool.unc.edu.

The Review of Financial Studies Spring 2002 Vol. 15, No. 1, pp. 243–288

and Ross (1985b) (CIR hereafter) has been developed, and Duffie and Kan (1996) clarify the primitive assumptions underlying this framework. There have been three major innovations in the evolution of ATSMs. The first innovation is the specification of a single state variable diffusion process driving ATSMs which can better explain the empirical stochastic process of the short rate or its volatility.[1] The second innovation is the extension of single state variable ATSMs to their counterparts with *orthogonal* multiple state variables. This innovation is motivated by empirical evidence which suggests that single-factor ATSMs are unable to explain the dynamics of the U.S. term structure, and is developed in two different branches. The first branch directly extends the single-factor representation of the short rate by introducing a stochastic central-tendency factor and/or stochastic volatility see Andersen and Lund (1997), Balduzzi et al. (1996), Chen (1996), and Jegadeesh and Pennachi (1996). The second branch specifies the short rate as an addition of several state variables see Chen and Scott (1992), Longstaff and Schwartz (1992), Sun (1992), Pearson and Sun (1994), and Knez, Litterman, and Scheinkman (1996). The advantage of the first branch is that it provides an economic interpretation for the underlying state variables. The second branch results in bond prices that are simply the product of single-factor bond prices due to the additive property of the state variables, facilitating the models' empirical analysis. Dai and Singleton (2000) show that the first branch of models can be suitably represented as special cases of the second branch after reparameterization when correlations among state variables are allowed. The final innovation is the extension of multifactor dynamic models through the incorporation of nontrivial correlations among the state variables, which is again motivated by empirical concerns. This theoretical extension is first introduced in Langetieg (1980) in Gaussian models and pioneered in Duffie and Kan (1996) and its importance in an empirical context is examined in Dai and Singleton. In addition, Dai and Singleton characterize the admissibility of ATSMs and explore a maximally flexible ATSM that empirically nests all other ATSMs as its subfamily.

Despite the above-mentioned innovations accumulated in the development of ATSMs and their relatively promising empirical performance [Dai and Singleton (2000)], there are several good reasons for researchers to consider term structure models that are not members of the ATSM family. First, as documented in Dai and Singleton, ATSMs have a theoretical drawback which hampers their empirical performance. The form of ATSMs requires a tradeoff between the structure of bond price volatilities and admissible nonzero

[1] The Vasicek model, which is based on the Ornstein–Uhlenbeck process (continuous-time AR(1) process), has a drawback in that it generates homoscedastic volatility of the short rate. The CIR model incorporates the property of heteroscedastic volatility of the short rate, where the volatility is a function of the level of the short rate. Pearson and Sun (1994) extends the CIR model by introducing a positive lower bound on the short rate.

conditional correlations of the state variables.[2] Let $A_m(n)$ denote an ATSM with m state variables with square-root processes and $n-m$ Gaussian factors, following the notation of Dai and Singleton. Admissibility of an ATSM requires nonnegative correlations among the m square-root factors. As such, an increase in m limits the flexibility of the ATSM in specifying conditional/unconditional correlations while giving more flexibility in specifying heteroscedastic volatility. Therefore we expect that the goodness-of-fit of ATSMs may be weakened in settings where state variables have pronounced conditional volatility and are simultaneously strongly negatively correlated. In related evidence, Duffee (2000) finds that ATSMs forecast future yield changes poorly; a martingale provides better yield forecasts. A second and related issue is that the results of Dai and Singleton suggest that there may be some omitted *nonlinearity* in the ATSMs since the pricing errors of the ATSMs are sensitive to the magnitude of the slope of the (swap) yield curve and highly persistent. Finally, the only ATSM that ensures a strictly positive nominal interest rate is $A_m(m)$, in which all state variables are square-root factors. Therefore ATSMs cannot simultaneously allow for negative correlations among the state variables and guarantee positivity of the nominal interest rate.

Since the ATSMs that Dai and Singleton (2000) examine are maximally flexible, the aforementioned potential drawbacks of ATSMs provide a motivation for the development of a nonaffine family of term structure models. Compared to the ATSMs, nonaffine term structure models have been relatively slow to develop. This family of models can be broadly classified into two subgroups. The first group includes the double square-root model of Longstaff (1989), the multivariate quadratic model of Beaglehole and Tenney (1991), the univariate quadratic model of Beaglehole and Tenney (1992), the squared-autoregressive-independent-variable nominal term structure (SAINTS) model of Constantinides (1992), the quadratic model of Karoui, Myneni, and Viswanathan (1992), and the generalized SAINTS model of Ahn (1995). A heuristic sketch demonstrates that these models have some structural similarities in that the state variables are characterized as Gaussian diffusions and the instantaneous interest rate is represented as a quadratic function of the state variables. Despite these similarities, no rigorous study has formally clarified the relationships and differences among these models. The second group is the nonaffine model developed by Ahn and Gao (1999). This model is based on state variables with inverted square-root diffusions, and is clearly distinguished from the first group of models.

[2] To be precise, Dai and Singleton (2000) discuss a trade-off between the *conditional variance* of state variables and the admissible structure of the *correlation matrix* for the state variables. However, given the unobservability of the state variables and an affine functional relationship between state variables and yields (including the short rate), of central importance is the time-varying volatilities of bond returns and the short rate rather than the state variables themselves.

This article theoretically investigates characteristics underpinning the first group of models, which we refer to as quadratic term structure models (QTSMs). Specifically, we develop a comprehensive QTSM, which is maximally flexible and thus encompasses all features of the diverse models mentioned above. This full-fledged QTSM has the potential to overcome the aforementioned limitations of ATSMs. We demonstrate that the QTSM maintains admissibility without sacrificing flexibility in modeling heteroscedastic volatility and negative correlation among factors. This feature of the QTSM results from the combination of Gaussian state variables and a quadratic relationship between the state variables and the yields (including the short rate). In addition, QTSMs belong to a family of nonaffine term structure models, and thus they have the potential to capture omitted nonlinearities documented in Dai and Singleton (2000). Finally, because of the quadratic functional form, QTSMs allow for strictly positive nominal interest rates without imposing restrictions on the correlation structure of state variables. As such, QTSMs accommodate characteristics that can potentially overcome the shortcomings of ATSMs.

We also formally explore how the all-encompassing QTSM can nest all other QTSMs as special cases. In particular, the method of derivation and description of the SAINTS model prevents direct comparison between the SAINTS model and alternatives, and thus makes it difficult to infer which of its features lead to superior or inferior empirical performance relative to the alternatives. With an invariant transformation, we demonstrate that, with reparametrization and certain restrictions, our all-encompassing QTSM can be reduced to the SAINTS. As a by-product of this analysis, we identify some exogenous restrictions on the market prices of factor risks imposed by the SAINTS model. These restrictions are inherited from a direct specification of the stochastic discount factor and, as a result, there is no economic reasoning behind them.

It is surprising that despite a decade of history of QTSMs, little rigorous empirical study of any subfamily of QTSMs has been undertaken. An empirical implementation of QTSMs is complicated by the need to estimate the parameters of unobservable stochastic processes. Unlike ATSMs, even in the single state variable case, the short rate is not a sufficient statistic for the term structure since the short rate and yields are quadratic functions of an unobserved state variable. Furthermore, since the models are specified in the continuous time domain, the estimation method must address issues of discretization bias [see, e.g., Aït-Sahalia (1996a)]. These issues have been a major hindrance to empirical implementation of a family of QTSMs. The only empirical study of a family of QTSMs of which we are aware is Lu (1999), which computes nonlinear filter estimates of a two-factor SAINTS model using Kitagawa (1987), and compares its goodness-of-fit with that of a two-factor orthogonal CIR model.

We use the efficient method of moments (EMM) of Gallant and Tauchen (1996) to estimate a wide variety of QTSMs. The EMM is a suitable estimation scheme for QTSMs since it can overcome the aforementioned difficulties surrounding their empirical implementation.[3] Following Dai and Singleton (2000), we simultaneously use time-series data on short- and long-term Treasury bond yields to explore QTSMs empirical properties. We investigate four different parameterizations of QTSMs: the full-fledged QTSM (QTSM1), the QTSM with orthogonal state variables but with interactions in determination of the short rate (QTSM2), the QTSM with orthogonal state variables and without interactions in determination of the short rate (QTSM3), and finally the SAINTS model (QTSM4). This classification of QTSMs is informative since each model is nested in the next more-flexible version (i.e., QTSM$i \subset$ QTSMj for $i > j$). This hierarchy lets us explore the sources of improvements in the goodness-of-fit of QTSMs.

Our specification tests indicate that the QTSM class of models provides a good description of the dynamics of zero-coupon bond yields. We find that the restrictions imposed by the SAINTS model of Constantinides (1992) result in strong rejection of the QTSM and that relaxing these restrictions dramatically improves the fit of the quadratic class of models. When we allow for correlation among the state variables in the full-fledged model, QTSM1, we find that the performance is improved further, and the QTSM provides a good fit for term structure dynamics. In contrast, our specification test results suggest that the maximally flexible ATSM investigated in Dai and Singleton (2000) cannot fit these data even as well as the orthogonal QTSM, QTSM3, despite incorporating correlations among the state variables.

The remainder of this article is organized as follows. In Section 2 we provide a general characterization of QTSMs, describing the framework for the model, nested cases, and a general equilibrium that supports QTSMs. In Section 3 we explore the canonical form of the model, which allows us to implement the model empirically. Section 4 provides a discussion of the data and EMM methodology that we use for examining the fit of the the QTSM. The empirical results of the EMM estimation and further measurement of the model's fit are provided in Section 5. We make concluding remarks in Section 6.

1. A Characterization of QTSMs

The economy is represented by the augmented filtered probability space $(\Omega, F, \mathscr{F}, P)$, where filtration $\mathscr{F} = \{\mathscr{F}_t\}_{0 \leq t \leq \mathscr{T}}$. We first assert the existence

[3] An alternative estimation scheme that accommodates both the unobservability of the state variables and the nonlinearity of the yields in the factors is the extended Kalman filter, which is employed, for example, in Claessens and Pennachi (1996). However, in general, the extended Kalman filter suffers from approximation error, which requires simulation-based correction. This is an indirect inference method [Gourieroux, Monfort, and Renault (1993)], which is equivalent to the EMM. However, the EMM is known to have computational advantages.

of a positive state-price density process, $M(t)$, which defines the canonical valuation equation:

$$x(t) = E_t^P\left[\frac{M(T)}{M(t)}x(T)\right], \tag{1}$$

where $x(t)$ is the price of an asset, $x(t,w): [0,\infty)\times\Omega \to \Re^+$, and $E_t^P[\cdot]$ denotes the expectation conditional on the information at time t, $\mathscr{F}_t$ under the physical probability measure P. We refer to $M(t,T) \triangleq \frac{M(T)}{M(t)}$ as the stochastic discount factor, which discounts payoffs at time T into time t value under the stochastic economy.

As shown by Harrison and Kreps (1979) and Harrison and Pliska (1981), under the assumption of a complete market, there is a unique equivalent martingale measure Q under which all money market scaled asset prices follow a martingale:

$$\frac{x(t)}{B(t)} = E_t^P\left[\frac{dQ(t,T)}{dP(t,T)}\frac{x(T)}{B(T)}\right] \triangleq E_t^P\left[\mathscr{N}(t,T)\frac{x(T)}{B(T)}\right] = E^Q\left[\frac{x(T)}{B(T)}\right], \tag{2}$$

where $B(T)$ denotes a money market account and $B(t) = \exp(\int_0^t r(s)\,ds)$, where $r(s)$ denotes the locally riskless instantaneous rate at time s. $\mathscr{N}(t,T) = (dQ(t,T)/dP(t,T))$ is called the Radon–Nikodym derivative in the literature, which is equivalent to the conditional stochastic discount factor, $M(t,T)$ when $r(s) = 0\ \forall\ s \in [0,\mathscr{T})$. Given the uniqueness of the stochastic discount factor, the equivalence of Equations (1) and (2) yields the relationship between the stochastic discount factor and the Radon–Nikodym derivative:

$$M(t,T) = \left(\frac{B(t)}{B(T)}\right)\mathscr{N}(t,T) = \left[\exp\left(-\int_t^T r(s)\,ds\right)\right]\mathscr{N}(t,T).$$

We assume that $x(T)$ is the *nominal* payoff of an asset, which results in $M(t,T)$ as the corresponding nominal stochastic discount factor. Constantinides (1992) demonstrates that the nominal stochastic discount factor is the product of the inverse of the gross inflation rate and the real stochastic discount factor.

Following Hansen and Richard (1987), we directly explore the stochastic process of the nominal stochastic discount factor, $M(t,T)$. This pricing kernel approach is popular in the existing term structure literature [see Constantinides (1992), Ahn and Gao (1999), and Dai and Singleton (2000)]. As is shown by Harrison and Kreps (1979), there always exists an equilibrium that supports any admissible stochastic discount factor. As such, we will also demonstrate a general equilibrium that supports the prespecified diffusion process of the stochastic discount factor in Appendix C.

1.1 QTSMs

In this section we establish an N-factor QTSM by directly specifying the time-series process of the nominal stochastic discount factor. This pricing kernel approach hinges on the following three assumptions regarding the stochastic differential equation (SDE) of the stochastic discount factor and the $N \times 1$ vector of state variables $Y(t)$.

Assumption 1. We represent the time-series process of $M(t)$ as the SDE

$$\begin{aligned}\frac{dM(t)}{M(t)} &= -r(t)\,dt + 1'_N \text{diag}[\eta_{0i} + \eta'_{1i}Y(t)]_N\, dw_N(t)\\ &= -r(t)\,dt + 1'_N\big[\big(\eta_0 + \eta_1 Y(t)\big) \circ dw_N(t)\big],\end{aligned} \tag{3}$$

where

$$\eta_0 = (\eta_{01}\eta_{02}\cdots\eta_{0N})'\eta_1 = (\eta_{11}\eta_{12}\cdots\eta_{1N})',$$

and $\circ$ is a Hadamard product, an element by element multiplication. $w_N(t)$ is an N-dimensional vector of standard Wiener processes which are mutually independent.[4]

Assumption 1 states that the diffusion in Equation (3) is represented as an affine function of the state variables. This specification is unique in the sense that the diffusion is determined by constants and the level of state variables. Notice that the drift in Equation (3) is $-r(t)$, which stems from the martingale property of the stochastic discount factor $M(t)$ [see Harrison and Kreps (1979) and Cox, Ingersoll, and Ross (1985a)].

Assumption 2. The nominal instantaneous interest rate is a quadratic function of the state variables:

$$r(t) = \alpha + \beta' Y(t) + Y(t)'\Psi Y(t), \tag{4}$$

where α is a constant, β is an N-dimensional vector, and Ψ is an $N \times N$ matrix of constants. We assume that $\alpha - \frac{1}{4}\beta'\Psi^{-1}\beta \geq 0_N$, and Ψ is a positive semidefinite matrix.

As such, the nominal interest rate is a generalized positive semidefinite quadratic form. This form for the nominal interest rate is the property which designates the model QTSMs, which are clearly distinguishable from ATSMs. The sign restrictions on the parameters are required to ensure the nonnegativity of the nominal interest rate. Since Ψ is positive semidefinite, we obtain the lower bound on the short rate, $\alpha - \frac{1}{4}\beta'\Psi^{-1}\beta$ when $Y(t) = -\frac{1}{2}\Psi^{-1}\beta$.[5]

---

[4] $\text{diag}[x_i]_N$ denotes an *N-dimensional* diagonal matrix with diagonal elements of x_is $(i = 1, 2, \ldots, N)$.

[5] The lower bound on the short rate, $\alpha - \frac{1}{4}\beta'\Psi^{-1}$, can be strictly positive, which may be a minimum value of the interest rate that monetary authority can allow for to accelerate economic growth.

This is one of the advantages of QTSMs since ATSMs cannot, in general, guarantee this desirable property. Only in the case of $A_n(n)$ models [(potentially correlated) multifactor CIR models] can the positivity of the short rate be assured.

Assumption 3. The SDEs of the state variables $Y(t)$ are characterized as multivariate Gaussian processes with mean reverting properties:

$$dY(t) = [\mu + \xi Y(t)]\,dt + \Sigma dz_N(t),$$

where μ is an N-dimensional vector of constants, ξ and Σ are N-dimensional square matrices. We assume that ξ is "diagonalizable" and has negative real components of eigenvalues. $z_N(t)$ is an N-dimensional vector of standard Wiener processes that are mutually independent. The correlation matrix between $dw_N(t)$ and $dz_N(t)$, $\text{cov}_t(dw_N(t),\ dz_N(t))$ is denoted by Υ, an N-dimensional square matrix of constants.

