

Term Structure Dynamics in Theory and Reality

Qiang Dai and Kenneth Singleton¹

This draft : July 18, 2002

¹Dai is with the Stern School of Business, New York University, New York, NY, dai@stern.nyu.edu. Singleton is with the Graduate School of Business, Stanford University, Stanford, CA 94305 and NBER, ken@future.stanford.edu. We are grateful to Andrew Ang, Pierre Collin-Dufresne, Robert Goldstein, Robert Jarrow, Suresh Sundaresen, Len Umantsev, the referees, and Editor Maureen O'hara for helpful comments and suggestions, and to Mariusz Rabus, Len Umantsev, and Wei Yang for research assistance; and for financial support from the Financial Research Initiative, The Stanford Program in Finance, and the Gifford Fong Associates Fund, at the Graduate School of Business, Stanford University.

Abstract

This paper is a critical survey of models designed for pricing fixed income securities and their associated term structures of market yields. Our primary focus is on the interplay between the theoretical specification of dynamic term structure models and their empirical fit to historical changes in the shapes of yield curves. We begin by overviewing the dynamic term structure models that have been fit to treasury or swap yield curves and in which the risk factors follow diffusions, jump-diffusion, or have “switching regimes.” Then the goodness-of-fits of these models are assessed relative to their abilities to: (i) match linear projections of changes in yields onto the slope of the yield curve; (ii) match the persistence of conditional volatilities, and the shapes of term structures of unconditional volatilities, of yields; and (iii) to reliably price caps, swaptions, and other fixed-income derivatives. For the case of defaultable securities we explore the relative fits to historical yield spreads.

1 Introduction

This paper is a critical survey of models designed for pricing fixed income securities (*FIS*) and their associated term structures of market yields.¹ Our primary focus is on the interplay between the theoretical specification of dynamic term structure models (*DTSMs*) and their empirical fit to historical changes in the shapes of yield curves. With this interplay in mind, we characterize *DTSMs* in terms of three primary ingredients: the *risk-neutral* distribution of the state variables or *risk factors*, the mapping between these risk factors and the short-term interest rate, and the factor risk premiums that (when combined with the first two) allow construction of the likelihood function of the historical bond yields. Then the roles of these ingredients in the goodness-of-fits of several widely studied *DTSMs* are assessed based on their matching of several first and second moment properties of bond yields and the prices of fixed-income derivatives.

We begin in Section 2 with an introduction to arbitrage-free pricing of both default-free and defaultable *FIS*. This discussion is directed more toward the needs of researchers interested in the empirical implementation of pricing models, than toward the mathematical details of valuation.² This is followed in Section 3 by a review of the key ingredients of several of the most widely studied continuous-time econometric specifications of *DTSMs*, including affine, quadratic-Gaussian, and non-affine stochastic volatility models. For the case of defaultable bonds, we review both “reduced-form” and “structural” models. Particular attention is given to the trade-offs between analytic tractability and richness of the yield distributions that can be achieved within these modeling frameworks. We conclude this section with a discussion of pricing in the presence of “regime shifts.”

The empirical fit of *DTSMs* to treasury and swap yield data is explored in Section 4. Using model-implied yields simulated from several popular families of *DTSMs*, we assess their goodness-of-fit by the degree to which they replicate: (1) historical regressions showing that holding period returns on bonds are predictable using yield-curve variables (e.g., Fama and Bliss [1987], Campbell and Shiller [1991]); and (2) the conditional volatilities of yields tend to be highly persistent (e.g., Brenner, Harjes, and Kroner [1996]) and the “term structure” of unconditional volatilities is hump-shaped. In Section 5 we discuss the empirical fit of *DTSMs* for defaultable securities, placing particular emphasis on whether these *DTSMs* generate credit spreads consistent with historical experience.

In Section 6 we look beyond the spot market to see how well the model-implied prices of fixed-income derivatives match up with their historical counterparts (implied volatilities). In the process, we are led to address the fit of *DTSMs* to two additional historical observations: (3) the correlations of the changes in the slopes of non-overlapping yield curve segments are small compared to the correlations of changes in the yields themselves (e.g., Rebonato and Cooper [1997]); and (4) implied volatilities of caps and swaptions appear to show variation that is independent of the variation in the underlying swap market (Heidari and Wu [2001])

¹*FIS* are contingent claims with promised cash flows that are known contractually at the inception of the contract. Issuers of *FIS* often have multiple securities outstanding with different maturities that are otherwise homogeneous with regard to their characteristics, in which case there is an (issuer-specific) *term structure* of market yields.

²Recent, more mathematically oriented, surveys of the theoretical term structure literature can be found in Back [1996], Sundaresan [2000], Gibson, LHabitant, and Talay [2001], and Yan [2001].

and Collin-Dufresne and Goldstein [2001a]).

There are several important segments of the fixed-income literature that we have chosen to omit from this survey in order to keep its scope manageable. Specifically, we largely restrict our attention to dynamic pricing models that have examined features of the *joint* distribution of long- and short-term bond yields in estimation and testing. This means that no attempt is made to systematically review the vast literature on descriptive, time-series models of interest rates (including the literature on short-term rates).³ Similarly, we discuss the literature on pricing fixed-income derivatives using *DTSMs*, but do not explore in depth the various specialized “pricing measures” used for pricing LIBOR-based derivatives, taking the current yield curve as given.⁴

2 Specification of *DTSMs*

This section overviews the conceptual foundations of modeling the dynamic properties of yield curves – specifying a *DTSM* in the sense used here. For ease of exposition, and because of its prominence in the theoretical literature on *DTSMs*, we assume that the state vector Y follows a continuous-time diffusion. Extensions of *DTSMs* to accommodate jumps and regime shifts are discussed in Section 3.

DTSMs are typically constructed from the following three key ingredients:

$I_{\mathbb{M}(P)}$: the time-series process for Y under the *pricing* measure $\mathbb{M}(P)$, induced by some numeraire price P ;

$I_{\mathbb{P}}$: the time-series process for Y under the *actual* measure \mathbb{P} ;

I_r : the functional dependence of $r(t)$ on Y .

$I_{\mathbb{M}(P)}$ and I_r are used for pricing and $I_{\mathbb{P}}$ and I_r are used to construct the moments of bond returns (under \mathbb{P}) used in estimation, so all three ingredients are essential for econometric analyses of *DTSMs*. The requirement that there be no arbitrage opportunities within a *DTSM* places only weak restrictions on the choice of $(I_{\mathbb{M}(P)}, I_{\mathbb{P}}, I_r)$. However, the computational demands of both pricing bonds and maximizing the estimation criterion function have typically led researchers to choose $(I_{\mathbb{M}(P)}, I_r)$ so that closed- or essentially closed-form solutions are obtained for zero-coupon bond prices. Additionally, with $(I_{\mathbb{M}(P)}, I_r)$ in hand, the specification of the risk premiums that link $I_{\mathbb{M}(P)}$ and $I_{\mathbb{P}}$ have often been chosen to maintain tractability in estimation, while preserving the condition of no arbitrage opportunities. We elaborate on these specification issues in Section 3.

³See Chapman and Pearson [2001] for a survey with extensive coverage of empirical studies of short-rate models.

⁴Musiela and Rutkowski [1997] and Dai and Singleton [2002a] survey large portions of the literatures on the “forward-rate” models of Heath, Jarrow, and Morton [1992], Brace, Gatarek, and Musiela [1997], and Miltersen, Sandmann, and Sondermann [1997].

2.1 DTSMs of Default-Free Bond Yields

Throughout this survey, we will use the concept of a pricing kernel, M , to describe the pricing mechanism. Initially, we specialize to a setting where the state of the economy is completely described by a Markovian state vector $Y(t)$, with

$$dY(t) = \mu_Y^{\mathbb{P}}(Y, t) dt + \sigma_Y(Y, t) dW(t), \quad (1)$$

where $\mu_Y^{\mathbb{P}}$ is an $N \times 1$ vector of drifts under \mathbb{P} and σ_Y is an $N \times N$ state-dependent factor-volatility matrix. In this diffusion setting, M_t can be written generically as

$$\frac{dM_t}{M_t} = -r_t dt - \Lambda_t' dW(t), \quad (2)$$

where $r_t = r(Y(t), t)$ is the instantaneous riskless rate, $W(t)$ is a vector of N independent Brownian motions, and $\Lambda_t = \Lambda(Y(t), t)$ is the N -vector of *market prices of risk*. For simplicity, we take the risk factors driving M and Y to be one and the same.

For a *FIS* with a dividend rate $h(Y(t), t)$ for $t \leq T$ and terminal payoff $g(Y(T))$ at date T , its price at date $t \leq T$ can be expressed in terms of the pricing kernel as

$$P(Y(t), t) = E_t \left[\int_t^T \frac{M(s)}{M(t)} h(Y(s), s) ds \right] + E_t \left[\frac{M(T)}{M(t)} g(Y(T)) \right], \quad (3)$$

where E_t denotes expectation conditioned on date t information under \mathbb{P} . The solution to (3) can be expressed equivalently as the solution to the following fundamental partial differential equation (*PDE*) (e.g., Duffie [1996])

$$\left[\frac{\partial}{\partial t} + \mathcal{A} \right] P_t - r_t P_t + h_t = 0, \quad (4)$$

where \mathcal{A} is the infinitesimal generator

$$\mathcal{A} = (\mu_t^{\mathbb{P}} - \sigma_{Yt} \Lambda_t)' \frac{\partial}{\partial Y_t} + \frac{1}{2} \text{Trace} \left[\sigma_{Yt} \sigma_{Yt}' \frac{\partial^2}{\partial Y_t \partial Y_t'} \right]. \quad (5)$$

Λ in this construction is interpreted as the market price of risk, because the expected excess return from holding a *FIS* with price P is given by

$$e_P(t) \equiv E \left[\frac{dP(t) + h(t)dt}{P(t)dt} - r_t \mid \mathcal{F}_t \right] = \frac{1}{P(t)} \frac{\partial P(t)}{\partial Y(t)'} \sigma_Y(t) \Lambda_t, \quad (6)$$

where \mathcal{F}_t denotes agents' information set at date t . The term pre-multiplying $\Lambda(t)$ in (6) is the volatility of P induced by volatility in Y . Thus, Λ is the vector of risk premiums required for each unit of volatility of the N risk factors.

The practical problem faced by researchers pricing *FIS* and their associated derivatives is one of computing the expectations in (3). The most widely applied approach to solving (3)

(equivalently (4)) involves a “change of measure” to obtain a more tractable expectation.⁵ To illustrate this approach we fix $T > 0$ and let $Z(t)$ and $P(t)$ denote the prices of two traded securities at date $t < T$ that, for simplicity, have no cash flows prior to date T . We view $P(t)$ as a numeraire price that defines an associated measure $\mathbb{M}(P)$ in the following sense. Letting $V(t) = \frac{Z(t)}{P(t)}$, from the counterparts of (4) for $P(t)$ and $Z(t)$ it follows that

$$0 = V_t + \mu_Y^{\mathbb{M}(P)'} V_Y + \frac{1}{2} \text{Tr} [\sigma_Y \sigma_Y' V_{YY'}], \quad \text{where } \mu_Y^{\mathbb{M}(P)} = \mu_Y^{\mathbb{P}} - \sigma_Y [\Lambda - \sigma_P] \quad (7)$$

and $\sigma_P \equiv (\sigma_Y' \partial P / \partial Y) / P$. Under regularity, the Feynman-Kac theorem applied to (7) implies

$$V(t) = E_t^{\mathbb{M}(P)} [V(T)] \Leftrightarrow \frac{Z(t)}{P(t)} = E_t^{\mathbb{M}(P)} \left[\frac{Z(T)}{P(T)} \right], \quad (8)$$

where the conditional expectation is taken with respect to a measure $\mathbb{M}(P)$ under which Y follows the process $dY(t) = \mu_Y^{\mathbb{M}(P)}(t) dt + \sigma_Y(t) dW^{\mathbb{M}(P)}(t)$, with $W^{\mathbb{M}(P)}$ a vector of standard Brownian motions under measure $\mathbb{M}(P)$. Under the measure $\mathbb{M}(P)$, the short rate r does not appear (in (7) and (8)) and the relative price $V(t)$ follows a Martingale. Different choices of $P(t)$ lead to different $\mu_Y^{\mathbb{M}(P)}(t)$ and therefore different *pricing measures* $\mathbb{M}(P)$.

Two choices of numeraire $P(t)$, that lead to two general-purpose pricing measures, are:

Risk-Neutral Measure $\mathbb{Q} = \mathbb{M}(e^{\int_0^t r(s) ds})$: $P(t)$ is the price of a continuously compounded bank deposit (that is, $P(t) = e^{\int_0^t r(s) ds}$), for which $\sigma_P = 0$ and the \mathbb{Q} -drift of Y is $\mu_Y^{\mathbb{Q}}(t) \equiv \mu_Y^{\mathbb{P}}(t) - \sigma_Y(t) \Lambda(t)$. From (8),

$$Z(t) = E_t^{\mathbb{Q}} \left[e^{-\int_t^T r(s) ds} Z(T) \right]; \quad (9)$$

$Z(t)$ is the present value of $Z(T)$ discounted by the riskless rate, so \mathbb{Q} is referred to as the *risk-neutral* measure.