The time-series process of the state variables is represented as a Gaussian process, which is characterized by steady-state long-term means of $-\xi^{-1}\mu$, a mean response matrix of $-\xi$, and a constant instantaneous covariance matrix $\Sigma\Sigma'$. ξ is assumed to be diagonalizable and to have negative eigenvalues in order to ensure the stationarity of the state variables.[6]

The SDEs of the state variables characterize the corresponding transition densities and marginal densities of $Y(t)$. Define U as the matrix of N eigenvectors and Λ as the diagonal matrix of eigenvalues,

$$U \triangleq [u_1\ \ u_2\ \ \cdots\ \ u_N] \text{ and } \Lambda \triangleq \text{diag}[\lambda_i]_N.$$

$U^{-1}\xi U = \Lambda$, since the diagonalizability of ξ ensures the linear independence of the eigenvectors. Then the transition densities for the state variables are represented as multivariate Gaussian densities (see Appendix A for a proof):

$$Y(t+\tau)|Y(t) \sim \text{MVN}_N(E[Y(t+\tau)|Y(t)],\ \ \text{var}[Y(t+\tau|Y(t)]), \tag{5}$$

where

$$E[Y(t+\tau)|Y(t)] = U\Lambda^{-1}[\Phi(\tau) - I_N]U^{-1}\mu + U\Phi(\tau)U^{-1}Y(t) \tag{6}$$

$$\text{var}[Y(t+\tau)|Y(t)] = U\left[\frac{v_{ij}(\exp((\lambda_i+\lambda_j)\tau)-1)}{\lambda_i+\lambda_j}\right]_{NN} U' \tag{7}$$

$$\Phi(\tau) \triangleq \text{diag}[\exp(\lambda_i\tau)]_N \tag{8}$$

$$\mathcal{V} \triangleq [v_{ij}]_N = U^{-1}\Sigma\Sigma' U'^{-1}. \tag{9}$$

[6] As is shown by Beaglehole and Tenney (1991), if all of the eigenvalues of ξ are negative and real, the conditional expectation exponentially decays toward the stable point. If ξ has negative complex eigenvalues, the conditional expectation of the state variables exhibit oscillatory decaying behavior toward the stable point. In contrast, in the case where a single eigenvalue of ξ has a positive real component (either real or complex), the state variables are nonstationary, either oscillating or drifting to some infinite value.

Note that there are two channels through which interdependencies among state variables are determined: (i) the off-diagonal terms of ξ which determine feedbacks in the conditional mean, and (ii) off-diagonal terms of Σ (along with ξ) which characterize the conditional covariances of the state variables. Therefore, when ξ and Σ are diagonal matrices, the diagonal is $\mathcal{V}$, which yields a diagonal covariance matrix.[7] Provided the admissibility conditions in Assumption 3 are satisfied, the steady-state distribution of $Y(t)$ is multivariate Gaussian with the following mean vector and covariance matrix:

$$E[Y(t)] = -\xi^{-1}\mu, \quad \text{var}[Y(t)] = U\left[-\frac{v_{ij}}{\lambda_i + \lambda_j}\right]U'.$$

We next turn to the distribution of the interest rate dictated by the QTSM. Appendix B derives the conditional and unconditional distributions, which can be represented as an infinite mixture of noncentral χ^2 distributions of the form[8]

$$\Pr[r = \alpha + Y(t)'\Psi Y(t) \le r_0] = \sum_{j=0}^{\infty} e_j \left[\chi^2_{N+2j}\left(\sum_{j=1}^{N}\omega_j^2\right) \le \frac{r_0 - \alpha}{\varepsilon}\right], \qquad (10)$$

where e_j, ω_j, and ε are defined in Appendix B. Appendix B also derives the first and second moments of the interest rates. This distribution is reduced to a noncentral χ^2 distribution only if the state variables $Y(t)$ are orthogonal. Therefore the functional form of the QTSM drives the distribution of the interest rate, which is different from either the SAINTS model or Beaglehole and Tenney (1992).

Based on these assumptions, we can solve for bond prices. Let $V(t, \tau)$ denote the nominal price at time t of a default-free bond that pays \$1 at time $T = t + \tau$. From the fundamental valuation equation [Equation (1)], we immediately know

$$V(t, \tau) = E_t^P[M(t, t+\tau)].$$

In order to solve this expectation, we write the SDE of the normalized bond price $Z(t, \tau) = V(t, \tau)/B(t)$,

$$\frac{dZ(t, \tau)}{Z(t, \tau)} = [a(t, \tau) - r(t)]\,dt + b(t, \tau)\,dz_N(t),$$

[7] When ξ is a diagonal matrix, $U = I_N$.

[8] For simplicity, we derive the distribution under the assumption of $\beta = 0_N$, which is one of the important conditions in the canonical form that we will discuss in Section 3.

and applying Ito's lemma leads to

$$a(t,\tau)=\left[\frac{1}{2}\mathrm{tr}\left(\Sigma\Sigma'\frac{\partial^2 V(t,\tau)}{\partial Y(t)\partial Y(t)'}\right)+\left(\frac{\partial V(t,\tau)}{\partial Y(t)}\right)'\right.$$
$$\left.\times[\mu+\xi Y(t)]+\frac{\partial V(t,\tau)}{\partial t}\right]\Big/V(t,\tau)$$
$$b(t,\tau)=\left(\frac{\frac{\partial V(t,\tau)}{\partial Y(t)}}{V(t,\tau)}\right)'\Sigma.$$

From Assumption 1, the product of the Radon–Nikodym derivative $\mathcal{N}(t,T)$ and the normalized bond price $Z(t,\tau)$ is written as

$$\frac{d[\mathcal{N}(t,t+\tau)Z(t,\tau)]}{\mathcal{N}(t,t+\tau)Z(t,\tau)}=[a(t,\tau)-r(t)+b(t,\tau)\Upsilon(\eta_0+\eta_1 Y(t)]\,dt$$
$$+b(t,\tau)\,dz_N(t)+\ell_N'[\eta_0+\eta_1 Y(t)]\circ dw_N(t).$$

Equation (2) asserts that $\mathcal{N}(t,t+\tau)Z(t,\tau)$ is a martingale, which leads to the expression for the excess return on the bond:

$$a(t,\tau)-r(t)=-b(t,\tau)\Upsilon[\eta_0+\eta_1 Y(t)],$$

and equivalently

$$\left[\frac{1}{2}\mathrm{tr}\left(\Sigma\Sigma'\frac{\partial^2 V(t,\tau)}{\partial Y(t)\partial Y(t)'}\right)+\frac{\partial V(t,\tau)}{\partial Y(t)'}[\mu+\xi Y(t)]+\frac{\partial V(t,\tau)}{\partial t}\right]\Big/V(t,\tau)$$
$$=r(t)-\left[\frac{\frac{\partial V(t,\tau)}{\partial Y(t)'}}{V(t,\tau)}\right]\Sigma\Upsilon[\eta_0+\eta_1 Y(t)]. \tag{11}$$

Equation (11) is a fundamental partial differential equation (PDE) for a bond price. Its left-hand side stands for the instantaneous expected return on the bond, which is derived from Ito's lemma. In contrast, the right-hand side expresses the instantaneous expected return as a sum of the instantaneous risk-free rate and the risk premium of the bond. In turn, the risk premium of the bond is a multiplication of two components. $\frac{\partial V(t,\tau)}{\partial Y(t)}/V(t,\tau)$ is a vector of sensitivities to the state variables, and $-\Sigma\Upsilon[\eta_0+\eta_1 Y(t)]$ represents the covariance between the state variables and the stochastic discount factor, $\mathrm{cov}_t(dY(t), dM(t)/dM(t))$, which is the market-wide price of factor risks. Since the stochastic discount factor is not observable, we are not able to separately identify Υ, η_0, and η_1. These parameter vectors or matrices, as well as Σ, are constant. This feature allows us to define new notation for the market price of risk

$$\delta_0\triangleq-\Sigma\Upsilon\eta_0,\qquad \delta_1\triangleq-\Sigma\Upsilon\eta_1,$$

which enables us to reexpress the total market price of risk as $\delta_0 + \delta_1 Y(t)$. By rearranging the terms of Equation (11) we get the implied risk-neutral valuation scheme:

$$\left[\frac{1}{2}\mathrm{tr}\left(\Sigma\Sigma' \frac{\partial^2 V(t,\tau)}{\partial Y(t)\partial Y(t)'}\right) + \frac{\partial V(t,\tau)}{\partial Y(t)'} \right.$$
$$\left. \times [(\mu - \delta_0) + (\xi - \delta_1)Y(t)] + \frac{\partial V(t,\tau)}{\partial t}\right] \Big/ V(t,\tau) = r(t). \quad (12)$$

The Girsanov theorem states that this PDE is consistent with the valuation scheme under the risk-neutral measure, or Q measure, under which the SDE of the state variables is written as

$$\begin{aligned} dY(t) &= [\mu - \delta_0 + (\xi - \delta_1)Y(t)]\,dt + \Sigma d\tilde{z}_N(t) \\ &\triangleq [\tilde{\mu} + \tilde{\xi}Y(t)]\,dt + \Sigma d\tilde{z}_N(t), \end{aligned}$$

where $\tilde{z}_N(t) = z_N(t) + \int_0^t \Sigma^{-1}(\delta_0 + \delta_1 Y(s))\,ds$. δ_0 and δ_1 adjust the constant vector and the response matrix of the drift of the state vector SDE.

We turn next to the pricing of interest rate contingent claims. Provided the aforementioned assumptions are satisfied, $V(t,\tau)$ is a solution for the fundamental PDE [Equation (12)] given the terminal condition, $V(t,0) = 1$. The solution is a exponential quadratic function of the state vector

$$V(t,\tau) = \exp[A(\tau) + B(\tau)'Y(t) + Y(t)'C(\tau)Y(t)],$$

where $A(\tau)$, $B(\tau)$, and $C(\tau)$ satisfy the ordinary differential equations (ODEs),

$$\begin{aligned} \frac{dC(\tau)}{d\tau} &= 2C(\tau)\Sigma\Sigma'C(\tau) + (C(\tau)(\xi - \delta_1) + (\xi - \delta_1)'C(\tau)) - \Psi \\ \frac{dB(\tau)}{d\tau} &= 2C(\tau)\Sigma\Sigma'B(\tau) + (\xi - \delta_1)'B(\tau) + 2C(\tau)(\mu - \delta_0) - \beta \\ \frac{dA(\tau)}{d\tau} &= \mathrm{tr}[\Sigma\Sigma'C(\tau)] + \frac{1}{2}B(\tau)'\Sigma\Sigma'B(\tau) + B(\tau)'(\mu - \delta_0) - \alpha, \end{aligned}$$

with the initial conditions $A(0) = 0$, $B(0) = 0_N$, and $C(0) = 0_{N\times N}$. These ODEs can be easily solved numerically. The yield-to-maturity, $yt(t,\tau)$, is defined as $-(\ln V(t,\tau))/\tau$,

$$yt(t,\tau) = \frac{1}{\tau}[-A(\tau) - B(\tau)'Y(t) - Y(t)'C(\tau)Y(t)].$$

The yield is a quadratic function of the state variables. Therefore the QTSMs are nonlinear models, a feature which is particularly attractive considering the nonlinearity strongly evidenced in Ahn and Gao (1999). Even in the case of a model with a single state variable ($N = 1$), the same level of the interest rate can generate different yield curves, depending on the sign of the state variable $Y(t)$. As such, the nominal interest rate is not a sufficient statistic for the underlying state variable in a single state variable case. This is a feature that distinguishes the QTSMs from the single-factor ATSMs, in which the interest rate is always a sufficient statistic for the underlying factors. The distribution of the yield is characterized as an infinite mixture of noncentral χ^2 distributions, which is similar to Equation (10).

We derive the QTSM based on an exogenous assumption regarding the SDE of the stochastic discount factor $M(t, T)$. In Appendix C, we analyze a general equilibrium that supports the QTSM.[9] One point to note is that this general equilibrium model does not elaborate on inflation. Therefore the model is based on a real economy rather than a nominal economy.

1.2 The nested models

The QTSM that is developed in this article is similar to the model of Beaglehole and Tenney (1991), but with significant refinements. Their model resides in a risk-neutral economy, and thus does not specify the market price of risk. Second, their solutions for bond prices require numerical integration in addition to the solution of a system of equations. In contrast, our framework simply involves the solution of a system of ODEs.[10] The QTSM can be reduced to a variety of existing models as its special cases. These models include the double square-root model of Longstaff (1989), the univariate quadratic model of Beaglehole and Tenney (1992), and the SAINTS model of Constantinides (1992). In addition, the QTSM also nests a specific version of CIR (1985b).

1.2.1 The univariate quadratic model [Beaglehole and Tenny (1992)]. This model was originally developed as a single-factor model. However, under the assumption of orthogonality of the state variables, the model can be suitably extended to a multiple state variable model. The required restrictions which reduce the QTSM are

$$\alpha = 0, \beta = \delta_0 = 0_N, \Psi,\ \xi,\ \Sigma,\ \delta_1 = \text{diagonal matrix.}$$

[9] Since the general equilibrium requires a wide variety of assumptions, such as the preferences of economic agents, production technologies, and budget constraints, there may be multiple equilibria which support a given prespecified stochastic discount factor. Therefore the general equilibrium that we demonstrate may be one of many potential equilibria.

[10] Beaglehole and Tenney (1991) derive solutions based on Green's functions. Their solutions are, however, impaired by errors. In their solution for a zero-coupon bond [Equation (16)], the coefficients of the state vector [which correspond to $B(\tau)$ and $C(\tau)$ in our article] are constant rather than a function of time to maturity.

1.2.2 The double square-root model [Longstaff (1989)]. This model was also developed as a single state variable model, which can be extended to a multiple state variable version under the assumption of orthogonality of the state variables. The solution of this model violates the viability condition. The restrictions below yield an admissible model as suggested by Beaglehole and Tenney (1992). The state variable $Y(t)$ does not have zero as a reflecting bound, but is unrestricted. The key feature of this model is that the state variables are not mean reverting:[11]

$$\alpha = 0 \quad \beta = \delta_0 = 0_N, \quad \Psi, \Sigma = \text{diagonal matrix}, \quad \mu \neq 0_N, \quad \xi = \delta_1 = 0_{N\times N}.$$

1.2.3 A special version of the CIR model. The (orthogonal) CIR model is an ATSM and thus does not seem to be consistent with the QTSM. However, a QTSM with certain restrictions can be consistent with a particular version of the CIR model. Put differently, these two apparently heterogeneous models can coincide under certain conditions. The restrictions required are

$$\alpha = 0, \quad \beta = 0_N, \quad \Psi, \xi, \Sigma, \delta_1 = \text{diagonal matrix}, \quad \mu = \delta_0 = 0_N \text{ (or } \mu = \delta_0).$$

This model corresponds to the multifactor CIR model wherein the SDE of the state variable $Y_i(t)^c \ \forall \ i = 1, 2, \ldots, N$ is represented as

$$dY_i(t)^c = \left[\frac{\sigma_i^{c2}}{4} + \xi_i^c Y(t)^c\right] dt + \sigma_i^c \sqrt{Y_i(t)^c}\, dz_i(t). \tag{13}$$

This equivalence is easily illustrated by the fact that $Y_i(t)^c = Y_i(t)^2$. That is, the restricted CIR model is a reparameterized model with a quadratic transformation of the state variables.[12] It is obvious to see that under the orthogonality of the state variables, the conditional densities of the QTSM as well as the CIR model are noncentral chi-squared distributions. They achieve these chi-squared distributions in different ways and coincide only when $Y_i(t)^c$ $(= Y_i(t)^2)$ follows a square-root process.

1.2.4 The SAINTS model [Constantinides (1992)]. It is not obvious to see how the QTSM nests the SAINTS model. The SAINTS model is based

[11] However, the interest rate exhibits mean reversion because of its quadratic form.

[12] Notice that Equation (13) does not satisfy the Feller condition, and thus zero is accessible. However, the zero boundary of $Y_i(t)$ is reflecting or temporarily sticking since $Y_i(t)^c$ is a quadratic function of a Gaussian process, that is, $Y_i(t)^2$, and the CIR solution is still valid. This case corresponds to the "unrestricted equilibrium" in Longstaff (1992), wherein the process of the state variables, $Y_i(t)^c$s, return immediately to positive values if they reach zero. When the state variable is zero, its process is *locally* deterministic, $dY_i(t)^c = (\sigma_i^{c2}/4)\, dt$.

on the following assumptions regarding the SDE of state vector $X(t)$ and the stochastic discount factor:

$$dX(t) = -\mathcal{K}X(t)\,dt + S\,dz_N(t) \tag{14}$$

$$M(t) = \exp\left[-ht + \sum_{i=1}^{N}(X_i(t)-c_i)^2\right], \tag{15}$$

where $\mathcal{K}$ and S are assumed to be diagonal matrices of constants, which results in orthogonality among the state variables. The state variables have trivial long-term means in Equation (14). Of greater interest, the stochastic discount factor $M(t)$ is represented as an exponential quadratic function of the state variables. This specification for the stochastic discount factor jointly determines the nominal interest rate and also the diffusion of the stochastic discount factor (and eventually the market price of risk). Applying Ito's lemma leads to the corresponding SDE for the stochastic discount factor:

$$\frac{dM(t)}{M(t)} = -\left[\sum_{i=1}^{N}\left\{2\left(\mathcal{K}_{ii}-S_{ii}^2\right)X_i(t)^2 + 2c_i\left(2S_{ii}^2-\mathcal{K}_{ii}\right)X_i(t) - S_{ii}^2 - 2c_i^2S_{ii}^2\right\} + h\right]dt$$
$$+2\sum_{i}^{N} S_{ii}(X_i(t)-c_i)\,dz_i(t). \tag{16}$$

Comparing Equation (16) to Equations (3) and (4), we note that there is an isomorphism in the two models in the specification of the interest rate and the diffusion terms of the stochastic discount factor. To explore how the QTSM nests the SAINTS model, we need to conduct an invariant transformation (see Appendix D). The result reveals that the QTSM can be reduced to the SAINTS model by imposing the following restrictions on the market prices of risk:

$$\beta = 0_N, \quad \Psi = I_N, \quad \xi, \Sigma, \delta_1 = \text{diagonal matrix } \delta_{0i}$$
$$= \pm\mu_i\left(\frac{\xi_{ii} \pm \sqrt{\xi_{ii}^2 - 2\Sigma_{ii}^2}}{\sqrt{\xi_{ii}^2 - 2\Sigma_{ii}^2}}\right), \; \delta_{1ii} = -\xi_{ii} \mp \sqrt{\xi_{ii}^2 - 2\Sigma_{ii}^2}.$$

Therefore the market prices of risk are restricted to be very specific functions of the structural parameters.[13] That is, the risk premia are determined by the parameters governing the time-series evolution of the interest rates, ξ and Σ. Note that these matrices of parameters describe the stochastic process of the interest rate, while the δs, the risk premia, are the parameters controlling the cross-sectional relationship among bonds at a given point in

[13] Because the description of the SAINTS model imposes restrictions which yield nonlinear equations to solve for the δs, there are multiple solutions for the the δs. For each state variable there are two alternative forms of restrictions induced by the $\pm$ signs.

time. Put differently, the structural matrices determine the evolution of the economy under the physical measure P, whereas the δs convert this evolution to its counterpart under the risk-neutral measure Q. Since the structural parameters themselves govern the conversion of the probability measures, the SAINTS model may perform poorly in fitting the cross-section of bond yields relative to the comprehensive QTSM, which is a testable hypothesis. The issue is that these restrictions are not the outcomes of economic reasoning, but rather the results of an exogenous specification of the stochastic discount factor in Equation (15). Therefore the latitude of the model may be limited. The QTSM can overcome this drawback because the model does not impose any restrictions on the δs.

2. Canonical Form of the QTSM

Dai and Singleton (2000) demonstrate that the fully specified ATSM does not lend itself to specification analysis since not all parameters are empirically identifiable under the assumption of unobservability of the state variables $Y(t)$. A similar problem occurs in the QTSM. However, the identifiability conditions for the QTSM are much simpler since its state variables have a homoscedastic diffusion matrix. We define a canonical representation of the QTSM which lends itself to empirical implementation. This model is a maximally flexible model which can be reduced to a wide variety of subfamilies with appropriate restrictions on its parameters.

Definition 1. We define the canonical form of the QTSM by adding the following restrictions on the QTSM that we develop in Section 2,

$$\Psi = \begin{bmatrix} 1 & \Psi_{12} & \cdots & \Psi_{1N} \\ \Psi_{12} & 1 & \cdots & \Psi_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ \Psi_{1N} & \Psi_{2N} & \cdots & 1 \end{bmatrix},$$

a symmetric matrix with diagonal terms of 1s. $\mu \geq 0$. In addition, $\alpha > 0$, $\beta = 0_N$, and ξ and δ_1 are lower triangular matrices. Σ is a diagonal matrix.

The restrictions required to identify the QTSM are much simpler than those needed to identify the ATSM. Any equivalent model which is defined in Section 2 can be converted into the canonical form by an invariant transformation which is defined in Appendix E. The assumption of $\beta = 0_N$ is necessary to have μ identifiable. In addition, $\beta = 0_N$, together with positive semidefiniteness of Ψ, is a suitable way to ensure the positivity of the nominal interest rate. Notice that the lower bound on the interest rate is α in the canonical form. The diagonal terms of Ψ are assumed to be 1s in order to make the matrix robust to a rescaling linear transformation. Finally, the triangularity of ξ and the diagonality of Σ are necessary since the covariance matrix of the state variables requires that only one of ξ and Σ can be fully specified.[14] The symmetric covariance matrix is unique up to $(N^2+N)/2$ elements, and thus we assume that ξ is lower triangular. These restrictions are minimal normalizations for econometric identification of the QTSM. Notice that the proposed canonical form is only one of many alternative equivalent forms. For example, we can assume that ξ is diagonal where Σ is lower or upper triangular. Similarly, $\mu = 0_N$ can be paired with letting β be a vector of free parameters. All of these alternative forms designate the same specification of the QTSM.

We conclude this section by highlighting the difference between the canonical QTSM and a canonical ATSM suggested by Dai and Singleton (2000). First, the canonical QTSM ensures a positive interest rate. In contrast, $A_m(n)$ cannot ensure this property unless $m = n$, the (potentially correlated) CIR model. If $m < n$, there exists one or more Gaussian state variables which may take on negative values. Since the interest rate is an affine function, the interest rate may become negative in some states. In contrast, the state variables in the QTSM are all Gaussian, which may take on negative values. However, the quadratic relationship between the interest rate and the state variables ensures the nonnegativity of the interest rate under the assumption of $\alpha \geq 0$ and positive semidefiniteness of Ψ.

In addition, the interest rate and bond prices in the QTSM exhibit heteroscedastic conditional volatilities. Even though the state variables themselves do not exhibit this feature, the SDE of the interest rate is represented as

$$dr(t) = \left[\mathrm{tr}(\Sigma^2\Psi) + 2(\mu + \xi Y(t))'\Psi Y(t)\right] dt + 2Y(t)'\Psi\Sigma dz_N(t).$$

Thus the conditional variance of the interest rate is a linear function of the state variables, as in the CIR model. In contrast, in the $A_m(n)$, as emphasized in Dai and Singleton (2000), the m state variables determine the stochastic volatilities. This stochastic volatility is typically achieved at the cost of flexibility in the specification of conditional/unconditional correlations among the state variables. Again, this undesirable trade-off between the structure of conditional volatilities and admissible nonzero conditional correlations of the state variables results from the affine structure of the model. In contrast, the QTSM does not result in less flexibility in specifying conditional correlations among the state variables, since conditional volatilities are induced by the quadratic structure rather than the processes of the state variables. Thus the unconditional correlations among the state variables in the QTSM can be either positive or negative without hampering the flexibility of volatility specification.

[14] Therefore the assumption that ξ is diagonal and Σ is lower or upper triangular results in the same specification.

2.1 Three-factor models considered

To better understand the QTSM, we explore subfamily models of the QTSM. We are particularly interested in investigating the source of potential empirical improvement of the QTSM by constructing a hierarchy of models in terms of their general flexibility. Following Dai and Singleton (2000), and the evidence in Knez, Litterman, and Scheinkman (1996), we fix $N = 3$ since the empirical literature suggests that three factors are required to describe the term structure. We examine four alternative subfamilies of the QTSM.

2.1.1 QTSM1: maximally flexible model. The maximally flexible model is the aforementioned canonical model with $N = 3$. The model requires the estimation of α, three off-diagonal elements of Ψ, three elements of μ, six elements of ξ, three elements of Σ, three elements of δ_0, and six elements of δ_1. Thus the total number of parameters to be estimated is 25. This model embodies a fully specified covariance matrix of the state variables and allows for interactions among state variables in the determination of the nominal interest rate (i.e., the off-diagonal terms of ξ can be nontrivial).

2.1.2 QTSM2: orthogonal state variables and interactions. The underlying assumption of QTSM2 is that ξ and δ_1 are diagonal. This assumption results in orthogonal state variables under the P measure as well as the Q measure. However, Ψ is not diagonal, resulting in interactions in the determination of the nominal interest rate. Since ξ and δ_1 are diagonal, six parameters are set to zero in QTSM1. The number of parameters in QTSM2 is then 19.

2.1.3 QTSM3: orthogonal state variables and no interactions. We impose the additional restriction in this case that Ψ is diagonal, that is, I_3. Thus there are no interactions among the state variables in the determination of the interest rate, which results in 16 parameters. An important advantage is that QTSM3 allows for fully closed-form solutions for bond prices:[15]

$$V(t,\tau) = \exp(-\alpha\tau)\Pi_{i=1}^{N} \times \left[\sum_{i=1}^{N} A_i(\tau) + \sum_{i=1}^{N} B_i(\tau) Y_i(t) + \sum_{i=1}^{N} C_i(\tau) Y_i(t)^2\right], \tag{17}$$

[15] The proof is available upon request from the authors.

where

$$
\begin{aligned}
A_i(\tau) &= \left[-\left(\frac{\mu_i-\delta_{0i}}{\varsigma_i}\right)^2\tau\right] \\
&\quad + \left[\frac{(\mu_i-\delta_{0i})^2(\exp(\varsigma_i\tau)-1)[(-2(\xi_{ii}-\delta_{1ii})+\varsigma_i)(\exp(\varsigma_i\tau)-1)+2\varsigma_i]}{\varsigma^3[(-\xi_{ii}+\delta_{1ii}+\varsigma_i)(\exp(2\varsigma_i\tau)-1)+2\varsigma_i]}\right] \\
&\quad + \frac{1}{2}\ln\left[\frac{2\varsigma_i\exp(-(\xi_{ii}+\delta_{1ii})+\varsigma_i)}{(-\xi_{ii}+\delta_{1ii}+\varsigma_i)(\exp(2\varsigma_i\tau)-1)+2\varsigma_i}\right] \\
B_i(\tau) &= -\frac{2(\mu_i-\delta_{0i})(\exp(\varsigma_i\tau)-1)^2}{\varsigma_i[(-\xi_{ii}+\delta_{1ii}+\varsigma_i)(\exp(2\varsigma_i\tau)-1)+2\varsigma_i]} \\
C_i(\tau) &= -\frac{(\exp(2\varsigma_i\tau)-1)}{(-\xi_{ii}+\delta_{1ii}+\varsigma_i)(\exp(2\varsigma_i\tau)-1)+2\varsigma_i},
\end{aligned}
$$

where $\varsigma_i=\sqrt{(-\xi_{ii}+\delta_{1ii})^2+2\Sigma_{ii}^2}$. Since Ψ is orthogonal and thus I_3, the number of parameters is 16.

2.1.4 QTSM4: the SAINTS model. As shown in Section 2, the SAINTS model is based on orthogonal state variables and no interaction among the state variables in the determination of the interest rate. Therefore QTSM3 is reduced to the SAINTS model when we impose the restrictions on the market price of risk specified in Section 2. Since there are two alternative restrictions on the market price of risk associated with each factor, there could be six ($=2N$) different forms of aggregate restrictions consistent with the SAINTS model. We will investigate a particular combination of restrictions:

$$
\delta_{0i}=-\mu_i\left(\frac{\xi_{ii}-\sqrt{\xi_{ii}^2-2\Sigma_{ii}^2}}{\sqrt{\xi_{ii}^2-2\Sigma_{ii}^2}}\right),\ \delta_{1ii}=-\xi_{ii}\sqrt{\xi_{ii}^2-2\Sigma_{ii}^2}\ \forall\, i=1,2,\text{ and } 3.
$$

We choose these particular restrictions based on the calibration of the model for the unconditional yield curve. Since the market prices of risk are not free parameters, the total number of parameters is 10.

3. Data and Methods

3.1 Term structure data

In order to investigate the implications of the QTSM for the term structure of interest rates, we utilize the dataset of McCulloch and Kwon (1993). These data are sampled at a monthly frequency and cover the period December 1946–February 1991. Although this sample omits the most recent data, we view its advantages as superior to its disadvantages. Since the dataset accounts for coupon payments, we observe a zero-coupon term structure that is the object of interest for our analysis. Furthermore, since the interest rate period after 1991 has been relatively stable, we suggest that our analysis does not omit regimes in the data that are of particular importance to gauge the model's fit.

For the purposes of the analysis of the model's ability to fit the term structure of interest rates, we utilize three yields: the 3-month and 12-month Treasury-bill yields and the 10-year bond yield. These maturities are similar to those examined in comparable studies, for example, Dai and Singleton (2000). All of the yields are treated as such in estimation; many past studies have used the 3-month Treasury-bill yield as a proxy for the short rate. However, given the evidence in Chapman, Long, and Pearson (1999), which suggests that use of the 3-month Treasury-bill as a proxy for the short rate may induce bias in estimation, we explicitly treat the 3-month Treasury-bill yield as a bond yield in our empirical application. As these yields cover short-, intermediate-, and long-term bonds, we feel that they provide a reasonable description of the term structure of interest rates at a given point in time. The data are plotted in Figure 1, which shows that the sample period covers a wide range of interest rate regimes, from very low levels in the late 1940s and 1950s to the high-rate regime of the early 1980s. Thus the sample captures periods of relative stability in interest rates as well as periods punctuated by high volatility.

3.2 The efficient method of moments

As mentioned previously, one of the defining features of the QTSM is that even in the single-factor case, the short rate is not a sufficient statistic for risk in the economy. In the presence of multiple state variables, this issue becomes more important. As a result, estimation of the parameters of the model is complicated by the need to estimate the parameters of an unobserved stochastic process. Furthermore, since the model is expressed in continuous time, it is necessary to avoid issues of discretization bias [Aït-Sahalia (1996a, b)]. Recent econometric advances have allowed researchers to address both of these issues through the use of simulated method of moments techniques. We specifically employ the efficient method of moments [EMM; Gallant and Tauchen (1996)] to estimate the parameters of the QTSM. This methodology has been used to estimate parameters of the short rate diffusion in Andersen and Lund (1997) and to investigate ATSMs in Dai and Singleton (2000).[16]

The EMM procedure can be thought of as a two-step process. The first step is fitting a consistent estimator of the conditional density of the observable data. Designate this approximation to the density as

$$\hat{f}_K(y_t|x_{t-1},\theta) = \frac{f(x_{t-1},y_t|\theta)}{f(x_{t-1}|\theta)}, \tag{18}$$

where y_t denotes the current observation of the observed process, x_{t-1} denotes lags of the process, and θ denotes the K-dimensional parameter vector of the

[16] A detailed discussion of the method in these contexts can be found in Andersen and Lund (1997).

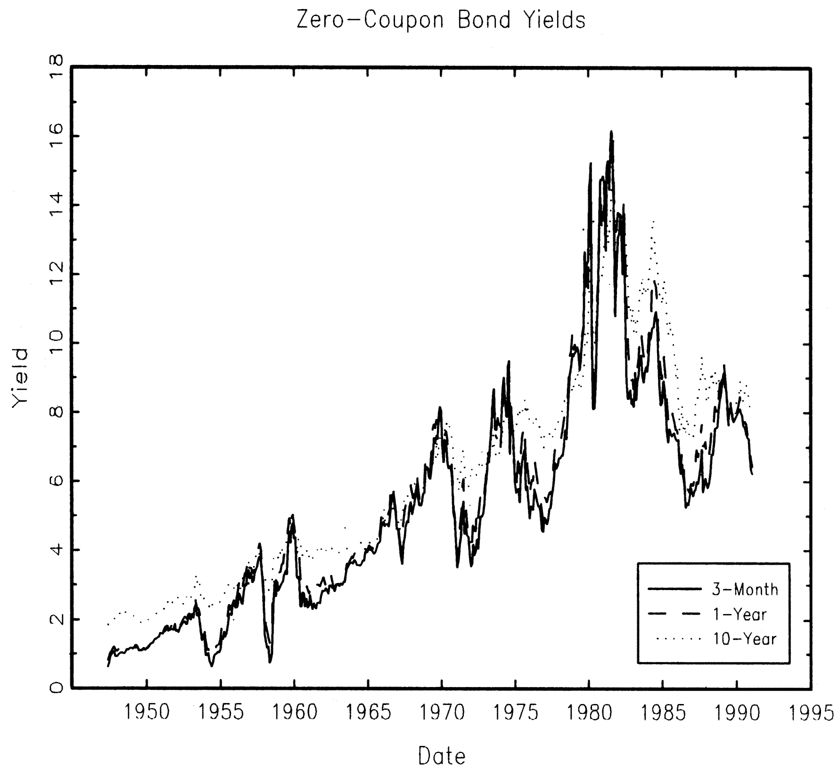

Figure 1
Treasury Yields
Plot of the 3-month, 1-year, and 10-year zero-coupon Treasury yield data over the period December 1946 February 1991. The data are sampled at a monthly frequency and are obtained from the McCulloch and Kwon (1993) dataset.

density approximation. We approximate this density using the seminonparametric (SNP) procedure of Gallant and Tauchen (1989). The procedure augments a Gaussian vector-autoregression (VAR) with the potential for ARCH innovations by a Hermite polynomial expansion to capture deviations from normality. Designating a demeaned transformation of y_t as $z_t = R^{-1}(y_t - \mu)$, the SNP approximation to the density is given by

$$\begin{aligned} h_K(z_t|x_{t-1}) &= \frac{f(z_t, x_{t-1})\phi(z_t)}{\int f(s, x_{t-1})\phi(s)\,ds} \\ f(z_t, x_{t-1}) &= \sum_{|\alpha|=0}^{K_z} \sum_{|\beta|=0}^{K_x} \left(a_{\alpha\beta} x_{t-1}^{\beta}\right) z_t^{\alpha} \qquad (19) \\ \phi(z_t) &\sim N(0, I). \end{aligned}$$

We fit an SNP model to the Treasury data using the procedure outlined in Gallant and Tauchen (1997). The authors suggest an upward-fitting strategy in which the parameters of parts of the SNP model are tuned to minimize the Schwartz (1978) criterion (BIC) and then are used as starting points for the fitting of the next part of the model. This method provides a fairly efficient way to fit the model. Our Schwartz-preferred fit is described by $\{L_\mu, L_r, K_z, I_z, K_x\} = \{1, 4, 4, 3, 0\}$. $L_\mu = 1$ implies that one lag of the data is sufficient to describe mean dynamics in the VAR, and $L_r = 4$ suggests that a fourth-order ARCH process describes the innovations to the process. $K_z = 4$ suggests that a fourth-order Hermite polynomial captures deviations from normality, and $I_z = 3$ indicates that the interaction terms in the orders of the polynomial are suppressed. Finally, $K_x = 0$ suggests that it is unnecessary to incorporate lags of the process in modeling the coefficients of the Hermite polynomial. This specification is quite similar to that of other SNP specifications in term structure studies. For example, Dai and Singleton (2000) find a specification of $\{L_\mu, L_r, K_z, K_x\} = \{1, 2, 4, 0\}$ describes a term structure of 6-month LIBOR, 2-year swap, and 10-year swap yields over the period 1987–1996. Our specification differs only in the ARCH term, which likely reflects our incorporation of an earlier period in our data sample.