Forward Measure $\mathbb{Q}^T \equiv \mathbb{M}(D(t, T))$: $P(t) = D(t, T)$, the price of a zero-coupon bond issued at date t and maturing at date T . In this case, $V(t)$ is a “forward price” of security $Z(t)$, so \mathbb{Q}^T is commonly referred to as the *forward measure*. Using the fact that $D(T, T) = 1$, (8) becomes

$$Z(t) = D(t, T) E_t^{\mathbb{Q}^T} [Z(T)], \quad (10)$$

where $E_t^{\mathbb{Q}^T} [\cdot]$ denotes conditional expectation under the measure \mathbb{Q}^T .

⁵An alternative approach, besides the direct application of numerical methods, is to derive the so-called *Green's function* $G(t, Y(t)|T, Y)$ which gives the price $P(t)$ of any *FIS* with dividend process $h(Y(t), t)$ and terminal payoff $g(Y(T))$ as

$$P(t) = \int_t^T ds \int dy G(t, Y(t)|s, y) h(y, s) + \int dy G(t, Y(t)|T, y) g(y).$$

Unfortunately, the derivation of the Green's function in analytical form or its numerical computation is often a non-trivial matter. See Steenkiste and Foresi [1999] for a discussion of the computational issues related to this approach for the case of affine jump-diffusions and Collin-Dufresne and Goldstein [1999] for the case of reflecting boundaries for the short-rate process r .

Which measure is chosen in practice, either among these two or with some other numeraire, depends on a researcher’s objective.

The financial industry tends to have a cross-sectional, as opposed to time-series, focus given the practical demands of “point-in-time” pricing systems. The numeraire, and associated measure $\mathbb{M}(P)$, are often chosen to give convenient closed-form or numerical solutions for (8) for various option payoffs $Z(T)$. Particularly in the case of such LIBOR-based instruments as caps, floors, and swaptions, pricing has tended to focus on forward measures \mathbb{Q}^T , with the numeraire chosen to be either the price of a LIBOR-based zero-coupon bond or a swap price. Of course, if two derivatives based on the same underlying risk factors are priced with different numeraires, then the resulting pricing models may be based implicitly on mutually inconsistent assumptions about the distributions of the risk factors (Brace, Gatarek, and Musiela [1997], Jamshidian [1997]).

Another issue that arises in practice is that trading desks often require that a model correctly “price” an entire yield curve before it will be used for pricing derivatives based on this curve. This consideration in part underlies the wide-spread use of forward-rate based models, which prescribe the risk-neutral dynamics of the forward curve (as in Heath, Jarrow, and Morton [1992]). In such models, the forward curve $f(t, \cdot)$, defined by $f(t, T) = -\partial \log D(t, T) / \partial T$, for any $T \geq t$, is an (observable) input into an arbitrage free pricing model. As typically implemented in industry, forward-rate models are silent about the time-series behavior of yields under \mathbb{P} and, therefore, are not within the family of *DTSMs* explored in depth in this survey.

In the *DTSMs* with fixed parameters explored subsequently, the dimension of Y , N , is typically small relative to the number of securities to be priced by M . One means of circumventing this limitation of dimensionality in yield-based models (those based on $(I_{\mathbb{Q}}, I_r)$), is to introduce time-dependent parameters that allow for point-in-time calibration of a low-dimensional factor model to an entire yield curve of spot yields or volatilities. (This is an easy “add-on” in most of the *DTSMs* discussed subsequently.) This practice is not without controversy, since recalibrating the parameters as the shape of the underlying yield curve or option volatilities change amounts to “changing the model.” Therefore, the resulting models are almost surely fraught with arbitrage opportunities from a dynamic perspective (Backus, Foresi, and Zin [1998], Buraschi and Corielli [2000], Brandt and Yaron [2001]).

In principle, these issues disappear if the entire yield curve is viewed as the state vector and its dynamic properties are explicitly modeled under both \mathbb{Q} and \mathbb{P} . Such high (possibly infinite) dimensional models, developed under the labels of “Brownian sheets” (Kennedy [1994]), “random fields” (Goldstein [2000]), and “stochastic string shocks” (Santa-Clara and Sornette [2001]), are discussed in more depth in Section 6.

2.2 *DTSMs* of Defaultable Bond Yields

If the issuer of an *FIS* might default prior to the maturity date T then, in addition to the risk of changes in r , both the magnitude and the timing of payoffs to investors may be uncertain. The effects of default risks on prices depends on how the default event is defined and the specification of recovery w_τ in the event of a default event. At the broadest level, the two most commonly studied default processes are those of *reduced-form* and *structural* models.

2.2.1 Reduced-Form Models

Reduced-form models treat default as governed by a counting (jump) process $z(t)$ with associated (possibly state-dependent) intensity process $\lambda^{\mathbb{P}}(t)$ and, as such, whether or not an issuer actually defaults is an *unpredictable* event. For pricing in this setting, we extend the formulation (2) of the pricing kernel to allow for a “jump” to the absorbing default state

$$\frac{dM_t}{M_t} = -r_t - \Lambda'_t dW_t - \Gamma_t(dz_t - \lambda_t^{\mathbb{P}} dt), \quad (11)$$

where $\Gamma_t = \Gamma(Y_t)$ is the market price of default risk, and let $w(Y(t))$ denote the recovery by holders of a *FIS* in the event of default. For a defaultable zero-coupon bond, issued at t and maturing at date T , with price $B(t, T)$, the absence of arbitrage opportunities implies that $B(t, T)M_t$ is a Martingale. This, in turn, implies that

$$\left[\frac{\partial}{\partial t} + \mathcal{A} \right] B(t, T) - (r_t + \lambda_t^{\mathbb{Q}})B(t, T) + w_t \lambda_t^{\mathbb{Q}} = 0, \quad (12)$$

where $\lambda_t^{\mathbb{Q}} = (1 - \Gamma_t)\lambda_t^{\mathbb{P}}$ is the “risk-neutral” intensity of arrival of default.

Comparing (12) with (4) we see that the defaultable security is priced using the default-adjusted discount rate $r_t + \lambda_t^{\mathbb{Q}}$. Intuitively, this is a consequence of the risk-neutral probability of an issuer surviving from date t to date s (presuming survival to date t) being $E_t^{\mathbb{Q}} [\exp\{-\int_t^s \lambda_u^{\mathbb{Q}} du\}]$. Using $r_t + \lambda_t^{\mathbb{Q}}$ in discounting accounts for both the time value of money and the need for the issuer to survive to receive payments. Moreover, we see that, even though we are pricing a zero-coupon bond, the possibility of a recovery in the event of default effectively introduces a dividend that is received at the rate $w_t \lambda_t^{\mathbb{Q}}$ (compare with h in (4)). Since $\lambda_t^{\mathbb{Q}} dt$ is the probability of default over the next instant of time and w_t is the recovery in the event of default, the dividend is the mean recovery rate due to default.

Importantly, as discussed in Artzner and Delbaen [1995], Jarrow, Lando, and Yu [2000] and Martellini and Karoui [2001], the requirement of no arbitrage places only weak restrictions on the risk premium $\Gamma(t)$ and, hence, on the mapping between $\lambda^{\mathbb{Q}}$ and $\lambda^{\mathbb{P}}$. Not only may $\lambda^{\mathbb{P}}$ and $\lambda^{\mathbb{Q}}$ differ in their current levels, they may also have different degrees of persistence, time-varying volatility, and one might jump while the other follows a continuous sample path. Thus, moving between $\lambda^{\mathbb{P}}$ and $\lambda^{\mathbb{Q}}$ is not analogous to the standard adjustment to the drift of r to obtain its risk-neutral representation.

Moreover, the instantaneous excess return on a defaultable zero-coupon depends on $\lambda^{\mathbb{P}}$ and *not* on $\lambda^{\mathbb{Q}}$:

$$e_{Bt} = \frac{1}{B(t, T)} \frac{\partial B(t, T)}{\partial Y'} \sigma_Y \Lambda_t + \frac{w_t - B(t, T)}{B(t, T)} \lambda_t^{\mathbb{P}} \Gamma_t. \quad (13)$$

Compared to equation (6), e_{Bt} has an extra component, $\frac{w_t - B(t, T)}{B(t, T)} \lambda_t^{\mathbb{P}} \Gamma_t$, representing compensation for the expected loss due to default. This component is the product of $\frac{w_t - B(t, T)}{B(t, T)}$, the percentage loss of value due to default; $\lambda_t^{\mathbb{P}}$, the actual default intensity; and Γ_t , the market price of default risk. Since bond prices reveal information only about $\lambda^{\mathbb{Q}}$, computing $\lambda^{\mathbb{P}}$ and e_{Bt} typically requires additional information about the \mathbb{P} -likelihood of an issuer defaulting.

The solution to (12) for $B(t, T)$ depends on what one assumes about recovery, w_t . Duffie and Singleton [1999] assume that investors lose an expected (risk-neutral) fraction $L_t^{\mathbb{Q}}$ of the market value of $B(t, T)$, measured just prior to the default event (fractional recovery of market value). In this case, $w_t = (1 - L_t^{\mathbb{Q}})B(t, T)$ and $B(t, T)$ solves the *PDE*

$$\left[\frac{\partial}{\partial t} + \mathcal{A} \right] - (r(t) + L^{\mathbb{Q}}(t)\lambda^{\mathbb{Q}}(t))B(t, T) = 0. \quad (14)$$

It follows that $B(t, T) = E_t^{\mathbb{Q}} \left[e^{-\int_t^T R_u du} \right]$, where $R_t \equiv r_t + \lambda_t^{\mathbb{Q}} L_t^{\mathbb{Q}}$ denotes the “default-adjusted” discount rate. The price of the zero-recovery security studied by Lando [1998] and Madan and Unal [1998] is obtained as the special case with $L^{\mathbb{Q}} = 1$.

Alternatively, Lando [1998] and Duffie [1998] assume recovery of w_τ at the time of default which (by analogy to (3)) leads to the pricing relation

$$B(t, T) = E_t^{\mathbb{Q}} \left[e^{-\int_t^T (r_s + \lambda_s^{\mathbb{Q}}) ds} \right] + E_t^{\mathbb{Q}} \left[\int_t^T e^{-\int_t^u (r_s + \lambda_s^{\mathbb{Q}}) ds} \lambda_u^{\mathbb{Q}} w_u du \right]. \quad (15)$$

With the face value of this bond normalized to unity, this recovery assumption is interpretable as fractional recovery of face value. Madan and Unal [1998] derive similar pricing relations for the case of junior and senior debts with different recovery ratios.

Finally, Jarrow and Turnbull [1995] assume a constant fractional recovery of an otherwise equivalent treasury security with the remaining maturity of the defaultable instrument. This recovery assumption has been less widely applied in the empirical literature than the preceding two recovery assumptions.

2.2.2 Structural Models

Structural models, in their most basic form, assume default at the first time that some credit indicator falls below a specified threshold value. The conceptual foundations for this approach were laid by Black and Scholes [1973] and Merton [1970, 1974]. They supposed that default occurs at the maturity date of debt provided the issuer’s assets are less than the face value of maturing debt at that time. (Default before maturity was not considered.) Black and Cox [1976] introduced the idea that default would occur at the first time that assets fall below a boundary D (which may or may not be the face value of debt), thereby turning the pricing problem into one of computing “first-passage” probabilities. For the case of exogenously given default boundary F , if in the event of default bondholders lose the fraction $L_T^{\mathbb{Q}}$ of par *at maturity*, then the price $B(t, T)$ becomes

$$B(t, T) = E_t^{\mathbb{Q}} \left[e^{-\int_t^T r_u du} (1 - L_T^{\mathbb{Q}} 1_{\{\tau < T\}}) \right] = D(t, T) \left[1 - L_T^{\mathbb{Q}} H^T(A_t/F, r_t, T - t) \right], \quad (16)$$

where $H^T(A_t/F, r_t, T - t) \equiv E_t^T[1_{\{\tau < T\}}]$ is the first-passage probability of default between dates t and T under the *forward* measure induced by the default-free zero price $D(t, T)$. Thus, $B(t, T)$ is the price of a riskless zero-coupon bond minus the value of a put option on the value of the firm. The value of a coupon bond in this setting is obtained by summing the present values of the promised coupons discounted by the relevant zero prices $B(t, T)$ on the coupon dates.

Pricing in models with endogenous default thresholds has been explored by Geske [1977], Leland [1994], Leland and Toft [1996], Anderson and Sundaresan [1996], Mella-Barral and Perraudin [1997], and Ericsson and Reneby [2001], among others. The endogeneity of F arises (at least in part) because equity holders have an option as to whether to issue additional equity to service the promised coupon payments. With F determined by the actions of equity holders and debtors, it becomes a function of the underlying parameter of the structural model. The models of Anderson and Sundaresan [1996], Mella-Barral and Perraudin [1997], and Ericsson and Reneby [2001] accommodate violations of absolute priority rules (equity holders experience non-zero recoveries, even though bond holders recover less than the face value of their debt).