The second step in the EMM process involves estimating a parameter vector for the term structure model. The procedure takes a set of initial starting values for the model and simulates a long set of data. In our case, we set the simulation length to $T = 50{,}000$. The SNP model is fit to the simulated data and the scores of the fitted model with respect to the SNP parameters are estimated. Designate the parameters of the structural model as ρ and the parameters of the SNP model as $\tilde{\theta}$. The scores of the fitted SNP model are used as moment conditions, $m'(\rho, \tilde{\theta})$, and the quadratic form

$$m'(\rho, \tilde{\theta})\widetilde{\mathcal{I}}^{-1} m(\rho, \tilde{\theta}) \tag{20}$$

is estimated, where $\widetilde{\mathcal{I}}^{-1}$ denotes the quasi-information matrix from quasi-maximum likelihood estimation of θ. The procedure is repeated until the quadratic form is minimized. Then a test of the specification of the SDE is formed through the test statistic

$$Tm'(\rho, \tilde{\theta})\widetilde{\mathcal{I}}^{-1}(\rho, \tilde{\theta}) \sim \chi^2_{K-J}, \tag{21}$$

where K denotes the dimension of θ and J denotes the dimension of ρ. The method uses all of the relevant moments of the conditional distribution and is therefore asymptotically as efficient as maximum likelihood, as shown in Gallant and Long (1997).

The final issue is the circumstances under which the market price of risk parameters, δ_0 and δ_1, can be identified. As argued by Dai and Singleton (2000), there are two sources of identification for the market prices of risk.

One source is a nonlinear mapping between the yield and the underlying Gaussian state variables. The other source is the assumption that the state variable follows a non-Gaussian process. Since the QTSM is a nonlinear model, there is a nonlinear mapping between zero-coupon yields and the underlying state variables. As such, the market prices of risk can be estimated.

4. Estimation of Term Structure Models

In this section we conduct tests of goodness-of-fit for the subclasses of the QTSM developed in this article. We repeat this assessment for two of the ATSMs investigated in Dai and Singleton (2000) and compare the ability of the QTSM to the ATSM class of models to fit term structure dynamics. We then perform further analysis on the models by examining their ability to match specific conditional moments of the data through the reprojection methodology described in Gallant and Tauchen (1998).

4.1 EMM specification tests

4.1.1 Quadratic term structure models. As discussed above, we examine four nested versions of the QTSM, designated QTSM1 (most general) through QTSM4 (most restrictive). Estimation results for the four models are presented in Table 1, which depicts parameter estimates and specification tests for each of the models discussed above. The first column presents results for QTSM4, the SAINTS model, the most restrictive model in our framework. The bottom rows of the table present χ^2 statistics for model fit and a z-statistic for model fit that is asymptotically standard normal and adjusted for degrees of freedom. The z-statistic for the SAINTS model is 59.515, suggesting a strong rejection of the overidentifying restrictions implied by the model.[17] The results of the estimation suggest that the restrictions on the prices of risk imposed by the model may significantly impact the model's ability to fit the yield curve. The restrictions are eased in QTSM3, and their impact on the model's fit is discussed below.

The second column presents results for QTSM3, the fully orthogonal QTSM, which allows for a closed-form expression for bond prices. Although the model is rejected by the data, its performance greatly improves on that of QTSM4, as evidenced by the z-statistic of 13.396. The main difference between this model and QTSM4 is the easing of restrictions on the prices of risk as functions of the SDE parameters. Some insight into the impact of these restrictions can be gained by examining the parameter estimates for ξ_{22} and Σ_{22}. QTSM4 restricts $\xi_{22}^2 > 2\Sigma_{22}^2$, which is violated by the parameter estimates shown in the table. The results for QTSM3 compared to QTSM4

[17] The z-statistic is calculated as $\frac{\chi^2 - df}{\sqrt{2df}}$ and represents a degrees of freedom normalization of the χ^2 statistic.

Table 1
Tests of three-factor QTSMs

Parameter	QTSM4	QTSM3	QTSM2	QTSM1
	Estimate (standard error)			
α	0.0373 (0.0524)	0.0176 (0.0062)	0.0180 (0.0059)	0.0338 (0.0053)
Ψ_{12}			0.0393 (0.0541)	−0.5847 (0.2926)
Ψ_{13}			0.0409 (0.1301)	−0.3866 (0.3586)
Ψ_{23}			−0.0555 (0.0250)	0.9101 (0.0731)
μ_1	0.0547 (0.0084)	0.0436 (0.1087)	0.0214 (0.2088)	0.0608 (0.1019)
μ_2	0.2937 (0.2549)	0.0007 (0.0001)	0.0006 (0.0000)	0.0007 (0.0000)
μ_3	0.0625 (0.0725)	0.0299 (0.0641)	0.2055 (0.1550)	0.1343 (0.1453)
ξ_{11}	−0.4030 (0.0005)	−1.5412 (0.2879)	−1.6815 (0.1850)	−1.7318 (0.2629)
ξ_{21}				0.0043 (0.0131)
ξ_{31}				2.7202 (1.1386)
ξ_{22}	−0.1690 (0.0378)	−0.0003 (0.0000)	−0.0003 (0.0000)	−0.0037 (0.0006)
ξ_{32}				−1.1898 (0.2570)
ξ_{33}	−3.9839 (0.5243)	−2.8718 (0.2408)	−2.7772 (0.3115)	−0.3411 (0.0521)
Σ_{11}^2	0.0805 (0.0001)	0.0173 (0.0037)	0.0147 (0.0037)	1.5143* (1.2167)
Σ_{22}^2	0.0023 (0.0019)	0.0009 (0.0001)	0.0008 (0.0001)	0.1543* (0.0472)
Σ_{33}^2	1.1673 (0.1273)	0.4924 (0.0428)	0.4980 (0.0420)	0.5916* (0.1132)
$\mu_1 - \delta_{01}$	−0.4784 (0.0000)	−0.0665 (0.1063)	0.0428 (0.2345)	−0.0416 (0.0762)
$\mu_2 - \delta_{02}$	0.2668 (0.0000)	−0.1197 (0.0187)	−0.1338 (0.0217)	0.0317 (0.0301)
$\mu_3 - \delta_{03}$	0.0573 (0.0000)	0.1398 (0.0144)	0.0851 (0.0341)	0.2534 (0.1451)
$\xi_{11} - \delta_{111}$	−0.0375 (0.0000)	1.6130 (0.1828)	1.7873 (0.0996)	−1.8465 (2.1800)
$\xi_{21} - \delta_{121}$				0.0226 (0.5891)
$\xi_{31} - \delta_{131}$				−2.4858 (1.8197)
$\xi_{22} - \delta_{122}$	−0.1548 (0.0000)	0.0358 (0.0056)	0.0328 (0.0074)	−0.2782 (0.0413)
$\xi_{32} - \delta_{132}$				−1.2022 (0.1988)
$\xi_{33} - \delta_{133}$	−3.6792 (0.0000)	0.6474 (0.0450)	0.6323 (0.0806)	−0.2459 (0.0215)
χ^2	508.120	122.599	114.504	51.297
df	32	26	23	17
z	59.515	13.396	13.491	5.882

$^* \times 10^{-4}$

Table 1 presents parameter estimates and goodness-of-fit tests for nested versions of the QTSM described in this article. The model and parameters are described in Section 2. The columns present parameter estimates for QTSM1, QTSM2, QTSM3, and QTSM4, as described in Section 3. The table also presents χ^2 statistics for the goodness-of-fit of the models and a z-statistic that adjusts for degrees of freedom across the models and is distributed $N(0, 1)$.

suggest that relaxing these restrictions is potentially quite important in the model's ability to fit the term structure.

The third column represents estimates for QTSM2, in which the SDE parameters are orthogonal, but the state variables interact in the determination of the short rate. The results indicate that this version of the model offers little improvement relative to QTSM3. Although the chi-squared statistic falls, the z-statistic of 13.491 suggests that the loss of degrees of freedom implied by the additional parameters more than offsets the improvement in fit. The easing of the restrictions on the short rate does not appear to materially impact the values of the coefficient estimates, which are close to those implied by QTSM3.

The final column presents estimates for QTSM1, the full-fledged QTSM. The results of this estimation suggest that allowing for correlation among the factors results in dramatic improvement in the model's fit. The z-value for the test of model fit falls to 5.882, which, despite indicating rejection of the model, represents a dramatic improvement relative to the 13.396 value for QTSM3. This evidence is broadly consistent with that presented in Duffie and Singleton (1997) and Dai and Singleton (2000). However, the estimates highlight important differences in the QTSM and the ATSM. Specifically, substituting the estimates of QTSM1 into Equation (8) suggests that the second state variable is positively correlated with the first and third state variables, whereas the first and third state variables are negatively correlated. Thus the QTSM does not require positive correlations among all state variables in contrast to the ATSM. This restriction is eased primarily due to the nonaffine structure between the interest rate and the state variables, and thus the correlation structure among the state variables may differ.[18]

Additional insight into the performance of the models can be derived from analyzing the scores of the best model fits with respect to the SNP parameter vector. Figure 2 presents t-ratios for the significance of the model scores. All of the models perform fairly well in capturing the mean dynamics of the VAR part of the SNP model; the SAINTS and orthogonal QTSMs each have t-ratios greater than 2.0 for one of the ψ parameters, which govern the mean of the VAR, whereas the remaining models have no significant t-ratios. However, the models cannot describe the ARCH innovations to the VAR and the Hermite polynomial terms. Of the 18 scores with respect to the τ terms, 12 are significant for the SAINTS model, 5 each for the two orthogonal QTSMs, and 4 for the full-fledged QTSM.

The scores with respect to the Hermite polynomial terms reveal more interesting patterns. In particular, the scores for the Hermite terms suggest that the models fail to capture the shape of the density for the short-term Treasury-bill yields. The SAINTS model is able to capture the scores with respect to A(2), A(5), A(8), and A(11), but cannot capture the remaining Hermite scores. This result suggests that the model generally captures the shape of the long-term bond density, but not the shorter-term instruments. At the opposite end of the spectrum, QTSM1 is able to capture most of the shape features of the density implied by the Hermite coefficients. The exceptions are A(7), A(10), and A(13) which suggest that the model has some difficulty in matching the shape of the density for the short-term yield. The performances of QTSM2 and QTSM3 are similar, and fall between the SAINTS model and the maximal model. This difficulty of the QTSMs in matching the shape of the short-term yield density may either reflect a drawback of the QTSM or be an outcome of the difference in institutional structure between the short-term bond market and the rest of the Treasury bond market, which is addressed in Knez, Litterman, and Scheinkman (1996).

In summary, the results of the goodness-of-fit tests suggest that easing the restrictions on the prices of risk imposed by the SAINTS model results in a

[18] A potential weakness of the QTSMs is their estimates of the lower bound on the short rate α. For example, QTSM4 and QTSM1 imply unreasonable lower bound of 3.73% and 3.38% respectively, counterfactual levels that are greater than some of the observed yields in the earlier part of the sample period.

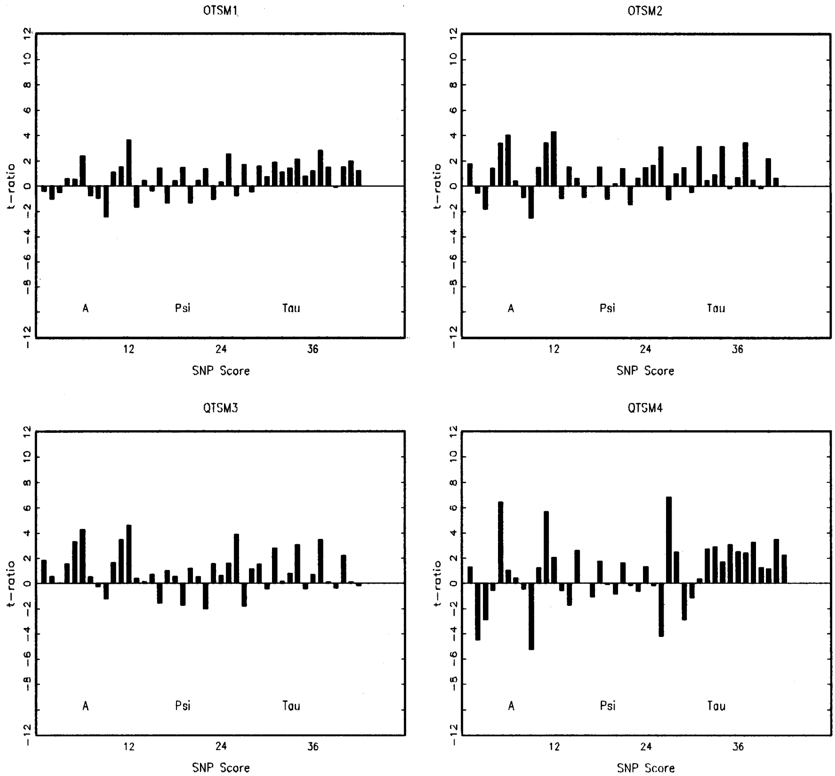

Figure 2
EMM Scores: QTSMs
t-ratio diagnostics for the scores implied by the nested versions of the QTSM. The first group of 12 t-ratios represent the scores with respect to A, the Hermite polynomial terms; A(2)–A(4) are linear terms, A(5)–A(7) are quadratic terms, A(8)–A(10) are cubic terms, and A(11)–A(13) are quartic terms. The next group of 12 represents the scores with respect to Ψ, the mean coefficients of the VAR part of the SNP estimation. The final group of 18 represents the scores with respect to τ, the ARCH part of the SNP VAR.

vast improvement in fit for the QTSM. Further improvements are made by relaxing the restriction of orthogonality among the state variables; allowing for unconditional correlation among the state variables dramatically improves model fit. However, even though the model performs well overall, it is not fully able to capture the dynamics of the term structure. In the next section we gauge the ability of the model to capture these dynamics relative to the ATSM class of models analyzed in Dai and Singleton (2000).

4.1.2 Affine term structure models. We estimate two ATSMs discussed in Dai and Singleton (2000). We estimate their preferred model, ATSM1, that allows for both conditional and unconditional correlation among the factors, and an orthogonal version of their model, ATSM2, the Chen (1996) model.

ATSM1 can be described as

$$d\begin{pmatrix} v(t) \\ \theta(t) \\ r(t) \end{pmatrix} = \begin{pmatrix} \kappa_{11} & 0 & 0 \\ \kappa_{21} & \kappa_{22} & 0 \\ \kappa_{31} & -\kappa_{33} & \kappa_{33} \end{pmatrix} \begin{pmatrix} \theta_1 - v(t) \\ \theta_2 - \theta(t) \\ \theta_2 - r(t) \end{pmatrix} dt + \begin{pmatrix} \Sigma_{11}\sqrt{v(t)} & 0 & 0 \\ 0 & \Sigma_{22}\sqrt{\theta(t)} & 0 \\ \Sigma_{31}\Sigma_{11}\sqrt{v(t)} & 0 & \sqrt{v(t)} \end{pmatrix} dz_N(t) \qquad (22)$$

or

$$dY(t) = \kappa(\theta - Y(t))\,dt + \Sigma(t)\,dz_N(t)$$

with prices of risk given by

$$\Sigma(t)\lambda,$$

where λ is a conforming vector of constants. In the orthogonal case, $\kappa_{21} = \kappa_{31} = \kappa_{32} = \Sigma_{31} = 0$.