2.2.3 Pricing with Two-Sided Default Risk

For the cases of interest rate forward and swap contracts, default risk is “two-sided” in the sense that a financial contract may go “into the money” to either counterparty, depending on market conditions, and so the relevant default processes for pricing change with market conditions. Duffie and Singleton [1999] show, in the context of reduced-form models, that this dependence of λ^Q and w_τ on the price $P(t)$ of the contract being valued renders the preceding reduced-form pricing models inapplicable, at least in principle.⁶ Fortunately, for at-the-money swaps (those used most widely in empirical studies of *DTSMs*), these considerations are negligible (Duffie and Huang [1996] and Duffie and Singleton [1997]). Hence, standard practice within academia and the financial industry is to treat such interest rate swaps as if they are bonds trading at par (\$1), with the discount rate R chosen to reflect the credit/liquidity risk inherent in the swap market.

Using this approximate pricing framework, the resulting discount curve, $-\log B(t, T)/(T-t)$, “passes through” short-term LIBOR rates. However, there is no presumption that long-term swap rates and LIBOR contracts reflect the same credit quality. They are in fact notably different (Sun, Sundaresan, and Wang [1993], Collin-Dufresne and Solnik [2000]). Nor is there a presumption that $R_t - r_t$ reflects only credit risk; liquidity risk may be as, if not more, important (see Grinblatt [1994] and Liu, Longstaff, and Mandell [2001] for discussions of liquidity factors in swap pricing).

3 Term Structure Models

The central role of *DTSMs* in financial modeling has led to the development of an enormous number of models, many of which are not nested. Initially, we focus on the case where Y follows a diffusion process and discuss four of the most widely studied families of *DTSMs*: affine, quadratic-Gaussian (*QG*), non-affine stochastic volatility, and structural defaultable bond pricing models. Then we step outside of the diffusion framework and discuss *DTSMs* with jumps and multiple “regimes.”

⁶ Under fractional recovery of market value, one can still express $B(t, T)$ as the solution to (14) but, because of the dependence of R on P , $B(t, T)$ solves a quasi-linear equation, instead of a more standard linear *PDE*, and prices must be obtained by numerical methods.

3.1 Affine Term Structure Models

The ingredients of affine term structure models are:

$I_{\mathbb{Q}}(A)$: Under \mathbb{Q} , the drift and volatility functions of the risk factors satisfy

$$\mu_Y^{\mathbb{Q}}(t) = \lambda_0 + \lambda_Y Y(t), \quad (17)$$

$$\sigma_Y(t)\sigma_Y(t)' = g_0 + \sum_{i=1}^N g_i Y_i(t), \quad (18)$$

where λ_0 is an $N \times 1$ vectors and λ_Y and g_i , $i = 0, 1, \dots, N$, are $N \times N$ matrices of constants. An equivalent characterization of $\sigma_Y(t)$ has $\sigma_Y(t) = \Sigma \sqrt{S(t)}$, where

$$S_{ii}(t) = \alpha_i + \beta_i' Y(t), \quad S_{ij}(t) = 0, \quad i \neq j, 1 \leq i, j \leq N, \quad (19)$$

and Σ is an $N \times N$ matrix of constants.

$I_{\mathbb{P}}(A)$: Given $\sigma_Y(t)$ satisfying (18), the requirement (17) determines the drift of Y under the actual measure, $\mu_Y^{\mathbb{P}}(t)$, once the market prices of risk $\Lambda(t)$ are specified, and vice-versa.

$I_r(A)$: The short rate is an affine function of Y :

$$r(t) = \delta_0 + \delta_Y' Y(t). \quad (20)$$

Duffie and Kan [1996] show that under $(I_{\mathbb{Q}}(A), I_r(A))$, the solution to the *PDE* (4) for $D(t, T)$ is exponentially affine:

$$D(t, T) = e^{\gamma_0(T-t) + \gamma_Y(T-t)' Y(t)}, \quad (21)$$

where γ_0 and γ_Y satisfy known ordinary differential equations (*ODEs*). Note that, in deriving the pricing relation (21), Duffie and Kan were silent on the properties of Y under \mathbb{P} . To obtain (21), essentially any specification of $\mu_Y^{\mathbb{P}}(t)$ and any arbitrage-free specification of $\Lambda(t)$ can be chosen, so long as the drift of Y under \mathbb{Q} is an affine function of Y .

In order for affine specifications to be *admissible*, restrictions must be imposed on the parameters to assure that the $S_{ii}(t) \geq 0$. To address this problem, for the case of N risk factors and $M (\leq N)$ factors driving the $S_{ii}(t)$, Dai and Singleton [2000] (hereafter *DS*) introduced the ‘‘canonical’’ model $\mathbb{A}_M(N)$ with $S_{ii}(t) = \sqrt{Y_i(t)}$, $i = 1, \dots, M$, and the remaining $N - M$ $S_{ii}(t)$ being affine functions of $(Y_1(t), \dots, Y_M(t))$. *DS* provide an easily verifiable set of sufficient restrictions on the parameters of $\mathbb{A}_M(N)$ to guarantee admissibility. The subfamily $A_M(N)$ ($M = 0, \dots, N$) of affine models was then defined to be all models that are nested special cases of the M^{th} canonical model or of *invariant* transformations of this model. Duffie, Filipovic, and Schachermayer [2002] show that *DS*’s canonical affine diffusion gives the most flexible affine *DTSM* on the state space $\mathbb{R}^M \times \mathbb{R}^{N-M}$. The union $\cup_{M=0}^N A_M(N)$ does not encompass all admissible N -factor affine models, however.

DS analyze the ‘‘completely affine’’ class of *DTSMs* with

$$\mu_Y^{\mathbb{P}}(t) = \kappa(\theta - Y(t)) \text{ and } \Lambda(t) = \sqrt{S(t)}\lambda_1, \quad (22)$$

where λ_1 is an $N \times 1$ vector of constants. In this case, both the \mathbb{P} -drift $\mu_Y(t)$ and \mathbb{Q} -drift $\mu^{\mathbb{Q}} = \mu_Y^{\mathbb{P}}(t) - \sigma_Y(t)\Lambda(t)$ are affine in $Y(t)$. A potentially important limitation of the specification (22) of Λ is that temporal variation in the instantaneous expected excess return on a $(T - t)$ -period zero bond, $e_D(t, T) = \gamma_Y(T - t)\Sigma S_t \lambda_1$ (see (6)), is determined entirely by the volatilities of the state variables through $S(t)$. Moreover, the sign of each $\Lambda_i(t)$ is fixed over time by the sign of λ_{1i} .

Duffee [2002] proposed the more flexible “essentially affine” specification of $\Lambda(t)$ that has

$$\Lambda_t = \sqrt{S_t} \lambda_1 + \sqrt{S_t^-} \lambda_2 Y_t, \quad (23)$$

where λ_1 is an $N \times 1$ vector and λ_2 is an $N \times N$ matrix, and⁷ $S_{ii,t}^- = (\alpha_i + \beta_i' Y_t)^{-1}$, if $\inf(\alpha_i + \beta_i' Y_t) > 0$, and zero otherwise. Within the canonical model for $A_M(N)$, the *inf* requirement implies that the first M rows of λ_2 are zero (corresponding to the M volatility factors). Thus, when $M = N$ (multi-factor *CIR* models), the “completely” and “essentially” affine specifications are equivalent – excess returns vary over time only because of time-variation in the factor volatilities.

However, when $M < N$, the essentially affine specification introduces the possibility that Y affects expected excess returns both indirectly through the $S_{ii}(t)$ and directly through the non-zero elements of $\lambda_2 Y_t$. Moreover, the signs of the $\Lambda_i(t)$ corresponding to the $N - M$ non-volatility factors may switch signs over time. (Those of the first M volatility factors have fixed signs as in *CIR*-style models.) The smaller is M , the more added flexibility is introduced by (23) over (22), though at the expense of less flexibility in matching stochastic volatility. The specification (23) preserves the property that the drifts of Y are affine under both \mathbb{Q} and \mathbb{P} .

Duarte [2001] extended the specification of Λ further to

$$\Lambda(t) = \Sigma^{-1} \lambda_0 + \sqrt{S_t} \lambda_1 + \sqrt{S_t^-} \lambda_2 Y_t, \quad (24)$$

where λ_0 is an $N \times 1$ vector of constants. The practical import of his extension is that the $\Lambda_i(t)$ corresponding to the M volatility factors in an $A_M(N)$ model may switch signs over time. Additionally, larger differences between the drifts of Y under the \mathbb{P} and \mathbb{Q} measures are accommodated, because $\mu^{\mathbb{P}}(t) = \kappa(\theta - Y(t)) + \Sigma \sqrt{S(t)} \Sigma^{-1} \lambda_0$ in Duarte’s model, as prescribed by (17). With this modification, Y follows a non-affine process under \mathbb{P} – the drift involves both the level and square-root of the state variables – but one that is nevertheless affine under \mathbb{Q} (so that the pricing relation (21) continues to hold).

Though the characterization of the families $A_M(N)$ using $(I_{\mathbb{Q}}(A), I_{\mathbb{P}}(A), I_r(A))$ is entirely in terms of the risk factors Y , each $A_M(N)$ is observationally equivalent to a large class of multi-factor models in which r remains one of the state variables, with the drift and conditional variance of r depending on other state variables (*DS*).⁸ For instance, the two-factor Gaussian (“Vasicek”) model studied by Langetieg [1980] is equivalent to the two-factor

⁷ The requirement that the i^{th} diagonal element of S_t^- be nonzero only if $\inf(\alpha_i + \beta_i' Y_t) > 0$ is necessary to rule out arbitrage that might arise if elements of Λ_t approach infinity as $(\alpha_i + \beta_i' Y_t)$ approaches zero (Cox, Ingersoll, and Ross [1985], Duffee [2002]).

⁸One implication of this equivalence is that many of the affine *DTSMs* examined in the literature are nested in the framework (17) - (22). Included in this set are those in Vasicek [1977], Langetieg [1980], Cox,

central-tendency model of r proposed by Beaglehole and Tenney [1991] and Balduzzi, Das, and Foresi [1998].

The development of affine models of defaultable bond prices proceeds in a similar manner. In the Duffie-Singleton framework with fractional recovery of market value, the default adjusted discount rate $R_t = r_t + \lambda_t^{\mathbb{Q}} L_t^{\mathbb{Q}}$ can be modeled as an affine function of the state Y_t . A researcher has the choice of modeling R directly or of building up a model of R from separate affine parameterizations of r and $\lambda^{\mathbb{Q}} L^{\mathbb{Q}}$.

Similarly, an affine model under fractional recovery of face value is obtained by assuming that $(r_t + \lambda_t^{\mathbb{Q}})$, $\lambda_t^{\mathbb{Q}}$, and $\ln w_\tau$ are affine functions of Y . The price of \$1 contingent on survival to date T (the first term in (15)) is known in closed form since $r_t + \lambda_t^{\mathbb{Q}}$ is an affine function of Y . The recovery claim is priced using the extended transform of Duffie, Pan, and Singleton [2000]:

$$E^{\mathbb{Q}} \left[e^{-\int_t^u (r_s + \lambda_s^{\mathbb{Q}}) ds} \lambda_u^{\mathbb{Q}} w_u \right] du = e^{\alpha_B(t,u) + \beta_B(t,u) \cdot Y(t)} [\hat{\alpha}(t,u) + \hat{\beta}(t,u) \cdot Y(t)], \quad (25)$$

where $\alpha_B(t,u)$, $\beta_B(t,u)$, $\hat{\alpha}_d(t,s)$, and $\hat{\beta}_d(t,s)$ are given by explicit formulas. Only the one-dimensional integral in (15) is computed numerically. Moreover, an equally tractable pricing model is obtained if the fractional loss $L_\tau^{\mathbb{Q}}$ is incurred at T , the original maturity of the bond (the convention used in most structural models). In this case, $w_\tau = (1 - L_\tau^{\mathbb{Q}})D(\tau, T)$ is the discounted recovery (from T to τ) and $L_\tau^{\mathbb{Q}}$ must be chosen judiciously to facilitate computation of (25).

Whether one is modeling r alone (for pricing default-free *FIS*) or r and $\lambda^{\mathbb{Q}}$ (for pricing defaultable securities), the choice of affine models will determine whether these variables remain strictly positive over the entire state space. Strictly speaking, negative values for either of these variables are not economically meaningful. However, the only family of affine diffusions that guarantees strictly positive $(r, \lambda^{\mathbb{Q}})$ are those in the family $A_N(N)$. With $M = N$, negative correlations among the Y 's are not admissible (*DS*), and $\lambda_2 = 0$ in (23) thereby restricting the state dependence of $\Lambda(t)$. Thus, the common practice of studying models in $A_M(N)$, with $M < N$, gives greater flexibility at specifying factor correlations and market prices of risk, at the expense of (typically small) positive probabilities of the realized r or $\lambda^{\mathbb{Q}}$ being negative.

Though Y is a latent state vector in most econometric studies of affine *DTSMs*, it is effectively observable if N of the $K \geq N$ bond yields used in estimation are priced perfectly by the model. The implied state Y_t^β is obtained by inverting the model for these N . In empirical implementations (see, e.g., Chen and Scott [1993], Pearson and Sun [1994], Duffie and Singleton [1997], and Honore [1998]), the remaining $K - N$ bonds are often assumed to be priced only up to an additive ‘‘pricing error.’’ In some cases, an alternative to introducing additive errors is to introduce a bond-specific yield factor ψ_t and to discount future cash flows for the bond in question using the adjusted short rate $r(t) + \psi(t)$ (see Duffie, Pedersen, and Singleton [2002] for an example in the context of pricing sovereign bonds).