Tests of the models' restrictions are presented in Table 2. Similar to the performance of the SAINTS model in the previous section, ATSM2 is rejected strongly in the data, with a z-statistic of 69.817. This evidence suggests that an affine model with uncorrelated state variables is not able to describe term structure dynamics very well. In contrast, ATSM1 performs vastly better with a z-statistic of 30.278. However, the model is still rejected strongly in the data and performs poorly relative to QTSM3, the orthogonal QTSM. The t-ratios for the scores of the model presented in Figure 3 provide some further insight into this failure. The model's main difficulty relative to the QTSM is in capturing the ARCH innovations; the scores with respect to several of the ARCH terms are quite large. This finding contrasts with Dai and Singleton (2000), who find that the model is able to accommodate these scores quite well. Much of this difference may be attributable to the sample period; whereas our sample encompasses the high interest rate and high volatility regime of the early 1980s, the Dai and Singleton study focuses on yields after 1987, wherein the interest rates have been relatively stable.

The results of the analysis for the affine and nonaffine models suggest that some of the restrictive features of the ATSM class, such as the limited correlation structure among state variables discussed above, hamper the ability of the models to fit the dynamics of the term structure. Relaxing the restrictions imposed by the affine functional form yields a vast improvement in model fit; the z-statistic of the full-fledged QTSM is 5.882, compared to 30.278 for the preferred ATSM.[19] In the next section we utilize the reprojection methodology of Gallant and Tauchen (1998) to further assess the ability of

[19] This result does not necessarily indicate the overall superiority of the QTSMs over ATSMs. This article adopts a different sample period and maturity points on the yield curve than Dai and Singleton (2000). Thus the correlation structure among the state variables may be more prominent in the sample period and chosen maturity points used in this article.

Table 2
Tests of three-factor ATSMs

Parameter	Estimate (standard error) ATSM2	Correlated (DS)
κ_{11}	0.7806 (0.0633)	0.2839 (0.0403)
κ_{21}		−0.0033 (0.0238)
κ_{22}	0.0035 (0.0008)	0.0017 (0.0008)
κ_{31}		−9.6384 (1.2872)
κ_{33}	5.5350 (0.2972)	9.5814 (0.4915)
θ_1	0.6723 (0.0451)	0.8188 (0.0929)
θ_2	7.7365 (0.4164)	7.7089 (0.8172)
Σ_{11}^2	0.1446 (0.0237)	0.1029 (0.0139)
Σ_{22}^2	0.0158 (0.0007)	0.0051 (0.0005)
Σ_{31}		−0.3895 (0.1348)
λ_1	−9.6125 (1.3393)	−9.4651 (2.1612)
λ_2	−1.4163 (0.2757)	−1.2472 (1.4057)
λ_3	−0.2241 (0.0881)	−0.7687 (0.9455)
χ^2	590.539	259.589
df	32	29
z	69.817	30.278

Table 2 presents parameter estimates and goodness-of-fit statistics for nested versions of a three-factor affine model, as described in Dai and Singleton (2000). The first column presents estimates of ATSM2, the Chen (1996) model, an orthogonal version of the general affine case. The second column presents estimates for the optimal model estimated in Dai and Singleton (2000), ATSM1. The form of the model is given by the physical measure process

$$dY(t) = \kappa(\theta - Y(t))\,dt + \Sigma(t)\,dW(t)$$

with prices of risk given by

$$\Sigma(t)\lambda.$$

χ^2 statistics for the goodness-of-fit of the models and a z-statistic that adjusts for degrees of freedom across the models and is distributed $N(0,1)$ are presented in the final three rows of the table. The parameter values are based on percentage yields.

the QTSM and the ATSM to describe the dynamics of the term structure. In particular, this methodology allows us to examine the models' implications for the conditional expectation and volatility of future yields.

4.2 Reprojection

We briefly summarize the reprojection method in this section; a complete discussion is provided in Gallant and Tauchen (1998). We denote the conditional density implied by the QTSM for observables as

$$\hat{p}(y_0|y_{-L},\ldots,y_{-1}) = p(y_0|y_{-L},\ldots,y_{-1},\hat{\rho}_n), \tag{23}$$

where $\hat{\rho}$ denotes the estimated model parameters. Although analytic expressions for Equation (23) are not available, an unconditional expectation, $\mathscr{E}_{\hat{\rho}_n}$ can be computed by generating a simulation $\{\hat{y}_t\}_{t=-L}^{N}$ from the system with parameters set to $\hat{\rho}_n$ and approximating

$$\mathscr{E}_{\hat{\rho}_n}(g) = \frac{1}{N}\sum_{t=0}^{N} g(\hat{y}_{t-L},\ldots,\hat{y}_t). \tag{24}$$

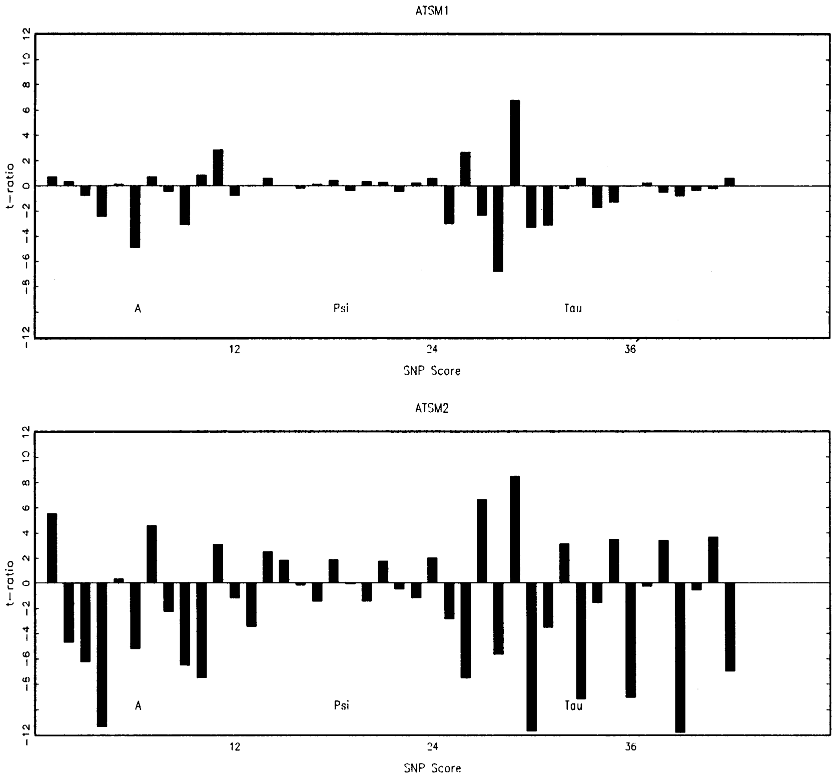

Figure 3
EMM Scores: ATSMs
t-ratio diagnostics for the scores implied by the nested versions of the ATSM. ATSM1 is the maximal ATSM of Dai and Singleton (2000), and ATSM2 is the orthogonal ATSM. The first group of 12 t-ratios represent the scores with respect to A, the Hermite polynomial terms; A(2)–A(4) are linear terms, A(5)–A(7) are quadratic terms, A(8)–A(10) are cubic terms, and A(11)–A(13) are quartic terms. The next group of 12 represents the scores with respect to Ψ, the mean coefficients of the VAR part of the SNP estimation. The final group of 18 represents the scores with respect to τ, the ARCH part of the SNP VAR.

With respect to unconditional expectation so computed, define

$$\hat{\theta}_K = \arg\max \theta \in \Re^{p_K} \mathscr{E}_{\hat{\rho}_n} \log f_K(y_0|y_{-L}, \ldots, y_{-1}, \theta), \tag{25}$$

where $f_K(y_0|y_{-L}, \ldots, y_{-1}, \theta)$ is the SNP conditional density of Equation (19). Let

$$\hat{f}_K(y_0|y_{-L}, \ldots, y_{-1}) = f_K(y_0|y_{-L}, \ldots, y_{-1}, \hat{\theta}_K). \tag{26}$$

Theorem 1 of Gallant and Long (1997) states that

$$\lim_{K\to\infty} \hat{f}_K(y_0|y_{-L}, \ldots, y_{-1}) = \hat{p}(y_0|y_{-L}, \ldots, y_{-1}).$$

Convergence is with respect to a weighted Sobolev norm that they describe. Of relevance here is that convergence in their norm implies that $\hat{f}_K$ as well as its partial derivatives in $(y_{-L}, \ldots, y_{-1}, y_0)$ converge uniformly over $\Re^\ell$, $\ell = M(L+1)$, to those of $\hat{p}$. The idea of reprojection is to study the dynamics of $\hat{p}$ by using $\hat{f}_K$ as an approximation; this result provides the justification.

Of immediate interest in eliciting the dynamics of observables are the first two one-step-ahead conditional moments:

$$\mathscr{E}(y_0|y_{-L}, \ldots, y_{-1}) = \int y_0 f_K(y_0|x_{-1}, \hat{\theta}_K)\, dy_0$$

and

$$\mathrm{var}(y_0|y_{-L}, \ldots, y_{-1}) = \int [y_0 - \mathscr{E}(y_0|x_{-1})][y_0 - \mathscr{E}(y_0|x_{-1})]' f_K(y_0|x_{-1}, \hat{\theta}_K)\, dy_0,$$

where $x_{-1} = (y_{-L}, \ldots, y_{-1})$.

Figure 4 plots $\mathscr{E}(y_0|y_{-L}, \ldots, y_{-1})$ with $(y_{-L}, \ldots, y_{-1})$ set, successively, to the values $(y_{t-L}, \ldots, y_{t-1})$ observed in the data for QTSM1. The plot suggests that the model fares quite well in capturing the first conditional moment of yield changes. Even in the high-rate regime of 1979–1982, the deviations in the rates predicted by the model from the observed yields are relatively small. The plot suggests that the model captures the long-bond dynamics somewhat better than the short-term Treasury-bill yield; deviations from the observed 3-month Treasury-bill rates are more apparent during this high-rate regime. This result is consistent with the evidence suggested by the t-ratios for the EMM estimation discussed above.

In Figure 5 we reproduce a similar plot to Figure 4 for ATSM1. Consistent with evidence in Duffee (2000), the ATSM is unable to adequately capture yield changes. The deviations in this case between the conditionally expected yields and actual yields are large, even outside of the high-rate regime of 1979–1982. As in the case of the QTSM, the model appears best able to fit the yields of the long bond at the expense of the fit for the 3-month bond. A significant degree of this difficulty in fitting the conditional expectation of yields may be due to the specification of the price of risk. Duffee (2000) and Dai and Singleton (2001) find that modifying the price of risk in the affine framework significantly improves the affine class' ability to capture yield changes.

Figures 6–8 plot the conditional volatilities implied by the fitted SNP model for the observed data, the conditional volatility implied by the QTSM, and the conditional volatility implied by the ATSM for the 3-month, 1-year, and 10-year Treasuries, respectively. The plots suggest that the QTSM is able to capture the shape of the SNP conditional volatility for all three bonds quite well. However, the model performs better in matching the *level* of conditional volatility better for some of the bonds than the others. In particular, the QTSM captures the conditional volatility of the 10-year bond quite well,

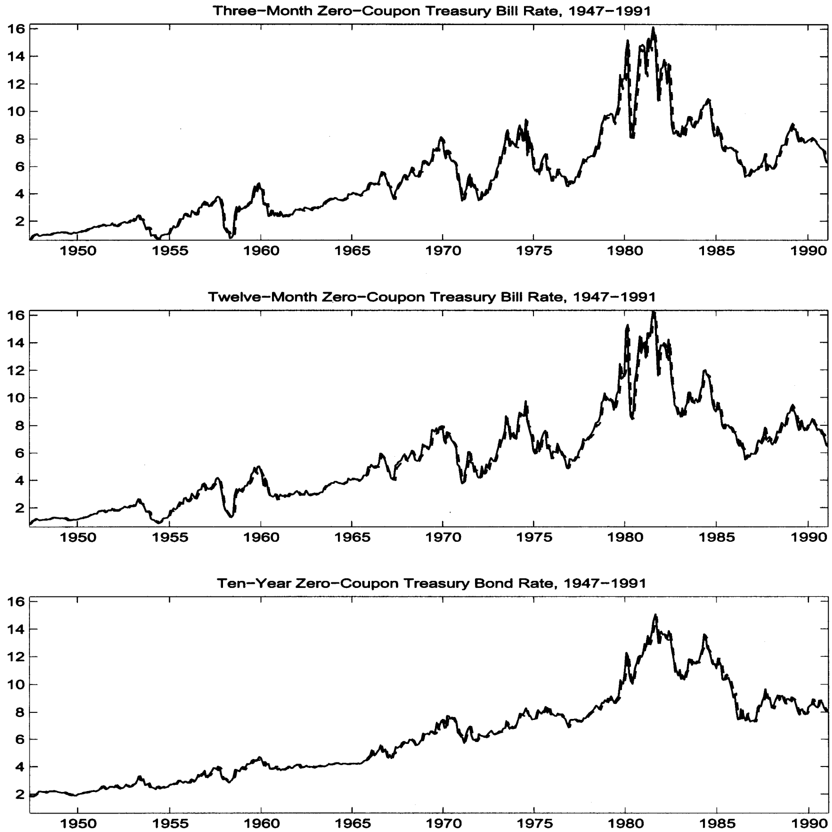

Figure 4
Reprojected Yields: QTSM
A plot of the reprojected bond yields from the full-fledged QTSM, QTSM1 against the actual yields observed in the data. The first figure represents the 3-month Treasury-bill yields, the second figure depicts 1-year Treasury-bill yields, and the third figure shows the 10-year Treasury bond yields.

matching both its shape and level. However, the model is not able to generate the level of the 1-year bond volatility. The level of the reprojected conditional volatility of the 3-month bond comes closer to matching that of the SNP conditional volatility than that of the 1-year bond, but does not match quite as well as that of the 10-year bond. Thus, as suggested by the EMM diagnostics, the QTSM's difficulties lie largely in capturing all of the conditional volatility features of the data.

The conditional volatility plots for the ATSM suggest that the model fares very poorly in capturing the conditional volatility of yield changes. The model is particularly poor in generating conditional volatilities that match the shape or level of the conditional volatility of the 3-month and 1-year Treasury-bill yields. The performance of the model with respect to

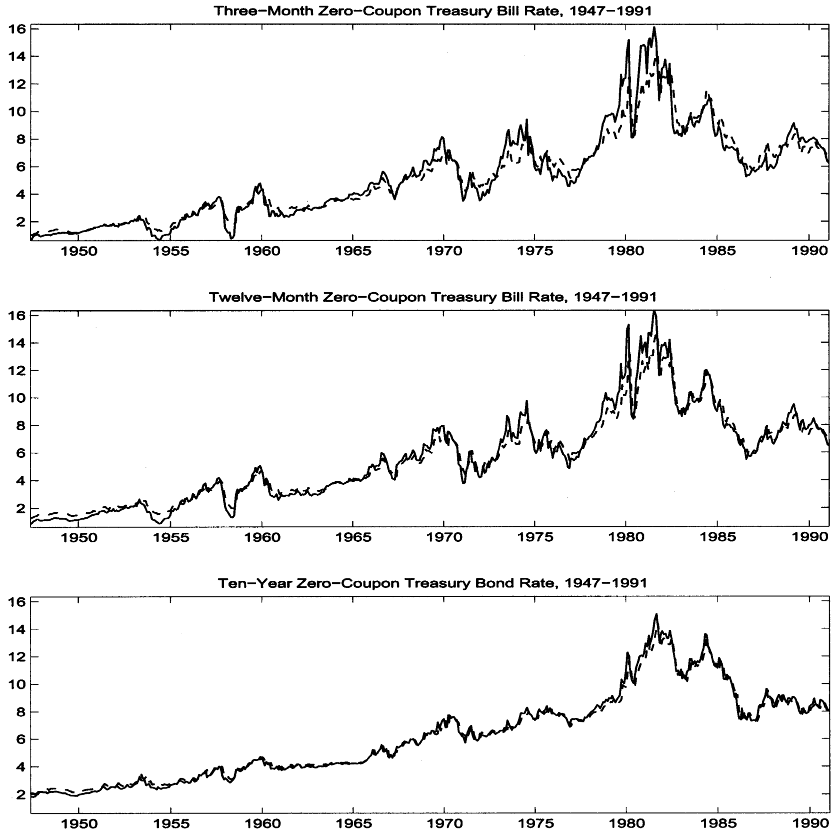

Figure 5
Reprojected Yields: ATSM
A plot of the reprojected bond yields from the ATSM estimated in Dai and Singleton (2000) against the actual yields observed in the data. The first figure represents the 3-month Treasury-bill yields, the second figure depicts 1-year Treasury-bill yields, and the third figure shows the 10-year Treasury-bond yields.

the 10-year bond yield is somewhat better, but still falls short of providing a good description of conditional volatility. Thus, as suggested earlier, the trade-off between heteroscedastic volatility and conditional correlation appears to sharply impact the ability of the ATSM to fit the data.