If all K bonds are assumed to be priced with errors, then filtering methods must be used to obtain fitted states Y^β . Outside of the Gaussian case, the optimal filters are nonlinear so the

Ingersoll, and Ross [1985], Brown and Dybvig [1986], Hull and White [1987], Longstaff and Schwartz [1992], Hull and White [1993], Chen and Scott [1993], Brown and Schaefer [1994], Pearson and Sun [1994], Chen [1996], Balduzzi, Das, Foresi, and Sundaram [1996], and Backus, Foresi, Mozumdar, and Wu [2001].

Kalman filters typically used are only approximations (Duan and Simonato [1999], Duffie and Stanton [2001]). Bobadilla [1999] finds that estimates of β can be sensitive to the parameterization of the pricing errors.

These extraction issues do not arise if the state variables are observed economic time series (e.g., macro-economic and yield curve variables). Critical in choosing yield curve variables as elements of Y is that the model maintain internal consistency – it must correctly “price” the state variables when they are known functions of the prices of traded securities. Duffie and Kan [1996] present a generic example of how this can be done in affine models. Imposing internal consistency in non-affine settings can be challenging, so internal consistency is often ignored (e.g., Boudoukh, Richardson, Stanton, and Whitelaw [1998]). Ang and Piazzesi [2000], Buraschi and Jiltsov [2001], Wu [2000], and Wu [2002] incorporate macro factors (for which these consistency issues do not arise) directly into affine term structure models.

The affine structure of Y also greatly facilitates estimation of *DTSMs*. The conditional characteristic function (*CCF*) of any affine diffusion Y (given Y_{t-1}) is known in closed form (Duffie, Pan, and Singleton [2000]). Therefore, if Y follows an affine diffusion under \mathbb{P} , then all of its conditional moments are also known. In particular, if a vector of zero-coupon bond yields is used in estimation, then method-of-moments is a feasible estimation strategy, since these yields are affine functions of Y (Liu [1997]). A special case is quasi-maximum likelihood estimation (Fisher and Gilles [1996]). Estimators for affine models of zero yields that are asymptotically efficient can be constructed directly from the *CCF* (Singleton [2001], Carrasco, Chernov, Florens, and Ghysels [2002]).

In the case of coupon bonds, *ML* estimation is feasible for those special cases– Gaussian (Jegadeesh and Pennacchi [1996]) and square-root (Chen and Scott [1993], Pearson and Sun [1994])– where the conditional density of Y_t given Y_{t-1} is known. More generally, Duffie, Pedersen, and Singleton [2002] propose an approximate *ML* estimator for affine *DTSMs*.

Estimation may not be as tractable when Λ is modeled differently than as in (23), even if the \mathbb{Q} -drift of Y remains affine. This is true, for example, for Duarte’s formulation which has the drift of Y being nonlinear under \mathbb{P} . In this case, as well as situations where the estimators in the preceding paragraph are not applicable, efficient estimates are obtained using the Monte Carlo maximum likelihood estimator of Pedersen [1995] (see also Brandt and Santa-Clara [2002]), the approximate *ML* estimator of Ait-Sahalia [2002], or the efficient simulated method-of-moments estimator proposed by Gallant and Tauchen [1996].

3.2 Quadratic-Gaussian Models

The quadratic-Gaussian family of *DTSMs* includes the models studied by Longstaff [1989], Beaglehole and Tenney [1991], Constantinides [1992], Leippold and Wu [2002], Ahn, Dittmar, and Gallant [2002], and Lu [2000] as special cases. The ingredients of *QG* models are:

$I_{\mathbb{Q}}(QG)$: Under \mathbb{Q} , the drift and volatility functions of the risk factors satisfy

$$\mu_Y^{\mathbb{Q}}(t) = \nu_0 + \nu_Y Y(t), \tag{26}$$

$$\sigma_Y = \Sigma, \text{ a constant matrix,} \tag{27}$$

where ν_0 is an $N \times 1$ vector and ν_Y is an $N \times N$ matrix of constants.

$I_{\mathbb{P}}(QG)$: Given $\sigma_Y(t)$ satisfying (27), the requirement (26) determines $\mu_Y^{\mathbb{P}}(t)$, once $\Lambda(t)$ is specified, and vice-versa.

$I_r(QG)$: The short rate is a quadratic function of Y :

$$r(t) = a + Y(t)'b + Y(t)'cY(t), \quad (28)$$

where a is a scalar, b is an $N \times 1$ vector, and c is an $N \times N$ matrix of constants.

Ahn, Dittmar, and Gallant [2002] and Leippold and Wu [2002] show that under these conditions the solution to the *PDE* (4) is an exponential quadratic function of Y :

$$D(t, T) = e^{\gamma_0(T-t) + \gamma_Y(T-t)'Y_t + Y_t'\gamma_Q(T-t)Y_t}, \quad (29)$$

where γ_0 , γ_Y , and γ_Q satisfy known ODEs.

The drift condition (26), along with the assumption that the diffusion coefficient $\sigma_Y(t)$ is the constant matrix Σ , imply that Y follows a Gaussian process under \mathbb{Q} . These assumptions do not restrict the drift of Y under \mathbb{P} , however. Subject to preserving no arbitrage, we are free to choose essentially any functional form for $\mu_Y^{\mathbb{P}}(t)$, so long as $\Lambda(t)$ is chosen so that the weighted difference $(\mu_Y^{\mathbb{P}}(t) - \Sigma\Lambda(t))$ is affine in Y . The special case examined in Ahn, Dittmar, and Gallant [2002] has $\Lambda(t) = \lambda_0 + \lambda_Y Y(t)$, which implies that Y is Gaussian under both \mathbb{P} and \mathbb{Q} . Interestingly, this is the same functional form as is obtained for the Gaussian affine model with $\Lambda(t)$ as in (23) since, in this case, $S(t)$ is a constant matrix. The effects of $\Lambda(t)$ on prices in the *QG* and Gaussian affine models are not equivalent, of course, because the mappings between Y and zero-coupon bond prices are different.

The SAINTS model proposed by Constantinides [1992] is shown by Ahn, Dittmar, and Gallant [2002] to be the special case of the *QG* model in which Σ is diagonal and $\mu_Y^{\mathbb{P}}(t) = -\mathcal{K}Y(t)$ with \mathcal{K} diagonal (the N risk factors are mutually independent); and the coefficients λ_0 and λ_Y determining $\Lambda(t)$ are constrained to be specific functions of the parameters governing the Y process. These constraints potentially render the SAINTS model much less flexible than the general *QG* model in representing the distributions of bond yields.

Estimation of *QG* models is complicated by the fact that there is not a one-to-one mapping from observed yields to the state vector Y , because of the quadratic dependence of r on Y . For example, in a one-factor *QG* model, the yield on a s -period, zero-coupon bond, y_s , is given by $y_s(t) = a_s + b_s Y(t) + c_s Y(t)^2$. Given $y_s(t)$ and under suitable parameter values (so that $b_s^2 - 4c_s(a_s - y_s(t)) > 0$), there are two roots to the above quadratic equation, corresponding to two possible values of the implied state variable:

$$Y(t) = \frac{-b_s \pm \sqrt{b_s^2 - 4c_s(a_s - y_s(t))}}{2c_s}. \quad (30)$$

Given this indeterminacy, filtering methods are called upon to estimate the model-implied Y . Ahn, Dittmar, and Gallant circumvent this problem by using simulated method of moments (which effectively gives observable Y through simulation). Then they use the reprojection methods proposed by Gallant and Tauchen [1998] to estimate $E[Y(t)|y(t-1), \dots, y(t-L)]$, with y being the vector of bond yields used in estimation. Lu [2000] uses the filtering density

$f(Y(t)|y(t), y(t-1), \dots, y(1))$ to compute the conditional expectation of $Y(t)$. Though the likelihood function for QG models can be written down in closed-form, we are not aware of any studies that directly implement the ML estimator.

QG models are also easily adapted to the problem of pricing defaultable securities by having both r and $\lambda^{\mathbb{Q}}$ be quadratic functions of Y . In this setting, QG models offer the flexibility, absent from affine models, of having strictly positive $(r, \lambda^{\mathbb{Q}})$ and negatively correlated state variables (see, e.g., Duffie and Liu [2001]).

3.3 Non-affine, Stochastic Volatility Models

Andersen and Lund [1997,1998] studied various special cases of the following non-affine three-factor model⁹

$$\begin{aligned} d \log v(t) &= \mu(\bar{v} - \log v(t))dt + \eta dW_v(t), \\ d\theta(t) &= \nu(\bar{\theta} - \theta(t))dt + \sqrt{\alpha_\theta + \beta_\theta \theta(t)} dW_\theta(t), \\ dr(t) &= \kappa(\theta(t) - r(t))dt + r(t)^\gamma v(t) dW_r(t), \end{aligned} \tag{31}$$

where (W_v, W_θ, W_r) are independent Brownian motions. In this model the volatility factor $v(t)$ follows a lognormal process and the instantaneous stochastic volatility of r , $r(t)^\gamma v(t)$, is affected both by v and r . These models do not have known, closed-form solutions for bond prices. Largely for this reason this formulation of stochastic volatility has been studied primarily in the context of econometric modeling of the short-rate and not as a $DTSM$.

Another non-affine, one-factor $DTSM$ has (Cox, Ingersoll, and Ross [1980], Ahn and Gao [1999])

$$dr(t) = \kappa(\theta - r(t))r(t) dt + r(t)^{1.5} dW_r(t). \tag{32}$$

This “three-halves” model gives a stationary r so long as κ and σ are greater than 0, and the conditional density $f(r_{t+1}|r_t)$ is known in closed form (Ahn and Gao [1999]). For this model, the market price of risk of r (as specified by Ahn and Gao) is $\Lambda(t) = \lambda_1/\sqrt{r(t)} + \lambda_2\sqrt{r(t)}$. Therefore, unlike in CIR -style affine models, $\Lambda(t)$ may change signs over time if λ_1 and λ_2 have opposite signs. Moreover, the \mathbb{P} and \mathbb{Q} drifts of r may be quite different. Thus, this model may achieve added flexibility for fitting yields relative to many affine $DTSMs$, though correlated, multi-factor extensions have yet to be worked out.

⁹Another non-affine model that has received considerable attention in the financial industry (but relatively little attention in academic research) has $\log r(t)$ following a Gaussian process. Perhaps the most well known version is the Black, Derman, and Toy [1990] model, along with its continuous-time limit in Black and Karasinski [1991]. A two-factor (multi-nomial) extension was studied by Peterson, Stapleton, and Subrahmanyam [1998]. An important limitation of this class of models, first noted by Hogan and Weintraub [1993], is that, for any positive time interval, expected rollovers of r are infinite, because the continuous compounding involves an exponential of an exponential. To circumvent this problem, Sandmann and Sondermann [1997] propose that one construct a “money-market” account as the numeraire for risk-neutral pricing using rates with a finite accrual period like a three-month rate.

3.4 Parametric Structural Models of Defaultable Bond Prices

Structural models of default combine an arbitrage-free specification of the default-free term structure with an explicit definition of default in terms of balance sheet information. The former is typically taken to be a standard affine *DTSM*. The new, non-trivial practical consideration that arises in implementing structural models is the computation of the forward, first-passage probability \mathbb{Q}^T .

A representative structural model has firm value A following a log-normal diffusion with constant variance and non-zero (constant) correlation between A and the instantaneous riskless rate r :

$$\frac{dA_t}{A_t} = (r - \gamma) dt + \sigma_A dW_{At}, \quad (33)$$

$$dr = \kappa(\mu - r) dt + \sigma_r dW_{rt}, \quad (34)$$

where $\text{corr}(dW_A, dW_r) = \rho$ and γ is the payout rate. Kim, Ramaswamy, and Sundaresan [1993] and Cathcart and El-Jahel [1998] adopt the same model for A , but assume that r follows a one-factor square-root ($A_1(1)$) process. Related structural models are studied by Nielsen, SaaRequejo, and Santa-Clara [1993] and Briys and de Varenne [1997].

The basic “Merton model” has: (i) the firm capitalized with common stock and one bond that matures at date T , (ii) a constant net payout rate γ and a constant interest rate r , and (iii) default occur when $A_T < F$, where F is constant. (Firms default only at maturity of the bond.) In actual applications of this model, a coupon bond is typically assumed to be a portfolio of zero-coupon bonds, each of which is priced using the Merton model. Geske [1977] extended Merton’s model to the case of multiple bonds maturing at different dates.

Building upon Black and Cox [1976], Longstaff and Schwartz [1995] allowed the issuer to default at any time prior to maturity of the bonds (not just at maturity) and replaced the assumption of constant r with the one-factor Vasicek [1977] model (34). Though the Longstaff-Schwartz model is in many respects more general than the Merton model, the latter is not nested in the former.¹⁰

In the Leland and Toft [1996] model, a firm continuously issues new debt with coupons that are paid from the firm’s payout γA . The default boundary is endogenous, because equityholders can decide whether or not to issue new equity to service the debt in the event that the payout is not large enough to cover the dividends. Their model gives a closed-form expression for $B(t, T)$ under the assumption of constant r . Anderson and Sundaresan [1996] and Mella-Barral and Perraudin [1997] solve simplified bargaining games to obtain close-form expressions for their default boundaries.