4.3 Discussion and interpretation

The results of the estimation above suggest that several features of the QTSM contribute to the fit of the observed bond yields. First, the general model improves upon the SAINTS model in that it loosens restrictions on the parameters governing the risk premia of the factors. As indicated by the substantial improvement in fit moving from the restricted to the unrestricted case, this

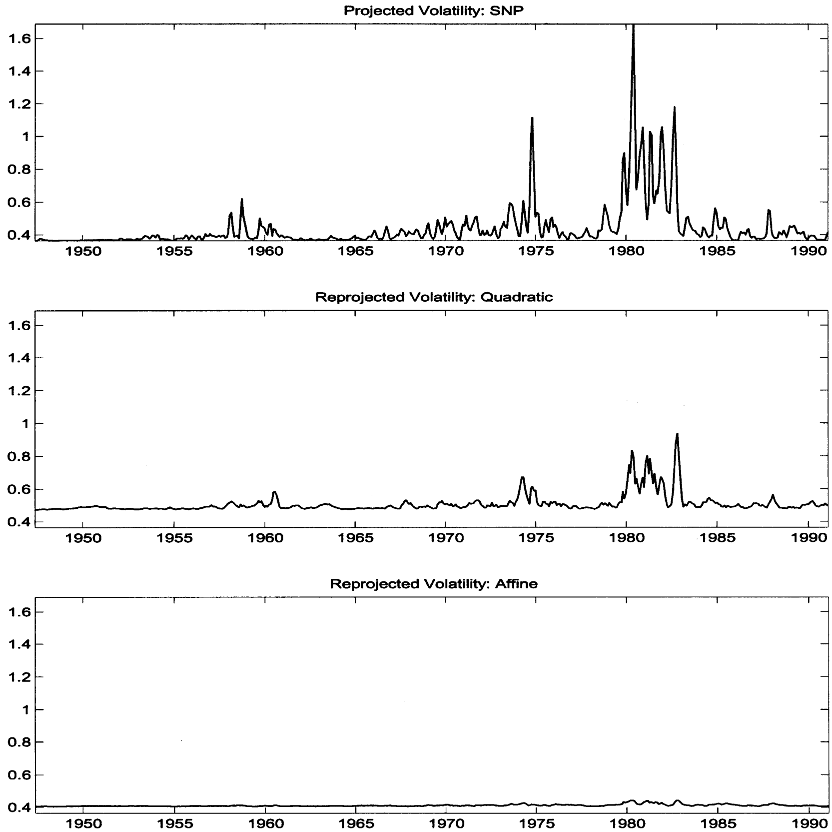

Figure 6
Reprojected Volatility: 3-Month Bill
A plot of the reprojected volatility of 3-month Treasury-bill yields. The first figure depicts the conditional volatility implied by the SNP fit to the data, the second figure shows the conditional volatility implied by QTSM1, and the third figure shows the conditional volatility implied by the ATSM estimated in Dai and Singleton (2000).

flexibility is quite important in describing the dynamics of bond yields. Second, when we allow for unconditional correlation among the factors, the fit is improved considerably. This result is similar to that shown in Duffie and Singleton (1997) and Dai and Singleton (2000). As we have shown in Figure 4, allowing for this unconditional correlation allows us to capture much of the dynamics of the observed term structure.

The model also fares well in contrast to the class of affine term structure models examined in Dai and Singleton (2000). This difference suggests that allowing for nonlinearity in the pricing of bond yields is quite important for describing the term structure. The result supports the findings of Aït Sahalia (1996b) and Stanton (1997), who provide evidence that the drift of the short rate process is nonlinear. However, evidence of the nonlinearity of the drift of

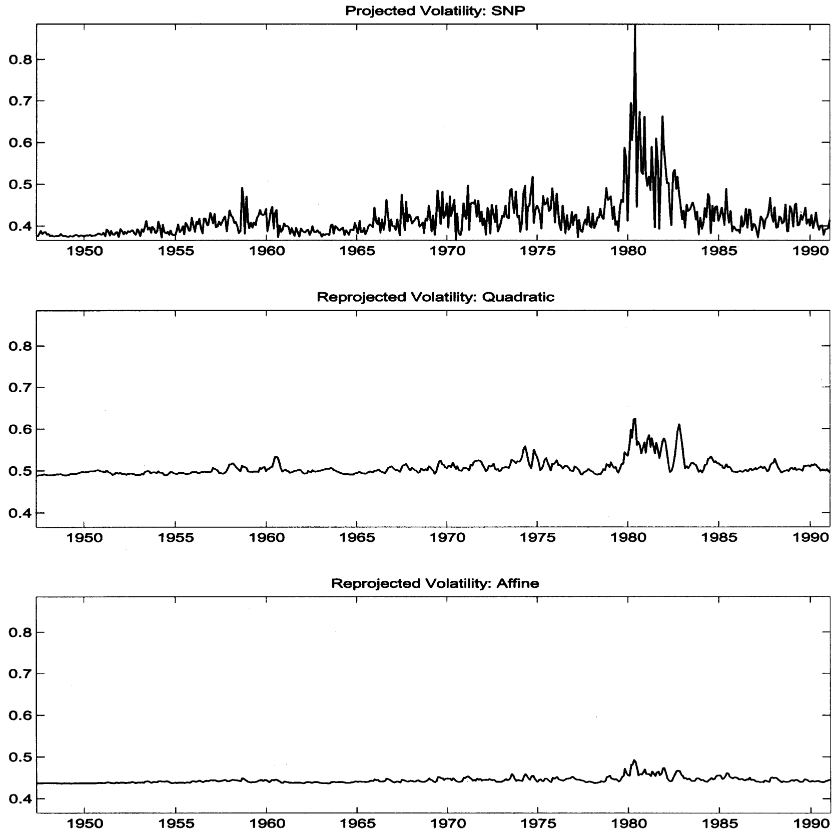

Figure 7
Reprojected Volatility: 1-Year Bill
A plot of the reprojected volatility of 1-year Treasury-bill yields. The first figure depicts the conditional volatility implied by the SNP fit to the data, the second figure shows the conditional volatility implied by QTSM1, and the third figure shows the conditional volatility implied by the ATSM estimated in Dai and Singleton (2000).

the short rate remains somewhat controversial; Chapman and Pearson (2000) provide evidence suggesting that this nonlinearity is not a robust stylized fact. Insofar as the nonaffine class of models nested in the QTSM framework implies a nonlinear drift for the short rate, the model is consistent with these findings. Furthermore, the results suggest that it is probable that the trade-off between flexibility in specifying volatility and correlation structure materially impacts the model's ability to fit the data.

However, the results do suggest that the model cannot fully capture the dynamics of the term structure. The EMM diagnostics provide some insight into the source of this failure. The model was not fully able to capture the ARCH and non-Gaussian features of the observed data. These results suggest that some further flexibility in modeling the diffusion of the process may

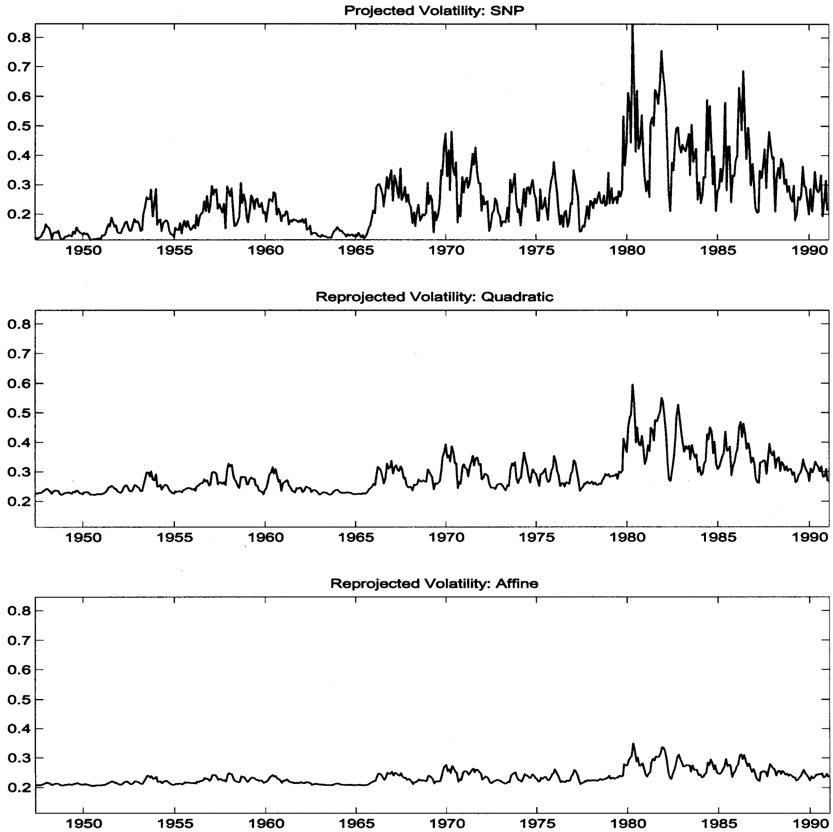

Figure 8
Reprojected Volatility: 10-Year Bond
A plot of the reprojected volatility of 10-year Treasury-bond yields. The first figure depicts the conditional volatility implied by the SNP fit to the data, the second figure shows the conditional volatility implied by QTSM1, and the third figure shows the conditional volatility implied by the ATSM estimated in Dai and Singleton (2000).

contribute to the fit of the model. One possibility is that a hybrid of affine and nonaffine models may better describe term structure dynamics. Alternatively, a nonaffine diffusion process may be necessary to fully describe the volatility of yield changes.

5. Conclusion

Much of the term structure literature has focused on ATSMs, models that specify bond yields as affine functions of underlying state variables. Although this class is popular due to its tractability and relatively straightforward empirical implementation, the models suffer from several noteworthy drawbacks. In particular, the affine form of the yields results in a trade-off between the structure of the correlation matrix for the state variables and their conditional variance. Further, the evidence in Dai and Singleton (2000) suggests that the class may fail to capture important nonlinearities in the data, and, in general, the framework cannot guarantee a positive nominal interest rate.

In contrast, models that specify yields as a quadratic function of the underlying state variables have received less theoretical and empirical attention. However, the QTSM class of models is attractive because the functional form of the models overcomes many of the drawbacks of ATSM. We derive a general form for the family of QTSMs which nests existing models of its class. We derive a particular equilibrium that supports the QTSM and represent the model in canonical form. The canonical form renders empirical implementation of the model tractable.

We estimate parameters for four versions of the QTSM and assess their goodness-of-fit using the EMM procedure of Gallant and Tauchen (1996). We first find that the restrictions imposed by the SAINTS model of Constantinides (1992) on the QTSM result in a strong rejection of the model. Easing these restrictions results in an orthogonal QTSM, which fits the term structure dynamics considerably better than the SAINTS model. The fit of the model is improved dramatically by allowing for unconditional correlations among the state variables. In this case, the QTSM provides a fairly good description of term structure dynamics and captures these dynamics considerably better than the preferred ATSM investigated in Dai and Singleton (2000).

The QTSM captures conditional expectations of future bond yields at both the long and the short end of the term structure quite well. It also is able to match the shape features of conditional volatility of bond yields across the spectrum of the term structure. However, the model is not able to generate the level of conditional volatility observed for the short- and intermediate-term bond yields. It is possible that either some combination of ATSM and QTSM may be able to accommodate the level of the volatility. Alternatively, the state variables may need to be nonaffine in order to generate sufficient conditional volatility, as in Ahn and Gao (1999). Our results suggest that the ability of a model to generate this conditional volatility is quite important for the fit of term structure models.

Appendix A: Distributions of State Variables

The SDE of the state vector is represented as

$$dY(t) = [\mu + \xi Y(t)]\,dt + \Sigma\, dz_N(t).$$

Since ξ is diagonalizable, that is, a *regular* matrix, spectral decomposition of ξ leads to

$$U^{-1}\xi U = \Lambda,$$

since each of the eigenvalues are regular.[20] We also define

$$\mathscr{V} = [v_{ij}]_{NN} = U^{-1}\Sigma\Sigma' U'^{-1}.$$

A.1 Conditional distribution

We prove that the time s joint distribution conditional upon time t information ($s > t$) is

$$Y(t+\tau)|Y(t) \sim \mathrm{MVN}_N\big(E[Y(t+\tau)|Y(t)], \mathrm{var}[Y(t+\tau)|Y(t)]\big),$$

where

$$\begin{aligned} E[Y(t+\tau)|Y(t)] &= U\Lambda^{-1}[\Phi(\tau)-I_N]U^{-1}\mu + U\Phi(\tau)U^{-1}Y(t) \\ \mathrm{var}[Y(t+\tau)|Y(t)] &= U\left[\frac{v_{ij}(\exp((\lambda_i+\lambda_j)\tau)-1)}{\lambda_i+\lambda_j}\right]_{NN} U' \\ \Phi(\tau) &\triangleq \mathrm{diag}[\exp(\lambda_i\tau)]_N. \end{aligned}$$

Proof. It is well known that multivariate Ornstein–Uhlenbeck processes follow multivariate Gaussian distributions. As such, we need to determine the vector of conditional means and conditional covariance matrix in order to define the distribution. First, we apply a nonsingular linear transformation of the factors, $Z(t) = U^{-1}Y(t)$, the SDE of which is characterized as

$$\begin{aligned} dZ(t) &= d(U^{-1}Y(t)) = \big[U^{-1}\mu + U^{-1}\xi UU^{-1}Y(t)\big]\,dt + U^{-1}\Sigma\, dz_N(t) \\ &= \big[U^{-1}\mu + \Lambda Z(t)\big]\,dt + U^{-1}\Sigma\, dz_N(t). \end{aligned}$$

The solution to the above SDE is [see Karatzas and Shreve (1991)]

$$Z(t+\tau) = \Phi(\tau)\left[Z(t) + \int_0^\tau \Phi(s)^{-1}U^{-1}\mu\, ds + \int_0^\tau \Phi(s)^{-1}U^{-1}\Sigma^{1/2}\, dz(t+s)\right],$$

where

$$\frac{d\Phi(\tau)}{d\tau} = \Lambda\Phi(\tau) \text{ and } \Phi(0) = I_N,$$

which leads to the solution $\Phi(\tau) = \mathrm{diag}[\exp(\lambda_i\tau)]_N$.

- *Conditional mean vector*

$$\begin{aligned} &E[Z(t+\tau)|Z(t)] \\ &\quad = \Phi(\tau)Z_t + \Phi(\tau)\int_0^\tau \Phi(s)^{-1}U^{-1}\mu\, ds \\ &\quad = \Phi(\tau)Z_t + \Phi(\tau)\big[\mathrm{diag}\big(-\Lambda^{-1}(\exp(-\Lambda\tau)-I_N)\big)\big]U^{-1}\mu \\ &\quad = \Phi(\tau)Z_t + \Lambda^{-1}[\Phi(\tau)-I_N]U^{-1}\mu, \end{aligned}$$

which leads to

$$\begin{aligned} E[Y(t+\tau)|Y(t)] &= UE[Z(t)|Z(0)] \\ &\qquad = U\Phi(\tau)U^{-1}Y(t) + U\Lambda^{-1}[\Phi(\tau)-I_N]U^{-1}\mu. \end{aligned}$$

[20] Since ξ has eigenvalues ξ_i with multiplicity m_k for $k = 1, 2, \ldots, s$ and $\sum_{k=1}^s m_k = N$, it has N eigenvectors that are linearly independent if and only if $\mathrm{rank}(\xi - \lambda_k I_N) = N - m_k\ \ \forall\ k = 1, 2, \ldots, s$; whereupon U is nonsingular and ξ is diagonalizable. Eigenvalue λ_k satisfying $\mathrm{rank}(\xi - \lambda_k I_N) = N - m_k$ is called a regular matrix. Therefore, when every eigenvalue is regular, ξ is diagonalizable.

- *Conditional covariance matrix*

$$\begin{aligned}\operatorname{var}[Z(t+\tau)|Z(t)] &= \Phi(\tau)\left[\int_0^\tau \Phi(s)^{-1}U^{-1}\Sigma\Sigma' U'^{-1}\Phi(s)^{-1}\,ds\right]\Phi(\tau)\\ &= \Phi(\tau)\left[\int_0^\tau \Phi(s)^{-1}\mathcal{V}\Phi(s)^{-1}\,ds\right]\Phi(\tau).\end{aligned} \tag{27}$$

Solving the argument in the integral yields

$$\Phi(s)^{-1}\mathcal{V}\Phi(s)^{-1} = \left[\exp((-\lambda_i - \lambda_j)s)v_{ij}\right]_{NN}. \tag{28}$$

Substituting Equation (28) into Equation (27) results in

$$\int_0^\tau \Phi(s)^{-1}\mathcal{V}\Phi(s)^{-1}\,ds = \left[-\frac{v_{ij}(\exp(-(\lambda_i+\lambda_j)\tau)-1)}{\lambda_i+\lambda_j}\right]_{NN},$$

which leads to

$$\begin{aligned}\operatorname{var}[Z(t+\tau)|Z(t)] &= \Phi(\tau)\left[-\frac{v_{ij}(\exp(-(\lambda_i+\lambda_j)\tau)-1)}{\lambda_i+\lambda_j}\right]_{ij}\Phi(\tau)\\ &= \left[\frac{v_{ij}(\exp((\lambda_i+\lambda_j)\tau)-1)}{\lambda_i+\lambda_j}\right]_{ij}.\end{aligned}$$

Finally, the conditional covariance matrix of the state variables is

$$\begin{aligned}\operatorname{var}[Y(t+\tau)|Y(t)] &= U\operatorname{var}[Z(t+\tau)|Z(t)]U'\\ &= U\left[\frac{v_{ij}(\exp((\lambda_i+\lambda_j)\tau)-1)}{\lambda_i+\lambda_j}\right]_{ij}U'.\end{aligned}$$

A.2 Unconditional distribution

Provided $\operatorname{Re}[\lambda_i] < 0\ \forall i = 1, 2, \ldots, N$, the steady-state multivariate distribution of the state vector is defined as

$$Y \sim \mathrm{MVN}_N(E[Y], \operatorname{var}[Y]),$$

where

$$\begin{aligned}E[Y] &= \lim_{\tau\uparrow\infty} E[Y(t+\tau)|Y(t)]\\ &= \lim_{\tau\uparrow\infty} U\Phi(\tau)U^{-1}Y(t) + \lim_{\tau\uparrow\infty} U\Lambda^{-1}[\Phi(\tau)-I_N]U^{-1}\mu\\ &= -U\Lambda^{-1}U^{-1}\mu\\ &= -\xi^{-1}\mu\\ \operatorname{var}(Y) &= \lim_{\tau\uparrow\infty}\operatorname{var}[Y(t+\tau)|Y(t)]\\ &= \lim_{\tau\uparrow\infty} U\left[\frac{v_{ij}(\exp((\lambda_i+\lambda_j)\tau)-1)}{\lambda_i+\lambda_j}\right]_{NN}U'\\ &= U\left[-\frac{v_{ij}}{\lambda_i+\lambda_j}\right]_{NN}U'.\end{aligned}$$

Appendix B: Distributions of the Instantaneous Interest Rate

In Assumption 2, the instantaneous nominal interest is written as

$$r(t) = \alpha + Y(t)'\Psi Y(t),$$

which is a quadratic function of the state variables. As shown in Appendix A, the unconditional as well as conditional distributions are multivariate normal densities. For brevity we use the symbols φ and V to denote the mean vector and covariance matrix, respectively, for either conditional or unconditional distributions. The probability density is given by

$$f_y(Y) = (2\pi)^{-\frac{1}{2}N}|V|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(Y-\varphi)'V^{-1}(Y-\varphi)\right).$$

We wish to identify the distribution of r, that is

$$\begin{aligned}\Pr\left[r = \alpha + Y(t)'\Psi Y(t) \le r_0\right] &= \Pr\left[Y(t)'\Psi Y(t) \le r_0 - \alpha\right] \\ &= \int_{Y'\Psi Y \le (r_0-\alpha)} f_y(Y)\, dY. \end{aligned} \tag{29}$$

V is a symmetric and positive semidefinite matrix. Therefore there exists a nonsingular lower triangular matrix L which factors V as $V = LL'$. Define $X_L \triangleq L^{-1}(Y-\varphi)$, which is a *standard* normal multivariate. Then we can reexpress Equation (29) as

$$\begin{aligned}&\Pr\left[r = \alpha + Y(t)'\Psi Y(t) \le r_0\right] \\ &\quad = (2\pi)^{-\frac{1}{2}N} \int_{(\varphi+LX_L)'\Psi(\varphi+LX_L)\le(r_0-\alpha)} \exp\left(-\frac{1}{2}X_L'X_L\right) dX_L. \end{aligned} \tag{30}$$

We define the diagonal matrix of eigenvalues of $L'\Psi L$ as Λ_L and its matrix of eigenvectors as U_L. Notice that since $L'\Psi L$ is symmetric, U_L is an orthogonal matrix, that is, $U_L'U_L = I_N$. Therefore $U_L'L'\Psi LU_L = \Lambda_L$. We define a new vector of transformed state variables $Z_L \triangleq U_L'X_L$. Since U_L is an orthonormal matrix, Z_L is also a standard normal multivariate. Then

$$\begin{aligned}(\varphi+LX_L)'\Psi(\varphi+LX_L) &= (\varphi+LU_LZ_L)'\Psi(\varphi+LU_LZ_L) \\ &= \varphi'\Psi\varphi + 2\varphi'\Psi LU_LZ_L + Z_L'U_L'L'\Psi LU_LZ_L. \end{aligned} \tag{31}$$

From $U_L'L'\Psi LU_L = \Lambda_L$, we can show that $\Psi LU_L = L'^{-1}U_L\Lambda_L$ and $\Psi = L'^{-1}U_L\Lambda_LU_L'L^{-1}$. Therefore we can rewrite Equation (31) as

$$\begin{aligned}(\varphi+LX_L)'\Psi(\varphi+LX_L) &= \varphi'L'^{-1}U_L\Lambda_LU_L'L^{-1}\varphi + 2\varphi'L'^{-1}U_L\Lambda_LZ_L + Z_L'\Lambda Z_L \\ &= (Z_L + U_L'L^{-1}\varphi)'\Lambda_L(Z_L + U_L'L^{-1}\varphi). \end{aligned}$$

Hence when $\omega \triangleq [U_L'L^{-1}\varphi]_N$, Equation (30) becomes

$$\begin{aligned}&\Pr[r = \alpha + Y(t)'\Psi Y(t) \le r_0] \\ &\quad = (2\pi)^{-\frac{1}{2}N} \int_{(Z_L+U_L'L^{-1}\varphi)'\Lambda_L(Z_L+U_L'L^{-1}\varphi)\le(r_0-\alpha)} \exp\left(-\frac{1}{2}Z_L'Z_L\right) dZ_L \\ &\quad = \Pr\left[\sum_{i=1}^{N} \lambda_{Li}(Z_{Li} - \omega_i)^2 \le r_0 - \alpha\right]. \end{aligned}$$

The characteristic function of $\mathscr{T}r(Z_L) = \sum_{i=1}^{N} \lambda_{Li}(Z_{Li} - \omega_i)^2$ is known as [see Johnson and Kotz (1970)]:

$$\phi_N(m; \Lambda_L; \omega) = \exp\left(\sum_{j=1}^{N} \frac{\mathrm{im}\,\lambda_{Lj}\omega_j^2}{1 - 2\mathrm{im}\,\lambda_{Lj}}\right) \Pi_{j=1}^{N}(1 - 2\mathrm{im}\,\lambda_{Lj})^{-\frac{1}{2}},$$

where i stands for the imaginary number. Therefore the probability density function of $\mathscr{T}r$ is represented as

$$f_{\mathscr{T}r}(r_0) = (2\pi)^{-1} \int_{-\infty}^{\infty} e^{-\mathrm{im}r_0} \exp\left(\sum_{j=1}^{N} \frac{\mathrm{im}\,\lambda_{Lj}\omega_j^2}{1 - 2\mathrm{im}\,\lambda_{Lj}}\right)$$
$$\times \Pi_{j=1}^{N}(1 - 2\mathrm{im}\lambda_{Lj})^{-\frac{1}{2}}\, dm. \tag{32}$$

Unfortunately the integral in Equation (32) does not allow for a closed-form solution. However, there are many alternative expansion methods which can be used to evaluate Equation (32). For example, Equation (32) can be represented as a mixture of noncentral chi-squared distributions. Thus the cumulative distribution can be written as [see Johnson and Kotz (1970)]

$$\Pr[r = \alpha + Y(t)'\Psi Y(t) \leq r_0] = \sum_{j=0}^{\infty} e_j \left[\chi^2_{N+2j}\left(\sum_{j=1}^{N} \omega_j^2\right) \leq \frac{r_0 - \alpha}{\varepsilon}\right],$$

where ε is any arbitrary constant, and

$$e_j = \begin{cases} \Pi_{k=1}^{N}(\varepsilon/\lambda_{Lk})^{\frac{1}{2}} & \text{if } j = 0 \\ (2j)^{-1}\sum_{k=0}^{j-1} h_{j-k}e_k & \text{if } j \geq 1 \end{cases},$$

which is a solution in a forward direction, and

$$h_j = \begin{cases} \sum_{k=1}^{N}\left(1 - \omega_k^2\right)\left(1 - \frac{\varepsilon}{\lambda_{Lk}}\right) & \text{if } j = 1 \\ \sum_{k=1}^{N}\left(1 - \frac{\varepsilon}{\lambda_{Lk}}\right)^j + k\varepsilon\sum_{k=1}^{N}\left(\frac{\omega_k^2}{\lambda_{Lk}}\right)\left(1 - \frac{\varepsilon}{\lambda_{Lk}}\right)^{j-1} & \text{if } j \geq 2 \end{cases}.$$

However, we have the closed-form expressions for the mean and the variance of the interest rates:

$$E[r] = \alpha + \varphi'\Psi\varphi + \sum_{j=1}^{N} \lambda_{Li} = \alpha + \varphi'\Psi\varphi + 1'_N(V \circ \Psi)1_N$$

$$\mathrm{var}[r] = 4\varphi' L^{'-1} U_L \Lambda_L^2 U_L' L^{-1}\varphi + 2\sum_{j=1}^{N} \lambda_{Li}^2.$$

As a final remark, the above distribution of the nominal interest rate will be reduced to the noncentral chi-squared distribution when $\Lambda_L = I_N$. This occurs when the state variables $Y(t)$ are orthogonal to each other: that is, ξ and Σ are diagonal matrices. In this case, $\mathscr{T}r(Z_L) = \sum_{j=1}^{N}(Z_L - \omega_i)^2$, where Z_L can be easily obtained by demeaning and rescaling the original state variables Y: that is, $L = \mathrm{diag}[1/V_{ii}]_N$. Therefore the SAINTS model specifies the noncentral chi-squared distribution, whereas the general QTSM dictates a much more complicated distribution for the interest rate.

Appendix C: A Supporting Equilibrium of the QTSMs

We assume that the market is complete, which is sufficient for the existence of a representative agent. Following CIR (1985b), we assume further that (1) the utility function of the representative agent is logarithmic: $\mathcal{U}(C(t)) = \exp(\rho)\ln C(t)$ and (2) production is governed by a constant-return-to-scale production technology with the following SDE:[21]

$$\frac{dq(t)}{q(t)} = \left[d + g'Y(t) + Y(t)'HY(t)\right]dt + [n_0 + n_1 Y(t)] \circ dw_N(t),$$

where $d > 0, g > 0$, and H is a positive definite matrix. As such, the expected return on the production technology is governed by a quadratic function of the state variables, while its diffusion is an affine function of the state variables. We denote the indirect utility function of the representative agent by $J(W(t), Y(t), t)$, where $W(t)$ is the wealth of the representative agent or the aggregate wealth of the economy. The assumption of logarithmic utility leads to

$$J_W(W(t), Y(t), t) = \frac{\exp(-\rho t)}{\rho} W_t^{-1},$$

which is independent of the state variables.

Due to the equivalence of the first derivatives of the direct and indirect utility functions, we can verify that the optimal consumption at time t is $C(t)^* = \rho W(t)$ when $\mathcal{T} = \infty$. Then the market clearing condition yields the following wealth process:

$$\frac{dW(t)}{W(t)} = W(t)\left[d + g'Y(t) + Y(t)'HY(t) - \rho\right]dt + W(t)[n_0 + n_1 Y(t)] \circ dw_N(t),$$

since the wealth of the economy is the production output adjusted for the consumption. Finally, the equivalence of the intertemporal marginal rate of substitution of consumption and the intertemporal marginal rate of transformation indicates that the stochastic discount factor implied by the equilibrium is $M(t,T) = J_W(T, W(T), Y(t))/J_W(t, W(t), Y(t))$. Applying Ito's lemma, we can endogenously determine the SDE for the stochastic discount factor:

$$\begin{aligned}
\frac{dM(t)}{M(t)} &= \frac{dJ_W(t, W(t), Y(t))}{J_W(t, W(t), Y(t))} \\
&= \left[1_N'(n_0 + n_1 Y(t))^2 - d - g'Y(t) - Y(t)'HY(t)\right]dt \\
&\quad - [n_0 + n_1 Y(t)] \circ dw_N(t) \\
&= -\left[(d - n_0' n_0) + (g' - n_0' n_1)Y(t) + Y(t)'\left[H - \sum_{i=1}^{N} n_{1i}' n_{1i}\right]Y(t)\right]dt \\
&\quad - [n_0 + n_1 Y(t)] \circ dw_N(t).
\end{aligned}$$

Finally, defining $\alpha \triangleq d - n_0' n_0$, $\beta \triangleq g - n_1' n_0$, $\Psi = H - \sum_{i=1}^{N} n_{1i}' n_{1i}$, $\eta_0 \triangleq -n_0$ and $\eta_1 \triangleq -n_1$ yields the desired result.

Appendix D: Equivalent Representation of the SAINTS Model

We explore the restrictions under which the QTSM is reduced to the SAINTS model. The SAINTS model is based on orthogonal state variables, and also designates no interaction terms

[21] As shown by Longstaff and Schwartz (1992), a single stochastic production technology results in an equivalent model to one with N production technologies.

among the state variables in the determination of the interest rate. Therefore we need the following restrictions:

$$\Phi, \xi, \Sigma, \delta_1 = \text{diagonal matrix}.$$

As such, the bond price formula is separable with respect to the state variables. Thus without loss of generality, we derive the restrictions based on a single-factor case.[22]

Writing the assumptions of a single-factor SAINTS model,

$$dX(t) = -\kappa_x X(t)\, dt + \sigma_x\, dz(t) \tag{33}$$

$$M(t) = \exp[-h\, t + (X(t) - c)^2]. \tag{34}$$

Note that a Wiener process $x_0(t)$ in the stochastic discount factor specified in Constantinides (1992) does not affect bond prices and is put in to allow for nonstationarity of the nominal stochastic discount factor. For brevity we do not incorporate it in Equation (34), since it will not change the solution for the term structure.

The SDE of the stochastic discount factor implied by Equation (34) is

$$\begin{aligned}\frac{dM(t)}{M(t)} &= \left[\sigma_x^2 - h - 2\kappa_x X(t)(X(t) - c) + 2\sigma_x^2 (X(t) - c)^2\right] dt \\ &\quad + 2\sigma_x (X(t) - c)\, dz(t).\end{aligned}$$

From the martingale property of the stochastic discount factor, the drift equals $-r(t)$ and thus

$$\begin{aligned}r(t) &= h - \sigma_x^2 + 2\kappa X(t)(X(t) - c) - 2\sigma_x^2 (X(t) - c)^2 \\ &= 2(\kappa_x - \sigma_x^2)\left[X(t) + \frac{c(2\sigma_x^2 - \kappa_x)}{2(\kappa_x - \sigma_x^2)}\right]^2 - \frac{c^2(2\sigma_x^2 - \kappa_x)^2}{2(\kappa_x - \sigma_x^2)} \\ &\quad + h - \sigma_x^2 - 2c^2\sigma_x^2.\end{aligned}$$

To ensure the positivity of the nominal interest rate, we assume

$$\kappa_x > \sigma_x^2 \text{ and } h > \sigma_x^2 + 2c^2\sigma_x^2 + \frac{c^2(2\sigma_x^2 - \kappa_x)^2}{2(\kappa_x - \sigma_x^2)} = \sigma_x^2 + \frac{c^2\kappa_x^2}{2(\kappa_x - \sigma_x^2)}.$$

Now we define a new state variable $Y(t)$, and the lower bound on the interest rate, α, by using an invariant affine transformation:

$$\begin{aligned}Y(t) &\triangleq \sqrt{2(\kappa_x - \sigma_x^2)}\left[X(t) + \frac{c(2\sigma_x^2 - \kappa_x)}{2(\kappa_x - \sigma_x^2)}\right] \\ \alpha &\triangleq -\frac{c^2(2\sigma_x^2 - \kappa_x)^2}{2(\kappa_x - \sigma_x^2)} + h - \sigma_x^2 - 2c^2\sigma_x^2.\end{aligned}$$

The SDE of the transformed state variable $Y(t)$ is

$$\begin{aligned}dY(t) &= \sqrt{2(\kappa_x - \sigma_x^2)}\, dX(t) \\ &= \kappa_x\left[\frac{c(2\sigma_x^2 - \kappa_x)}{\sqrt{2(\kappa_x - \sigma_x^2)}} - Y(t)\right] dt + \sqrt{2(\kappa_x - \sigma_x^2)}\sigma_x\, dz(t).\end{aligned}$$

[22] Since it is a single factor case, we will suppress the subscript i, an index of the state variable hereafter.