A typical feature of structural pricing models is that the value of the firm diffuses continuously over time. This has the counter-factual implication that yield spreads on short-maturity, defaultable bonds will be near zero, since it is known with virtual certainty whether or not an issuer will default over the next short interval of time. As shown by Duffie and Lando [2001], more plausible levels of short-term spreads are obtained in structural models

¹⁰ The Merton model gives a closed form solution for defaultable zero-coupon bond prices. Longstaff and Schwartz provided an approximate numerical solution for $B(t, T)$ in their setting. Subsequently, Collin-Dufresne and Goldstein [2001b] provided an efficient numerical method for computing the $B(t, T)$.

by making the assumption that bondholders measure firm's assets with error. Once measurement errors are introduced, this basic structural model becomes mathematically equivalent to an intensity-based, reduced-form model.

Equally importantly, many structural models also imply that credit spreads tend to asymptote to zero with increasing maturity. This is a consequence of the fact that (risk-neutrally) A drifts away from the default boundary at the rate r so the forward probability of default, $H^T(A_t/F, r_t, T - t)$ converges to zero as $T \rightarrow \infty$. To address this counter-factual implication, Collin-Dufresne and Goldstein [2001b] and Tauren [1999] attribute a target debt/equity ratio to issuers, and Ericsson and Reneby [2001] assume a positive growth rate of total nominal debt.

3.5 Bond Pricing with Jumps

There is growing evidence that jumps are an important ingredient in modeling the distribution of interest rates. For instance, Das [2002], Johannes [2000], and Zhou [2001] find that jump-diffusion models fit the conditional distribution of short-term interest rates better than the nested diffusion models they examine. Reduced-form models for the pricing of bonds (defaultable or default-free) are easily extended to allow Y to follow a jump-diffusion

$$dY(t) = \mu_Y(Y) dt + \sigma_Y(Y) dW(t) + \Delta Y dZ(t), \quad (35)$$

where Z is a Poisson counter, with state-dependent intensity $\{\lambda^{\mathbb{P}}(Y(t)) : t \geq 0\}$ that is a positive, affine function of Y , $\lambda^{\mathbb{P}}(Y) = l_0 + l_Y Y$; and ΔY is the jump amplitude with distribution $\nu^{\mathbb{P}}$ on \mathbb{R}^N . If the jump risk is priced, then a compensated jump term also appears in the pricing kernel with a possibly state-dependent coefficient $\Gamma(\Delta Y, Y)$ representing the market price of jump risk:

$$\frac{dM_t}{M_t} = -r(Y) dt - \Lambda(Y)' dW(t) - [\Gamma(\Delta Y, Y) dZ(t) - \gamma(Y) \lambda^{\mathbb{P}}(Y) dt], \quad (36)$$

where $\gamma(Y) = \int \Gamma(x, Y) d\nu^{\mathbb{P}}(x)$ is the conditional \mathbb{P} -mean of Γ . From (36), the risk-neutral distribution of the jump size and the risk-neutral jump arrival rate are given by

$$d\nu^{\mathbb{Q}}(x) = \frac{1 - \Gamma(x, Y)}{1 - \gamma(Y)} d\nu^{\mathbb{P}}(x), \quad \lambda^{\mathbb{Q}}(Y) = (1 - \gamma(Y)) \lambda^{\mathbb{P}}(Y). \quad (37)$$

Although in general $\Gamma(x, Y)$ may depend on both Y and the jump amplitude x , and therefore $\gamma(Y)$ may be state-dependent, in most implementations Γ is assumed to be a constant.

These expressions simplify further if we can write $M_t = M(Y_t, t)$. This is possible, for example, in equilibrium pricing models where M represents marginal utility that depends only on the current state. In this case, Ito's lemma implies that

$$r_t = -\frac{1}{M_t} \left[\frac{\partial}{\partial t} + \mathcal{A}^{\mathbb{P}} \right] M_t, \quad \Lambda_t = -\sigma_t' \frac{\partial}{\partial Y_t} \log M_t, \quad \Gamma_t(x) = 1 - \frac{M(Y_t + x, t)}{M(Y_t, t)}. \quad (38)$$

Note in particular the link between the market price of jump risk Γ and M : the sign of M (read "marginal utility") depends on whether a jump in Y raises or lowers marginal

utility relative to the pre-jump value. Moreover, the risk-neutral jump arrival rate and the risk-neutral distribution of the jump size are given by

$$\lambda(Y_t, t) = \frac{\int M(Y_t + x, t) d\nu_t^{\mathbb{P}}(x)}{M(Y_t, t)} \lambda^{\mathbb{P}}(Y_t, t), \quad d\nu_t^{\mathbb{Q}}(x) = \frac{M(Y_t + x, t)}{\int M(Y_t + x', t) d\nu_t^{\mathbb{P}}(x')} d\nu_t^{\mathbb{P}}(x).$$

If Y is an affine-jump diffusion under the risk-neutral measure, with the risk-neutral drift and volatility specifications being affine as in (17) - (19)), and the “jump transform” $\varphi(c) = \int_{\mathbb{R}^N} \exp(c \cdot x) d\nu^{\mathbb{Q}}(x)$, for c an N -dimensional complex vector, is known in closed form, then the *PDE* defining the zero prices $D(t, T)$ admits a closed-form solution (up to *ODEs*) as an exponential-affine function of Y , just as in the case of affine diffusions (Duffie, Pan, and Singleton [2000]). Care must be taken in specifying $\varphi(c)$ to make sure that Y remains an admissible process. For instance, for those risk factors that follow square-root diffusions in the absence of jumps, it appears that the added jump must be positive to assure that this factor never become negative. Das and Foresi [1996] and Chacko and Das [2001] present illustrative examples of affine bond and bond-option pricing models with jumps.

State variables with jumps have received relatively less attention in the empirical literature on *DTSMs*. One of the earliest affine models with jumps is that of Ahn and Thompson [1988] who extend the equilibrium framework of Cox, Ingersoll, and Ross [1985] to the case of Y following a square-root process with jumps. Brito and Flores [2001] develop an affine jump-diffusion model, and Piazzesi [2001] develops a mixed affine-QG model, in which the jumps are linked to the resetting of target interest rates by the Federal Reserve (see also Das [2002]).

Within a reduced-form model of defaultable bonds, Collin-Dufresne and Solnik [2000] have the mean loss rate $s = \lambda^{\mathbb{Q}} L^{\mathbb{Q}}$ following a Gaussian jump-diffusion model with constant arrival intensity for jumps in $\lambda^{\mathbb{Q}} L^{\mathbb{Q}}$. This formulation allows credit spreads to be negative, even without the jump component.

In a structural pricing model, Zhou [2000] added the possibility of a jump in assets A with *i.i.d.* amplitudes at independent Poisson arrival times, thereby allowing for A to pass through the default threshold (F in our basic formulation) either through continuous fluctuations of the Brownian motion or by jumps.

3.6 *DTSMs* with Regime Shifts

A potential limitation of diffusion models of the risk factors Y is that the resulting *DTSMs* may not match the higher-order moments of bond yields. While adding jumps would help in this regard, jump-diffusions may not generate *persistent* periods of “turbulent” and “quiet” bond markets. To fit such patterns in historical yields, Hamilton [1988], Gray [1996], and Ang and Bekaert [2001], among others, have had some success using switching-regime models.

Motivated by these descriptive studies, this section introduces the possibility of multiple economic regimes into the affine family of *DTSMs*. We begin by presenting a continuous-time treatment of regime-switching in affine *DTSMs* that allows for changes over time in the distribution of the state vector and in the market prices of risk. This discussion extends the complementary treatment in Landen [2000] by parameterizing the pricing kernel under

the measure \mathbb{P} (in addition to under \mathbb{Q}) and allowing for state-dependent probabilities of transitioning between regimes.

The evolution of “regimes” is described by an $(S + 1)$ -state continuous-time *conditionally Markov chain* $s_t : \Omega \rightarrow \{0, 1, \dots, S\}$ with a state-dependent $(S + 1) \times (S + 1)$ rate or generator matrix $R_t^{\mathbb{P}} = [R_{ij,t}^{\mathbb{P}}]$ in which all rows sum to zero. (See Bielecki and Rutkowski [2001] for formalities.) Intuitively, $R_{ij,t}^{\mathbb{P}} dt$, $i \neq j$ represents the probability of moving from regime i to regime j over the next short interval of time. The subsequent discussion is simplified notationally by introducing $(S + 1)$ regime indicator functions $z_t^j = 1_{\{s_t=j\}}$, $j = 0, 1, \dots, S$, with the property that $E[dz_t^j | s_t, Y_t] = R_t^{\mathbb{P}j} dt$, where $R_t^{\mathbb{P}j} \equiv R^{\mathbb{P}j}(s_t; Y_t, t) = \sum_{i=0}^S z_t^i R_{ij,t}^{\mathbb{P}}$.

To introduce regime-switching into a bond pricing model, we assume that the pricing kernel can be written as $M_t \equiv M(s_t; Y_t, t) = \sum_{j=0}^S z_t^j M(s_t = j; Y_t, t)$. (As noted in the case of jumps, having $M(s_t; \cdot)$ depend only on Y implicitly constrains the state-dependence of M .) Then, using Ito’s lemma,

$$\frac{dM_t}{M_t} = -r_t dt - \Lambda'_t dW_t - \sum_{j=0}^S \Gamma_t^j \left(dz_t^j - R_t^{\mathbb{P}j} dt \right). \quad (39)$$

$\Lambda(s_t; Y_t, t)$ is the market price of diffusion risk and $\Gamma^j(s_t; Y_t, t)$ is the market price of a shift from the current regime s_t to regime j an instant later. Under this formulation of M_t , $\Gamma^j(s_t = i; Y_t, t) = [1 - M(s_t = j; Y_t, t) / M(s_t = i; Y_t, t)]$. Therefore, $\Gamma^i(s_t = i; Y_t, t) = 0$ and

$$(1 - \Gamma^i(s_t = j; Y_t, t))(1 - \Gamma^j(s_t = i; Y_t, t)) = 1, \quad 0 \leq i, j \leq S. \quad (40)$$

Thus, there are only $\frac{1}{2}S(S + 1)$ free market prices of risk for regime shifts. In particular, for a two-regime model ($S = 1$) there is only one free market price of regime switching risk, representing the ratio of the pricing kernels for the two regimes.

The risk-neutral distribution of the short-rate is governed by the relations $r_t^i \equiv r(s_t = i; Y_t, t) = \delta_0^i + Y'_t \delta_Y^i$, and the assumption that (risk-neutrally) Y follows an affine diffusion with regime-dependent drifts and volatilities

$$\begin{aligned} \mu^{\mathbb{Q}}(s_t; Y_t, t) &\equiv \sum_{j=0}^S z_t^j \kappa^{\mathbb{Q}j} (\theta^{\mathbb{Q}j} - Y_t), \\ \sigma(s_t; Y_t, t) &= \sum_{j=0}^S z_t^j \text{diag}(\alpha_k^j + Y'_t \beta_k^j)_{k=1,2,\dots,N}. \end{aligned} \quad (41)$$

where δ_0^i and α_k^i are constants, $\kappa^{\mathbb{Q}i}$ is a constant $N \times N$ matrix, and δ_Y^i , $\theta^{\mathbb{Q}i}$, and β_k^i are constant $N \times 1$ vectors. When a regime shifts, the conditional moments of Y change, but its sample path remains continuous.

Letting $D(t, T) \equiv D(s_t, Y_t; t, T)$, we can write $D(t, T) \equiv \sum_{j=0}^S z_t^j D^j(t, T)$, where $D^j(t, T) \equiv D(s_t = j, Y_t; t, T)$. No arbitrage, which requires that $\mu^D(t, T) = r_t D(t, T)$ for all $0 \leq s_t \leq S$ and all admissible Y_t , implies that the $D^j(t, T)$ satisfy the $(S + 1)$ PDEs

$$\left[\frac{\partial}{\partial t} + \mathcal{A}^i \right] D^i(t, T) + \sum_{j=0}^S R_{ij,t}^{\mathbb{Q}} D^j(t, T) - r^i(Y_t, t) D^i(t, T) = 0, \quad 0 \leq i \leq S, \quad (42)$$

where \mathcal{A}^i is the counterpart to (5) for regime $s_t = i$, $R_{ij,t}^{\mathbb{Q}} = (1 - \Gamma^j(s_t = i; Y_t, t))R_{ij,t}^{\mathbb{P}}$ if $j \neq i$, and $R_{ii,t}^{\mathbb{Q}} = -\sum_{j \neq i} R_{ij,t}^{\mathbb{Q}}$. ($R_t^{\mathbb{Q}}$ is the rate matrix of the conditionally Markov chain under the risk-neutral measure \mathbb{Q} .) In general, the matrix $R_t^{\mathbb{Q}}$ is not diagonal. Therefore, these $S+1$ PDEs are coupled, and the $(D^i(t, T) : 0 \leq i \leq S)$ must be solved for jointly. The boundary condition is $D(T, T) = 1$ for all s_T , which is equivalent to $(S + 1)$ boundary conditions: $(D^i(T, T) = 1 : 0 \leq i \leq S)$.

An affine regime-switching model with a closed-form solution for zero-coupon bond prices is obtained by specializing further to the case where: $R_t^{\mathbb{Q}}$ is a constant matrix, and $\kappa^{\mathbb{Q}i}$, δ^i , and β_k^i are independent of i . Under these assumptions,

$$D^i(t, T) = e^{\gamma_{0i}(T-t) + \gamma_Y(T-t)'Y_t}, \quad 0 \leq i \leq S, \quad (43)$$

where the $\gamma_{0i}(\cdot)$ and $\gamma_Y(\cdot)$ are explicitly known up to a set of ODEs. Note that in this specialized environment regime-dependence under \mathbb{Q} enters only through the ‘‘intercept’’ term $\gamma_{0i}(T - t)$; the derivative of zero-coupon bond yields with respect to Y does not depend on the regime. Though admittedly strong, these assumptions do allow for Y to follow a general affine diffusion and for the \mathbb{P} -rate matrix $R^{\mathbb{P}}$ to be state-dependent.

In both respects, this formulation extends the one-factor, continuous-time formulation of Naik and Lee [1997] (as well as Proposition 3.2 of Landen [2000]). Even with regime-switching, it may be empirically more plausible to allow for multiple, correlated factors risk factors. Moreover, Naik and Lee assume constant market prices of regime-shift risk ($\Gamma^j(t)$ are constants), and obtain regime-independence of the risk-neutral feedback matrix $\kappa^{\mathbb{Q}j}$ under the stronger assumption that the actual matrix $\kappa^{\mathbb{P}j}$ is independent of j .

In sorting out the added econometric flexibility of these model relative to single-regime, affine DTSMs, it is instructive to examine the implied excess returns on a $(T - t)$ -period zero-coupon bond. Based on the pricing kernel (39), for current regime $s_t = i$, we have

$$\mu_{Dt}^i - r_t^i = \sigma_{Dt}^{i'} \Lambda_t^i - \sum_{j=0}^S \Gamma^j(s_t = i) \left[1 - \frac{D^j(t, T)}{D^i(t, T)} \right] R_t^{\mathbb{P}j}, \quad (44)$$

where σ_{Dt}^i is the diffusion vector in regime i for $D^i(t, T)$. If $\Gamma^j(s_t = i) = 0$, for all $j = 0, 1, \dots, S$, then excess returns may still be time varying for two reasons: (i) state-dependence of Λ_t and/or $\sigma_Y(t)$ (as in single-regime models), and (ii) the possibility that either of these constructs might shift across regimes. It is the latter added source of flexibility that the previous literature on DTSMs with regime shifts has relied on to improve goodness-of-fit over single-regime DTSMs.

By allowing for priced regime-shift risk ($\Gamma^j(s_t) \neq 0$), we see from (44) that our formulation introduces an additional source of variation in excess returns. This is true even if $R^{\mathbb{P}}$ is a constant (non-state dependent) matrix. Of course, allowing $R^{\mathbb{P}}$ to be state-dependent, while maintaining the assumption of constant $R^{\mathbb{Q}}$ for computational tractability, would add additional flexibility to this model.

3.7 Discrete-Time DTSMs

There is a limited amount of (mostly) academic research that takes a relatively non-parametric approach to specifying M in discrete time. A sequence of cash flows is then priced under

\mathbb{P} using (3). Backus and Zin [1994] parameterize $-\log(M_t/M_{t-1})$ as an infinite order, moving average process with *i.i.d.* normal innovations. This formulation accommodates richer dynamics than a Gaussian diffusion model and is easily extended to multiple factors, but it abstracts from time-varying volatility. More recently, Brandt and Yaron [2001] parameterize $-\log(M_t/M_{t-1})$ as a Hermite polynomial function of (Y_t, Y_{t-1}) , where Y_t is an observable state vector. Their model extends the Backus-Zin specification of M by allowing for non-normality and time-varying conditional moments, but it is more restrictive in requiring that the pricing kernel depend only on (Y_t, Y_{t-1}) and that Y be observable. Similarly, Lu and Wu [2000] model the state-price density using a semi-non-parametric density based on Hermite polynomial expansions.

These semi-parametric approaches, though flexible, often present their own challenges. Specifically, it may be difficult to verify that the parameters of the pricing kernel are identified from bond yield series. Also, if the state variables are taken to be functions of observable bond yields, then internal consistency requires that the same functions of the model-implied yields must recover the state vector. This consistency is not always easily imposed.

Sun [1992] and Backus, Foresi, and Telmer [1998] develop discrete-time counterparts to Gaussian and *CIR*-style affine *DTSMs* by working with an Euler approximation to the state vector. In the latter case, these models are approximate in that the conditional distribution of Y must be truncated to assure non-negativity of the conditional variance of Y . Gourieroux, Monfort, and Polimenis [2002] develop exact, discrete-time versions of affine *DTSMs* by modeling the conditional distributions of Y and M directly, potentially as functions of a long history of Y . To our knowledge, these models have not been extended to allow for market prices of risk as rich as (23) or (24).

Bansal and Zhou [2002] develop discrete-time, regime-switching models that relax the assumptions that $\kappa^{\mathbb{Q}^i}$ and β_t^i are regime-independent, while maintaining the assumption that $R^{\mathbb{P}} = R^{\mathbb{Q}}$, a constant matrix (the $\Gamma^j(t) = 0$). Their model does not admit a closed-form exponential affine solution, so they proceed by linearizing the discrete-time Euler equations and then solving the resulting linear relations for prices.

Note that, following the construction of Hamilton, Bansal and Zhou [2002] specify the conditional mean and variance of $\log \frac{M_{t+1}}{M_t}$ as functions of (s_{t+1}, Y_t) . This is different from the discrete-time formulation of our continuous-time regime-switching model which has the pricing kernel dependent on s_t , not on the future realization of the regime state s_{t+1} . Conceptually, having $\Lambda(t)$ depend on s_t , and not s_{t+1} , seems more natural if it is to be interpreted as the market price of risk, since the latter must be in agents' information set at the time prices are set, date t . Notwithstanding this reservation, the econometric implications of the Bansal-Zhou and our formulations may be very similar, because the continuous-time limit of their formulation is a special case of our continuous-time formulation.

4 Fitting *DTSMs* to Swap and Treasury Yields

We begin our more indepth exploration of the empirical fit of *DTSMs* by reviewing their applications to the treasury and swap yield curves. As a means of organizing our discussion, we focus on the following empirical observations:

LPY: Letting $y^n \equiv -\log D(t, t+n)/n$ denote the yield on an n -period zero-coupon bond, linear projections of $y_{t+1}^{(n-1)} - y_t^n$ on the slope of the yield curve, $(y_t^n - r_t)/(n-1)$, give fitted coefficients ϕ_{nT} that are negative, often increasingly so for longer maturities.

Campbell and Shiller [1991] and Backus, Foresi, Mozumdar, and Wu [2001], among others, document significantly negative ϕ_{nT} , particularly for large n , using U.S. Treasury bond data over the past fifty years (see Table 1). Using U.S. data, the pattern *LPY* is most pronounced (and statistically significant) for sample periods that include the change in monetary operating procedures during 1979-1983. However, notably, ϕ_{nT} is consistently negative across simple periods including prior and subsequent to this monetary “experiment,” though (no doubt due in part to the shorter sample period) the standard error bands are also larger.¹¹

Maturity	12	24	36	48	60	120
Campbell-Shiller (1991) 1952-1978	-0.672 (.598)	-1.031 (.986)	-1.210 (1.187)	-1.272 (1.326)	-1.483 (1.442)	-2.263 (1.869)
Campbell-Shiller (1991) 1952-1987	-1.381 (.683)	-1.815 (1.151)	-2.239 (1.444)	-2.665 (1.634)	-3.099 (1.749)	-5.024 (2.316)
Backus, et. al. (2001) 1970-1995	-1.425 (.825)	-1.705 (1.120)	-1.190 (1.295)	-2.147 (1.418)	-2.433 (1.519)	-4.173 (1.985)
Backus, et. al. data 1984-1995	0.206 (.527)	-0.001 (1.013)	-0.295 (1.358)	-0.478 (1.610)	-0.566 (1.811)	-0.683 (2.593)

Table 1: The estimated slope coefficients ϕ_{nT} in the regression of $y_{t+1}^{(n-1)} - y_t^n$ on $(y_t^n - r_t)/(n-1)$. The maturities n are given in months and estimated standard errors of the ϕ_{nT} are given in parentheses.

Looking outside the U.S., the pattern *LPY* is not nearly so pronounced, if there at all. Among the studies that document different patterns, primarily for European countries, are Hardouvelis [1994], Kugler [1997], Gerlach and Smets [1997], Evans [2000], and Bekaert and Hodrick [2001].

While these projections have been discussed most widely in the context of tests of the “expectations theory” of the term structure, we follow the recent literature on *DTSMs* and treat these projections as descriptive statistics that an empirically successful *DTSM* should match.¹² Focusing on the ϕ_{nT} naturally raises the issue of whether their large negative values are a consequence of small-sample bias induced by highly persistent interest rate levels and spreads. The Monte Carlo studies in Bekaert, Hodrick, and Marshall [1997] and Backus, Foresi, Mozumdar, and Wu [2001] suggest that the bias is toward *larger* values of ϕ_{nT} , thereby reinforcing the conclusion that $\phi_n < 0$ for large n .

¹¹ Preliminary analysis of weekly swap data over the period 1988-2000 gives values of ϕ_{nT} somewhat larger than -1 . for U.S. data.

¹² There is substantial empirical evidence that the assumption of constant risk premiums maintained in the expectations theory is inconsistent with historical data. Some of the central references on this topic are Fama [1984a,1984b], Fama and Bliss [1987], and Campbell and Shiller [1991]. Bekaert and Hodrick [2001] argue that the past use of large-sample critical regions, instead of their small-sample counterparts, may have overstated the evidence against the expectations theory.

CVY Conditional volatilities of changes in yields are time-varying and and persistent. Moreover, the term structures of unconditional volatilities of U.S. swap and treasury yields have recently been hump-shaped.

There is substantial evidence that bond yields exhibit time-varying conditional second moments (e.g., Ait-Sahalia [1996], Brenner, Harjes, and Kroner [1996], Gallant and Tauchen [1998]). Other than in the case of the Gaussian affine and basic log-normal models, *DTSMs* typically build in time-varying volatility, a property that is naturally central to the reliable valuation of many fixed-income derivatives. Thus, challenge *CVY* is not whether yields exhibit time-varying volatility, but rather whether there is *enough* model-implied variation in volatility (both in magnitude and persistence) to match historical experience.

Another important dimension of *CVY* is that the term structure of *unconditional* volatilities of (changes in) bond yields has tended to be hump-shaped over the past ten to fifteen years (see, e.g., Litterman, Scheinkman, and Weiss [1988]). Plotting the volatilities of zero-coupon treasury bond yields against maturity over the period 1983 - 1998 shows a hump that peaks around two to three years in maturity. (Figure 1).¹³ A very similar pattern of volatilities is obtained using U.S. dollar fixed-for-variable rate swap yields for the post 1987 period (for which data is available). Interestingly, it appears that this hump at two years was not a phenomenon observed for the entire post World War II period in the U.S. Figure 1 also displays the term structure of volatilities for the subperiod 1954 - 1978 (the period of the “monetary experiment” from 1979 - 1981 was omitted), during which the volatilities were both smaller and their term structure was less humped.

Single-factor models, as traditionally formulated, are unlikely to be successful in matching these patterns (see Brown and Dybvig [1986], Brown and Schaefer [1994], Backus, Foresi, Mozumdar, and Wu [2001]).¹⁴

4.1 *DTSMs* and Challenge *LPY*

Fisher [1998], Backus, Foresi, Mozumdar, and Wu [2001], Roberds and Whiteman [1999], Duffee [2002], and Dai and Singleton [2002b], among others, have examined whether affine *DTSMs* provide an explanation for the failure of the “expectations theory.”¹⁵ Drawing

¹³Annualized volatility is measured as the standard deviation of changes in the logarithms of bond yields, scaled up by the number of observations per year.

¹⁴Another issue is whether the linear drift specifications in one-factor models are appropriate. Evidence supporting a nonlinear conditional mean for the short rate is discussed in Ait-Sahalia [1996] and Stanton [1997]. In principle, the finding of non-linear drifts for one-factor models could be a consequence of misspecifying the number of factors. However, the non-parametric analyses in Boudoukh, Richardson, Stanton, and Whitelaw [1998] and Balduzzi and Eom [2000] suggest that the drifts in both two- and three-factor models of treasury yields are also nonlinear. In spite of this evidence, it does seem that having multiple factors in linear models is more important, at least for hedging purposes, than introducing non-linearity into models with a smaller number of factors (see, e.g., Balduzzi and Eom [2000]). Perhaps for this reason, or because of the computational demands of pricing in the presence of nonlinear drifts, attention continues to focus primarily on *DTSMs* with linear drifts for the state variables.

¹⁵Dai [2002] and Wachter [2002] examine the puzzle *LPY* in the context of non-affine macro-economic models in which agents preferences exhibit habit formation. McCallum [1994] and Kugler [1997] propose resolutions of the puzzle *LPY* in the context of linear monetary policy rules.