We define

$$\mu \triangleq \frac{\kappa_x c(2\sigma_x^2 - \kappa_x)}{\sqrt{2(\kappa_x - \sigma_x^2)}}, \quad \xi \triangleq -\kappa_x, \quad \sigma \triangleq \sqrt{2(\kappa_x - \sigma_x^2)}\sigma_x. \tag{35}$$

Thus we have transformed the specification for the interest rate and the SDE of the state variable of the SAINTS model into the counterparts of the QTSM. Now we can explore the restrictions on the market price of risk imposed by the SAINTS model. First we solve for σ_x and c as functions of the structural parameters of the QTSM.

From the definition of σ in Equation (35), we can solve a nonlinear equation, which yields the solution

$$\sigma_x = \sqrt{\frac{-\xi \pm \sqrt{\xi^2 - 2\sigma^2}}{2}}. \tag{36}$$

Notice that $\xi^2 - 2\sigma^2 = (\kappa_x - 2\sigma_x^2)^2$, which is asserted to be nonnegative. In addition, the definition of μ in Equation (35) yields

$$c = \frac{\mu\sigma}{-\xi\sqrt{\frac{-\xi \pm \sqrt{\xi^2 - 2\sigma^2}}{2}}(\pm\sqrt{\xi^2 - 2\sigma^2})}. \tag{37}$$

The diffusion term of the stochastic discount factor can be rewritten as

$$\begin{aligned} 2\sigma_x(X(t) - c) &= 2\sigma_x\left[\frac{Y(t)}{\sqrt{2(\kappa_x - \sigma_x^2)}} - \frac{c(2\sigma_x^2 - \kappa_x)}{2(\kappa_x - \sigma_x^2)} - c\right] \\ &= -\frac{c\kappa_x\sigma_x}{\kappa_x - \sigma_x^2} + \frac{2\sigma_x}{\sqrt{2(\kappa_x - \sigma_x^2)}}Y(t). \end{aligned}$$

Using Equations (36) and (37), we can write η_0 and η_1 as a function of the structural parameters of $Y(t)$,

$$\begin{aligned} \eta_0 &= -\frac{c\kappa_x\sigma_x}{\kappa_x - \sigma_x^2} = -\frac{\mu}{\sigma}\left(\frac{-\xi \pm \sqrt{\xi^2 - 2\sigma^2}}{\pm\sqrt{\xi^2 - 2\sigma^2}}\right) \\ \eta_1 &= \frac{2\sigma_x}{\sqrt{2(\kappa_x - \sigma_x^2)}} = \frac{-\xi \pm \sqrt{\xi^2 - 2\sigma^2}}{\sigma}. \end{aligned}$$

Finally, the market price of risk is defined as

$$-\mathrm{cov}_t\left(dY(t), \frac{dM(t)}{M(t)}\right) = -\sigma[\eta_0 + \eta_1 Y(t)].$$

Therefore

$$\begin{aligned}
\delta_0 &= -\sigma\eta_0 \\
&= -\sigma\left[-\frac{\mu}{\sigma}\left(\frac{-\xi \pm \sqrt{\xi^2 - 2\sigma^2}}{\pm\sqrt{\xi^2 - 2\sigma^2}}\right)\right] \\
&= \mu\left(\frac{-\xi \pm \sqrt{\xi^2 - 2\sigma^2}}{\pm\sqrt{\xi^2 - 2\sigma^2}}\right) \\
\delta_1 &= -\sigma\eta_1 \\
&= -\sigma\left[\frac{-\xi \pm \sqrt{\xi^2 - 2\sigma^2}}{\sigma}\right] \\
&= \xi \mp \sqrt{\xi^2 - 2\sigma^2}.
\end{aligned}$$

In summary, there are two possible identifications of the market price of risk imposed by the SAINTS model, which are

	δ_0	δ_1
First case	$\mu\left(\frac{-\xi + \sqrt{\xi^2 - 2\sigma^2}}{\sqrt{\xi^2 - 2\sigma^2}}\right)$	$\xi - \sqrt{\xi^2 - 2\sigma^2}$
Second case	$\mu\left(\frac{\xi + \sqrt{\xi^2 - 2\sigma^2}}{\sqrt{\xi^2 - 2\sigma^2}}\right)$	$\xi + \sqrt{\xi^2 - 2\sigma^2}$

Thus, for a general N-factor model, there are $2N$ alternative forms of restrictions on the market price of risk imposed by the SAINTS model.

Appendix E: Invariant Transformations and Normalizations

The maximally flexible QTSM should be *unique* in the sense that any transformation or rescaling cannot reproduce the SDEs of state variables, the instantaneous interest rate, and hence bond prices. In other words, the full-fledged QTSM should not allow for any *invariant transformation*. Otherwise there exists an *equivalent class* of the maximal model obtained by invariant transformations of the maximal model, and the maximal model itself is underidentified, which invalidates the empirical implementation of the model. As such, we will first examine a wide variety of invariant transformations and then we will show that the maximal model is uniquely defined, that is, robust to invariant transformations.

E.1 Invariant transformations

Following the study of Dai and Singleton (2000), we explore invariant transformations, which refer to transformations and rescaling of state vectors and parameter vectors without changing the instantaneous short rate/bond prices. We consider a QTSM with state vector, Brownian motions, and parameter vectors given by $\theta \triangleq [Y(t), z_N(t), (\alpha, \beta, \Psi, \mu, \xi, \Sigma, \delta_0, \delta_1)]$.

- **An affine transformation:** The only admissible transformation of the state vectors is an affine transformation since any nonaffine transformation of the state variables results in non-Gaussian state variables. An affine transformation refers to $\mathcal{T}_A = \vartheta + \mathcal{L}Y(t)$, where ϑ

is an $N \times 1$ vector and $\mathscr{L}$ is an $N \times N$ nonsingular matrix. This results in

$$\mathcal{T}_A\theta = \begin{bmatrix} \vartheta + \mathscr{L}Y(t) \\ z_N(t) \\ \big(\alpha - \beta'\mathscr{L}^{-1}\vartheta + \vartheta'\mathscr{L}'^{-1}\Psi\mathscr{L}^{-1}\vartheta, \beta'\mathscr{L}^{-1} - 2\vartheta'\mathscr{L}'^{-1}\Psi\mathscr{L}^{-1}, \\ \mathscr{L}'^{-1}\Psi\mathscr{L}^{-1}, \mathscr{L}\mu - \mathscr{L}\xi\mathscr{L}^{-1}\vartheta, \mathscr{L}\xi\mathscr{L}^{-1}, \mathscr{L}\Sigma, \\ \mathscr{L}(\delta_0 - \delta_1\mathscr{L}^{-1}\vartheta), \mathscr{L}\delta_1\mathscr{L}^{-1}\big) \end{bmatrix}.$$

- **An orthonormal rotation:** An orthonormal transformation, $\mathcal{T}_{\mathscr{O}}$, is defined by an $N \times N$ orthonormal matrix $\mathscr{O}$ (i.e., $\mathscr{O}\mathscr{O}' = I_N$), such that

$$\mathcal{T}_{\mathscr{O}}\theta = \big[Y(t), z_N(t), (\alpha, \beta, \Psi, \mu, \xi, \Sigma\mathscr{O}, \Sigma\mathscr{O}\Sigma^{-1}\delta_0, \Sigma\mathscr{O}\Sigma^{-1}\delta_1)\big].$$

  This orthonormal rotation results in the unidentifiability of all N entries in Σ.
- **A permutation:** A permutation, $\mathcal{T}_P$, reorders the state vector since state variables are not observable.

E.2 Normalizations

In order to identify the QTSM in the presence of invariant transformations, we need to impose the following restrictions on the parameters.

E.2.1 Correlation structure of the state variables. Since ξ and Σ jointly determine the correlation structure of the state variables, they are not separately identifiable. Following Dai and Singleton (2000), we assume that Σ is diagonal. In addition, the elements of ξ are not fully identifiable because the correlation matrix is symmetric; the QTSM is invariant with respect to the orthonormal rotation, $\mathcal{T}_O$. Thus we assume that ξ is lower triangular.

E.2.2 Linear invariant transformation. Consider a nonsingular linear transformation $X(t) = \mathcal{T}_A = \mathscr{L}Y(t)$. Then, we can represent the nominal interest rate and the SDE of the rotated state vector $X(t)$:

$$r(t) = \alpha + \beta'\mathscr{L}^{-1}X(t) + X(t)'\mathscr{L}'^{-1}\Psi\mathscr{L}^{-1}X(t)$$
$$dX(t) = \big[\mathscr{L}\mu + \mathscr{L}\xi\mathscr{L}^{-1}X(t)\big]\,dt + \mathscr{L}\Sigma\,dz_N(t).$$

Since Σ is diagonal, the only $\mathscr{L}$ which maintains the orthogonality of the diffusion matrix of $X(t)$, $\mathscr{L}\Sigma$ is an orthogonal matrix. Notice that $\mathscr{L}\xi\mathscr{L}^{-1}$ is still lower triangular, since $\mathscr{L}$ is diagonal. For Ψ to be identifiable, we assume that the diagonal terms of Ψ are 1. Then the only $\mathscr{L}$ which makes $\mathscr{L}'^{-1}\Psi\mathscr{L}^{-1}$ diagonal is either I_N or $-I_N$. Assuming $\mu \geq 0$ asserts that the only admissible $\mathscr{L}$ is I_N.

E.2.3 Level of state variable. Consider $X(t) = \mathcal{T}_A = \vartheta + I_N Y(t)$. Then the nominal interest rate and the SDE of the mean-shifted state vector $X(t)$ are represented as

$$r(t) = (\alpha - \beta'\vartheta + \vartheta'\Psi\vartheta) + (\beta' - 2\vartheta'\Psi)X(t) + X(t)'\Psi X(t)$$
$$dX(t) = \big[(\mu - \xi\vartheta) + \xi X(t)\big]\,dt + \Sigma\,dz_N(t).$$

In order to identify μ, we assume $\beta = 0_N$. These N restrictions result in an identification of μ, since the only admissible ϑ is then 0.

References

Ahn, D.-H., 1995, "A Generalized Squared Autoregressive Intertemporal Term Structure Model," working paper, University of North Carolina, Chapel Hill.

Ahn, D.-H., and B. Gao, 1999, "A Parametric Nonlinear Model of Term Structure Dynamics," *Review of Financial Studies*, 12, 721–762.

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

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

Andersen, T., and J. Lund, 1997, "Estimating Continuous-Time Stochastic Volatility Models of the Short-Term Interest Rate," *Journal of Econometrics*, 77, 343–377.

Balduzzi, P., S. Das, S. Foresi, and R. Sundaram, 1996, "A Simple Approach to Three Factor Affine Models of the Term Structure," *Journal of Fixed Income*, 6, 43–53.

Beaglehole, D., and M. Tenney, 1991, "General Solutions of Some Interest Rate-Contingent Claim Pricing Equations," *Journal of Fixed Income*, 1, 69–83.

Beaglehole, D., and M. Tenney, 1992, "A Nonlinear Equilibrium Model of the Term Structure of Interest Rates: Corrections and Additions," *Journal of Financial Economics*, 32, 345–454.

Chapman, D., J. Long, and N. Pearson, 1999, "Using Proxies for the Short Rate: When Are Three Months Like an Instant?," *Review of Financial Studies*, 12, 763–806.

Chapman, D., and N. Pearson, 2000, "Is the Short Rate Drift Actually Nonlinear?," *Journal of Finance*, 55, 355–388.

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*, Blackwell Publishers, Oxford, UK.

Chen, R.-R., and L. Scott, 1992, "Pricing Interest Options in a Two-Factor Cox-Ingersoll-Ross Model of the Term Structure," *Review of Financial Studies*, 5, 613–636.

Claessens, S., and G. Pennachi, 1996, "Estimating the Likelihood of Mexican Default from the Market Prices of Brady Bonds," *Journal of Financial and Quantitative Analysis*, 31, 109–126.

Constantinides, G., 1992, "A Theory of the Nominal Structure of Interest Rates," *Review of Financial Studies*, 5, 531–552.

Cox, J., J. Ingersoll, and S. Ross, 1985a, "An Intertemporal General Equilibrium Model of Asset Prices," *Econometrica*, 53, 363–384.

Cox, J., J. Ingersoll, and S. Ross, 1985b, "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53, 385–406.

Dai, Q., and K. Singleton, 2000, "Specification Analysis of Affine Term Structure Models," *Journal of Finance*, 55, 1943–1978.

Dai, Q., and K. Singleton, 2001, "Expectations Puzzles, Time-Varying Risk Premia, and Dynamic Models of the Term Structure," working paper, Stanford University.

Duffee, G., 2000, "Term Premia and Interest Rate Forecasts in Affine Models," working paper, University of California, Berkeley.

Duffie, D., and R. Kan, 1996, "A Yield-Factor Model of Interest Rates," *Mathematical Finance*, 6, 379–406.

Duffie, D., and K. Singleton, 1997, "An Econometric Model of the Term Structure of Interest Rates," *Journal of Finance*, 52, 1287–1321.

Gallant, A. R., and J. Long, 1997, "Estimating Stochastic Differential Equations Efficiently by Minimum Chi-Squared," *Biometrika*, 84, 125–141.

Gallant, A. R., and G. Tauchen, 1989, “Seminonparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Implications,” *Econometrica*, 57, 1091–1120.

Gallant, A. R., and G. Tauchen, 1996, “Which Moments to Match,” *Econometric Theory*, 12, 657–681.

Gallant, A. R., and G. Tauchen, 1997, “SNP: A Program for Nonparametric Time Series Analysis,” working paper, Duke University.

Gallant, A. R., and G. Tauchen, 1998, “Reprojecting Partially Observed Systems with Application to Interest Rate Diffusions,” *Journal of the American Statistical Association*, 93, 10–24.

Gourieroux, C., A. Monfort, and E. Renault, 1993, “Indirect Inference,” *Journal of Applied Econometrics*, 8, 85–118.

Hansen, L. P., and S. Richard, 1987, “The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models,” *Econometrica*, 55, 587–614.

Harrison, M., and D. Kreps, 1979, “Martingales and Arbitrage in Multiperiod Security Markets,” *Journal of Economic Theory*, 20, 381–408.

Harrison, M., and S. Pliska, 1981, “Martingales and Stochastic Integrals in the Theory of Continuous Trading,” *Stochastic Processes and Their Applications*, 11, 215–260.

Jegadeesh, N., and G. Pennachi, 1996, “The Behavior of Interest Rates Implied by the Term Structure of Eurodollar Futures,” *Journal of Money, Credit, and Banking*, 26, 426–446.

Johnson, N., and S. Kotz, 1970, *Continuous Univariate Distributions*, Houghton Mifflin, New York.

Karatzas, I., and S. Shreve, 1991, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, New York.

Karuoi, N. E., R. Myneni, and R. Viswanathan, 1992, “Arbitrage Pricing and Hedging of Interest Rate Claims with State Variables: I. Theory,” working paper, Universitè de Paris VI and Stanford University.

Kitagawa, G., 1987, “Non-Gaussian State-Space Modeling of Nonstationary Time Series,” *Journal of the American Statistical Association*, 82, 1032–1063.

Knez, P., R. Litterman, and J. Scheinkman, 1996, “Explorations into Factors Explaining Money Market Returns,” *Journal of Finance*, 49, 1861–1882.

Langetieg, T., 1980, “A Multivariate Model of the Term Structure,” *Journal of Finance*, 35, 71–97.

Longstaff, F., 1989, “A Nonlinear General Equilibrium Model of the Term Structure of Interest Rates,” *Journal of Finance*, 23, 1259–1282.

Longstaff, F., and E. Schwartz, 1992, “Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model,” *Journal of Finance*, 47, 1259–1282.

Lu, B., 1999, “An Empirical Analysis of the Constantinides Model of the Term Structure,” working paper, University of Michigan.

McCnlloch, J. H., and H. Kwon, 1993, “U.S. Term Structure Data, 1947–1991,” working paper, Ohio State University.

Pearson, N., and T.-S. Sun, 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.

Schwarz, G., 1978, “Estimating the Dimension of a Model,” *Annals of Statistics*, 6, 461–464.

Stanton, R., 1997, “A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk,” *Journal of Finance*, 52, 1973–2002.

Sun, T.-S., 1992, “Real and Nominal Interest Rates: A Discrete-Time Model and its Continuous Time Limit,” *Review of Financial Studies*, 5, 581–612.

Vasicek, O., 1977, “An Equilibrium Characterization of the Term Structure,” *Journal of Financial Economics*, 5, 177–188.