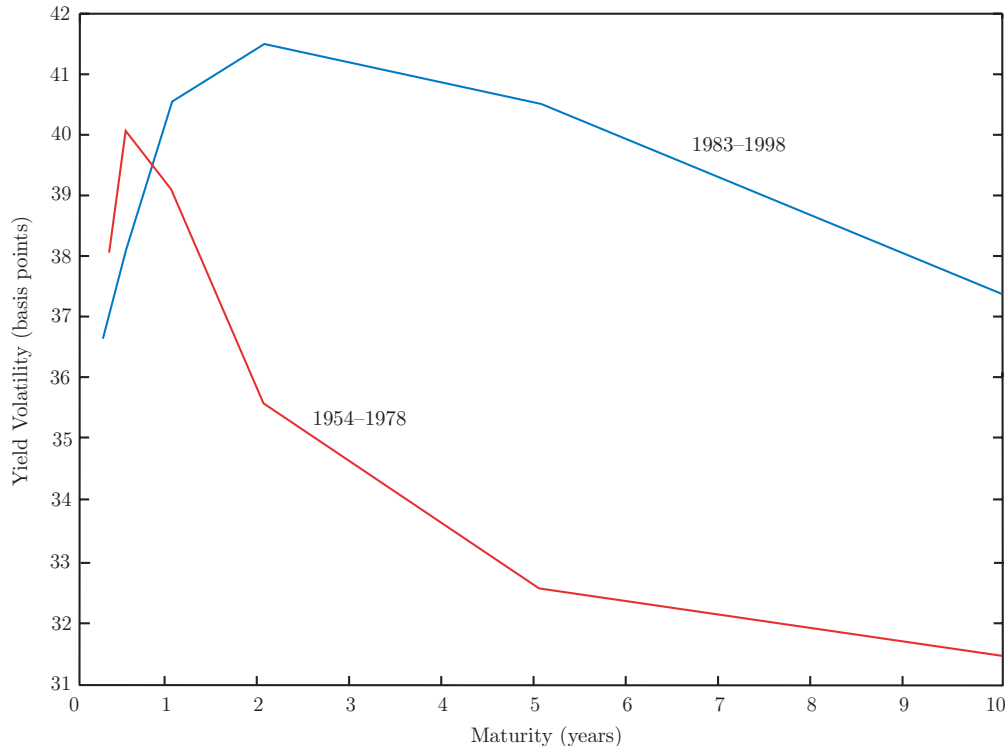


Figure 1: Term structures of volatilities of yields on zero-coupon U.S. treasury bonds based on monthly data from 1954 through 1998.

upon the analysis in Dai and Singleton [2002b], Figure 2 displays displays the *population*¹⁶ ϕ_n implied by canonical three-factor Gaussian ($A_{0C}(3)$) and square-root or *CIR*-style ($A_{3C}(3)$) models, along with the ϕ_{nT} (label *LPY*) estimated with the Fama-Bliss data set. Model $A_{0C}(3)$ was estimated using the extended risk-premium specification (23), while model $A_{3C}(3)$ was fit with the more restrictive specification (22) as required by the no-arbitrage condition.

Model $A_{3C}(3)$ embeds the most flexible specification of factor volatilities (within the affine family), but requires the relatively restrictive risk premium specification (22). From Figure 2 we see that the fitted ϕ_{nT} form (approximately) a horizontal line at unity, implying that the multi-factor *CIR* model fails to reproduce the downward sloping pattern *LPY*. The empirical analysis in Duffee and Stanton [2001] suggests that this failure of *CIR*-style models extends to the special case of (24) with $\lambda_2 = 0$. Thus, it seems that it is not enough to free up the mean of $\Lambda(t)$ in “completely affine” *DTSMs* to match *LPY*; the dynamics of Λ must also be changed as in (23). On the other hand, the Gaussian $A_{0C}(3)$ model, which gives maximum flexibility both in specifying the dynamic properties of the market prices of risk and the factor correlations, is successful at generating ϕ_n that closely resemble *LPY*.

Whether a *DTSM* matches *LPY* speaks to whether it matches the \mathbb{P} -dynamics of yields,

¹⁶Population ϕ_n are computed using a long simulated time-series from the estimated *DTSM*. Monte Carlo evidence reported in Dai and Singleton [2002b] suggests that similar patterns are obtained by computing the average estimates of the ϕ_n from simulated samples of lengths equal to that of the Fama-Bliss data set.

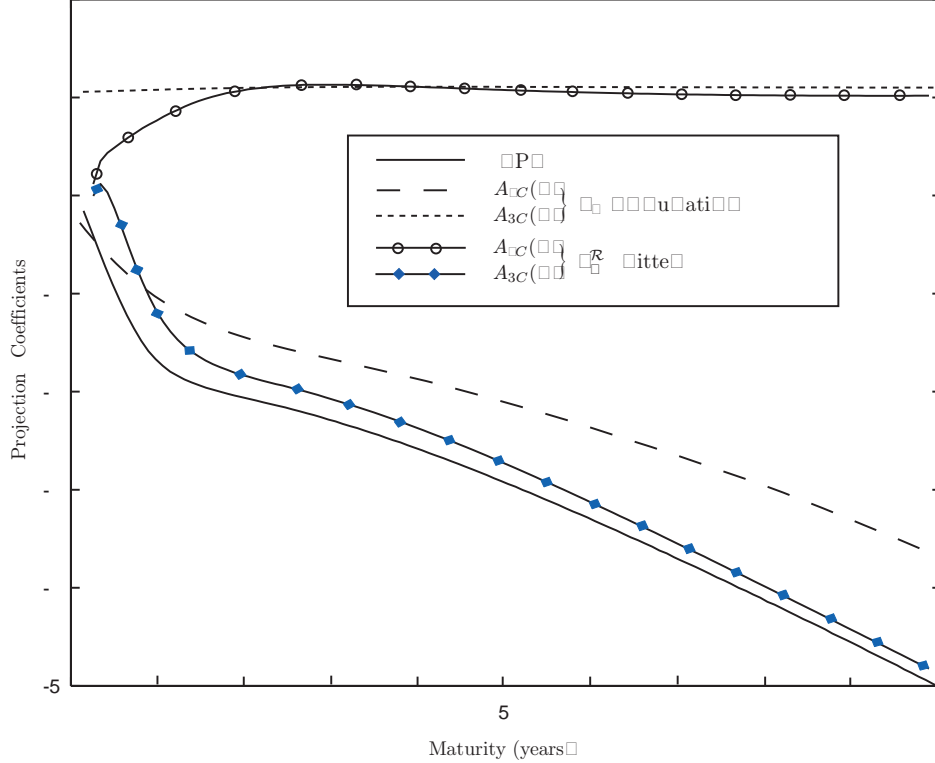


Figure 2: Unadjusted sample and model-implied population projection coefficients ϕ_n . Risk-adjusted sample projection coefficients $\phi^{\mathcal{R}}$. Source: Dai and Singleton [2002].

but it does not directly address whether a *DTSM* matches the \mathbb{Q} -dynamics. The latter is addressed by introducing the yield (c_t) and forward (p_t) “term premiums”

$$c_t^n \equiv y_t^n - \frac{1}{n} \sum_{i=0}^{n-1} E_t[r_{t+i}], \quad p_t^n \equiv f_t^n - E_t[r_{t+n}], \quad (45)$$

and regressing $(y_{t+1}^{n-1} - y_t^n - (c_{t+1}^{n-1} - c_t^{n-1}) + p_t^{n-1}/(n-1))$ onto $(y_t^n - r_t)/(n-1)$. A correctly specified *DTSM* should produce a coefficient $\phi_{nT}^{\mathcal{R}}$ of unity, for all n (Dai and Singleton [2002b]). From Figure 2 it is seen that model $A_{3C}(3)$ gives $\phi_{nT}^{\mathcal{R}}$ that are nearly the same as the historical estimates *LPY*; it is as if model $A_{3C}(3)$ has constant risk premiums. In contrast, the $\phi_{nT}^{\mathcal{R}}$ implied by model $A_{0C}(3)$ are approximately unity, at least for $n \geq 2$ years.

These findings highlight the demands placed on risk premiums in matching the first-moment properties of bond yields. Though the risk premium specification (22) has been widely adopted in econometric studies of affine *DTSMs*, it appears to be grossly inconsistent with *LPY*. Even with Duffee’s extended specification (23), only the Gaussian model appears to match both *LPY* and the requirement that $\phi^{\mathcal{R}} = 1$.

Turning to multi-factor *QG* models, they share essentially the same market prices of risk as the extended Gaussian model, so we would expect them to be equally successful at matching *LPY*. Lu [2000] computes the ϕ_n implied by a three-factor *SAINTS* model and, while they become increasingly negative with larger n , they appear to be too small relative to

the ϕ_{nT} . This may be a consequence of the restrictive nature of $\Lambda(t)$ in the *SAINTS* model. Using LIBOR/swap data over a more recent sample period, Leippold and Wu [2001] find that a general two-factor *QG* model generates patterns forward-rate projections consistent with history.

The structure of risk premiums in regime-switching models also appears to be central to their flexibility in matching *LPY*. All of the empirical studies we are aware of adopt the relatively restrictive risk premium specification (22) within each regime, and assume that regime-shift risk is not priced. Naik and Lee [1997] and Evans [2000] have the market price of risk being proportional to volatility, with the same proportionality constant across regimes. In Naik and Lee, this implies that regime-dependence of the bond risk premium is driven entirely by the regime-dependence of volatility. Evans only allows the long-run means, and not the volatility, of the state variables to vary across regimes. In contrast, Bansal and Zhou [2002] allow the market price of risk to vary across regimes, through both the regime-dependence of volatility and the regime-dependence of the proportionality constant.

Interestingly, Evans' two-factor *CIR*-style model (an $A_2(2)$ model with two regimes) fails to reproduce the historical estimates of ϕ_n from U.K. data (see his Table 6). In contrast, Bansal and Zhou, who study a two-factor *CIR* model with two regimes using U.S. data, generate projection coefficients consistent with the pattern *LPY* in Figure 2. Taken together, these findings suggest that having multiple regimes may overcome the limitations of $A_N(N)$ models in matching *LPY* outlined above, so long as the factor volatilities and risk premiums vary independently of each other across regimes. Yet unaddressed in this literature is whether having the market prices of diffusion risk changing across regimes would be empirically important in the presence of priced regime-shift risk.

4.1.1 *DTSMs* and Challenge *CVY*

A hump-shaped term structure of yield volatilities is inconsistent with the theoretical implications of both one-factor affine and *QG DTSMs*. This is essentially a consequence of mean-reversion of the state.

In multi-factor models, a humped-shaped volatility curve can be induced either by negative correlation among the state variables or hump-shaped loadings on the state variables Y in the mapping between zero coupon yields and Y . Fitted yields from both affine and *QG DTSMs* typically fit the volatility hump (e.g., *DS* and Leippold and Wu [2001]), so long as yields on bonds with maturities that span the humps are used in estimation.

The economic reasons for the different shapes in Figure 1 remain largely unexplored, though the differing patterns pre- and post-1979 are suggestive. In a study of U.S. treasury yields over the period 1991-95, Fleming and Remolona [1999] found that the term structure of “announcement effects” – the responses of treasury yields to macroeconomic announcements – also have a hump-shaped pattern that peaks around two to three years. Moreover, they fit two-factor affine models in which r mean-reverts to a stochastic long-run mean and found that the model-implied announcement impact curves were also humped-shaped. Might it be that investors' attitudes toward macroeconomic surprises following the monetary experiment in the late 1970's changed, much like what happened in option markets following the “crash” of October, 1987? Bekaert, Hodrick, and Marshall [2001] explore the possibility of associated

“peso” effects on yield curve behavior.

Piazzesi [2001]’s analysis lends support to a monetary interpretation of the volatility hump. Her econometric model, which is essentially a four-factor mixed affine-QG model with jumps, not only matches the humped-shape volatility pattern in Figure 1, but also the “snake” shape of volatility between zero and two years (steeply declining volatility from zero to six months and then rising volatility to two years). This study provides a rich structural (monetary) interpretation of the need for a fourth factor to capture the very short end of the LIBOR curve (on the need for four factors, see also Liu, Longstaff, and Mandell [2001]).

The second aspect of *CVY* is the degree of model-implied time-varying volatility relative to what we find in the historical data. To set a historical benchmark for comparing models we estimated GARCH(1,1) (Bollerslev [1986]) models for the five-year yields using historical data.¹⁷ Next we computed *ML* estimates of the canonical $A_{1C}(N)$ models ($N = 2, 3$), based on the risk premium specification (23), with the six-month and two- and ten-year yields (two- and ten-year yields when $N = 2$) assumed to be fit perfectly by the model. Then we refit the same GARCH model using simulated yields from these models. In the case of swaps, we simulated twenty years of weekly data (1040 observations); while in the case of treasury zero-coupon data, we simulated twenty years of monthly data (240 observations). We selected the five-year yield, because it lies between the two- and ten-year yields that are matched perfectly at the implied state variables. The results are displayed in Table 2.

The swap sample, 1987 - 2000, covers a period of relative tranquility in interest rates, compared to the period of the late seventies and early eighties. The five-year yields implied by model $A_{1C}(3)$ exhibit comparable volatility characteristics to the historical data. In contrast, the model $A_{1C}(2)$ substantially understates the degree of volatility persistence in the five-year swap yield. So moving from two to three factors makes a substantial difference in matching the persistence in stochastic volatility during this period, *even though in both cases only Y_1 drives the factor volatilities*.

GARCH(1,1)	$\bar{\sigma}$	α	β
Swap-Sample	0.005 (.001)	0.126 (.038)	0.657 (.062)
Model $A_{1C}(2)$	0.012 (.003)	0.102 (.040)	0.235 (.209)
Model $A_{1C}(3)$	0.008 (.000)	0.126 (.027)	0.793 (.024)
Treasury-Sample	0.016 (.005)	0.165 (.058)	0.749 (.069)
Model $A_{1C}(3)$.000 (.000)	0.146 (0.075)	0.605 (.188)
Model $A_{1R}(3)$.000 (.000)	0.164 (0.070)	NA

Table 2: *ML* estimates of GARCH(1,1) parameters using historical and simulated time series of swap and treasury yields. The GARCH model has $\sigma_t^2 = \bar{\sigma} + \alpha u_t^2 + \beta \sigma_{t-1}^2$, where u_t is the innovation from an AR(1) representation of the level of the five-year yield. Standard errors are given in parentheses.

¹⁷ Clearly one could use a much richer parameterization of conditional volatility than a GARCH(1,1) model— the semi-non-parametric density proposed by Gallant and Tauchen [1996] is one such parameterization. Our goal here was to simply compute a descriptive measure of persistence in volatility that could be used to compare models. Going beyond this basic comparison is an interesting topic for future research.

The Fama-Bliss sample of treasury zero-coupon bond yields covers a twenty-five year period that includes the monetary experiment in the late seventies. Nevertheless, we see that the implied GARCH(1,1) estimates from model $A_{1C}(3)$ again match those in the sample quite closely. To see whether the specification $\Lambda(t)$ affects these results, we re-estimated model $A_{1C}(3)$ with $\Lambda(t)$ given by (22) instead of (23). The results for treasury data— Model $A_{1R}(3)$ in Table 2— are striking, with the volatility in Model $A_{1R}(3)$ showing almost no persistence. The best fitting model was an ARCH(1).¹⁸

The structure of conditional volatility in QG models has been explored extensively in Ahn, Dittmar, and Gallant [2002] and Ahn, Dittmar, Gao, and Gallant [2000] (see also Lu [2000]). They argue that there is a significant difference between three-factor QG and affine models along this dimension, with QG models doing a much better job of matching the conditional variation in the historical U.S. treasury data as captured in their descriptive (semi-non-parametric auxiliary) model. However, their reference affine model was the preferred affine $A_1(3)$ model examined in DS which is based on specification (22) of $\Lambda(t)$. As we have just seen, using instead the specification (23) substantially affects the model-implied yield volatilities. Thus, the magnitude of the gap, or indeed whether there is a gap at all, between the fits of affine and QG models to the dynamic properties of yield volatilities may warrant further exploration.

5 Empirical Studies of Credit Spreads

The empirical literature on corporate bond pricing is relatively in its infancy. Besides overall assessments of goodness-of-fit, this literature has tended to focus on reproducing the negative correlation between treasury rates and corporate spreads observed historically (Duffee [1998]), and on understanding the determinants of spread variation.

5.1 Reduced-Form Models

Within the family of reduced-form models with fractional recovery of market value, Duffee [1999] examined corporate spreads to the U.S. Treasury curve, with $D(t, T)$ described by an $A_2(2)$ affine model for (Y_1, Y_2) and the instantaneous credit spread $s_t = \lambda_t^Q L_t^Q$ given by an affine function of (Y_1, Y_2, Y_3) . For Lehman Brothers data on trader quotes for non-callable corporate bonds, he found a model-implied negative correlation between corporate yield spreads and U.S. treasury rates, consistent with his earlier descriptive analysis. The average error in fitting non-callable corporate bond yields was less than ten basis points. Similarly, Collin-Dufresne and Solnik [2000] have r_t following a Gaussian $A_0(2)$ model and s_t following a Gaussian jump-diffusion model with constant jump intensity. Treating the U.S. treasury curve as the reference curve and using yields on LIBOR contracts as the defaultable securities, they find that the correlation between r_t and s_t is also negative.

Though the preceding studies treated s_t as a purely latent process, observable state variables are easily incorporated into reduced-form models by letting one or more of the

¹⁸ We repeated this calculation for several simulated time series of length twenty years and obtained qualitatively similar results.

Y 's be an observable economic time series. In this manner, it is possible to capture part of the “spirit” of structural models within a reduced-form setting by having $\lambda^{\mathbb{Q}}L^{\mathbb{Q}}$ depend on information about the balance sheet of an issuer. For instance, Bakshi, Madan, and Zhang [2001a] examine a fractional recovery of market value, reduced-form model in which the default-adjusted discount rate R is an affine function of r and $Y_t =$ firm leverage, $R_t = \alpha_0 + \alpha_r r_t + \alpha_Y Y_t$. The reference rate r was determined by a Gaussian $A_0(2)$ model and the credit factor Y was assumed to follow a mean-reverting Gaussian diffusion. Upon estimating the models using Lehman Brothers data, they found that (after accommodating interest rate risk) higher leverage increases the default-adjusted discount rate, with leverage-related credit risk being more pronounced for long- than for short-dated corporate bonds. Bakshi, Madan, and Zhang [2001b] compare the relative fits of reduced-form models under various recovery timing conventions in a recovery of face value model. Janosi, Jarrow, and Yildirim [2000] also study a two-factor model for R in which r follows a one-factor Gaussian process and the credit factor is related to the level of the *S&P500* equity index.

Affine models have also be used in the pricing of sovereign bonds. Merrick [1999] calibrates a discrete-time model (to Russian and Argentinian bonds) that can be reinterpreted as a model with a constant (state independent) intensity. More generally, Keswani [2002] and Pagès [2000] apply special cases of the recovery of market value model to data on Latin American Brady bonds, and Dullmann and Windfuhr [2000] apply a similar framework to price European government credit spreads under the EMU.

These models presume that holders of sovereign debt face a single credit event – default with liquidation upon default – and, in particular, do not allow for restructurings and the associated write-downs of face value. Duffie, Pedersen, and Singleton [2002] develop a model allowing for both write-downs and for bonds to have idiosyncratic credit factors. After fitting this model to data on Russian MinFin bonds around the Russian default in August, 1998, they found substantial evidence for significant write-downs after the 1998 default.

5.2 Structural Models

Empirical implementation of structural models requires one to confront more directly the (often very) complex capital structures of issuers. Among the issues that must be addressed are: the components of the capital structure to be included in A and F ; measurement of asset volatility σ_A and the correlation ρ between A and r ; and the choice of recovery ratios. The empirical implementations of structural models have varied widely in their resolutions of these issues.¹⁹

Jones, Mason, and Rosenfeld [1984] implemented the Merton model for a sample of callable coupon bonds for the sample period 1977 - 1981. They found absolute *pricing* errors of 8.5%: model prices were too high so spreads were too narrow. Ogden [1987] looked at

¹⁹As a precursor to econometric implementation of structural models, Huang and Huang [2000] inquire whether the credit risk inherent in structural models is likely to account for the observed spreads in investment grade bonds. Calibrating a structural model to both balance sheet information and historical default rates simultaneously, they conclude that, within their models, credit risk accounts for a small fraction of the observed spreads for investment grade bonds. Their models are more successful at explaining the larger spreads for junk bonds.

primary market prices for bonds over the period 1973 - 1985. The Merton model under-predicted spreads by an average of 104 basis points. A major limitation of both of these studies was their use of callable bond prices.

More recently, Lyden and Sariniti [2000] used Bridge data, which provides actual transactions prices on non-callable bonds for both financial and non-financial firms. For the Merton model they found mean absolute errors in yield spreads of 80+ basis points. The model-implied spreads were particularly low (bonds were over-priced) for small firms and long maturities. For a two-factor Longstaff-Schwartz model, in which default occurs when A falls below the par value of outstanding bonds and any recovery is a constant fraction w of par value, they obtained roughly the same results. Moreover, the findings for the L - S model were largely insensitive to the assumed value of ρ . In interpreting these results, one should perhaps bear in mind that their sample included both financial and non-financial firms, with quite heterogeneous leverage ratios.

Ericsson and Reneby [2001] reach a more optimistic conclusion about the fit of a structural model with endogenous default and leverage ratios that reflect growth in both debt and equity values. Using maximum likelihood methods to estimate the parameters of their firm-value process, they obtain unbiased out-of-sample spread predictions of yield spreads for the non-callable debt of a small sample of 50 industrial firms.

The most comprehensive empirical comparison of structural models to date is provided by Eom, Helwege, and Huang [2001], who examine versions of the Merton, Geske [1977], Leland and Toft [1996], Longstaff and Schwartz [1995], and Collin-Dufresne and Goldstein [2001b] models. They restrict their sample to industrial firms with relatively simple capital structures, comprised largely of equity and non-callable debt. Consistent with previous studies, the Merton model predicts spreads that are too small, as does Geske's model though to a lesser extent. Interestingly, the Leland-Toft, Longstaff-Schwartz, and Collin-Dufresne-Goldstein models all tend to *over-predict* spreads, though in different ways. The Leland-Toft model over-predicts spreads for virtually all ratings and maturities; the Longstaff-Schwartz model gives excessive spreads for relatively risky bonds, while yielding spreads that are too small for relatively safe bonds; and the Collin-Dufresne-Goldstein model with mean-reverting leverage ratios reduces the under-prediction of spreads on safe bonds, while still over-predicting spreads on average. Eom, Helwege, and Huang [2001] conclude that the major challenge to improving the fit of structural models is to "raise the average predicted spread relative to the Merton model, without overstating the risks associated with volatility, leverage or coupon. (page 3)"

6 *DTSMs* and Derivatives Prices

With the growing availability of time-series data on the implied volatilities on fixed-income derivatives, comparisons of *DTSM*-implied to market prices of derivatives is increasingly being used in assessing goodness-of-fit. Particular attention has been given to the pricing of caps/floors and swaptions in the LIBOR/swap markets, no doubt in part because of the size and importance of these markets.

6.1 Risk-Factor Distributions and Derivatives Pricing

Much of the recent literature applying *DTSMs* to the pricing of derivatives has focused on two features of the distributions of swap rates and implied volatilities on LIBOR-based derivatives. First, Heidari and Wu [2001] and Collin-Dufresne and Goldstein [2001a] find that there is substantial variation in the implied volatilities of swaptions that appears to be unrelated to variation in the underlying swap rates. That is, it appears as if there are risk factors underlying changes in derivatives prices that are independent of those underlying shifts in the swap curve. To some the presence of “unspanned factors” is puzzling, because swaptions are derivative contracts based on the underlying swap rates.

Second, though financial theory predicts a close link between the prices of caps and swaptions (as they are both LIBOR-based derivatives), developing a model that simultaneously prices both contracts has proved challenging. Explanations for this “swaption/cap puzzle” often focus on the nature of the model-implied factor volatilities and/or correlations and their roles in determining prices. For instance, Rebonato and Cooper [1997] and Longstaff, Santa-Clara, and Schwartz [2001a] compare the correlations among forward swap rates with those implied by low-dimensional factor models and find that the correlations implied by the models are much larger than those in the data. Brown and Schaefer [1999] and Carverhill [2002] find similar results using Treasury strip yields.²⁰

We can anticipate the difficulty standard *DTSMs* will have in matching yield correlations by comparing historical and model-implied correlations among weekly changes in the yield spreads for non-overlapping segments of the U.S. dollar swap yield curve. The correlation “3-2/4-3,” for example, in Table 3 represents the correlation of daily changes in the 3yr-2yr swap spread with changes in the 4yr-3yr spread. The rows labeled “2 PC” and “4 PC” present the corresponding correlations for fitted spreads from projections onto the first two and four principal components (PCs), respectively. Notably, even using four PC’s the segment correlations are larger than their sample counterparts, and the match is much worse using only two PC’s (in the spirit of a two-factor *DTSM*).

Segment	3-2/4-3	4-3/5-4	5-4/7-5	7-5/10-7
Historical	0.34	0.09	0.13	0.14
2 PC	0.99	0.99	0.99	0.99
4 PC	0.81	0.96	0.84	0.32

Table 3: Correlations of changes in swap yield spreads for various yield-curve segments. 10-7,3-2, for example, indicates the correlation between changes in the 10-7 year yield spread and the 3-2 year yield spread.

Not surprisingly, when we compute model-implied segment correlations from the affine $A_M(N)$ models with $N \leq 3$, using swap data, they are all substantially larger than their

²⁰It is well known that a small number of principal components explains well over 90% of the variation in yields across the maturity spectrum for U.S. treasury yields (e.g., Litterman and Scheinkman [1991]), as well as for many foreign bond markets (e.g., Singleton [1995], Driessen, Melenberg, and Nijman [2000]). This evidence serves as a primary motivation for examining low-dimensional state vectors in empirical studies of *DTSMs*.

historical counterparts. The same is true for the fitted, relative to the historical, treasury yields from Ahn, Dittmar, and Gallant [2002]’s study of QG models.

If swaptions and caps have different sensitivities to a model’s (in)ability to match yield curve segment or forward-rate correlations, then this could resolve the pricing puzzles. However, the literature is not fully in agreement about the relative responses of the prices of caps and swaptions to changes in factor volatilities or correlations. After addressing the pricing of swaptions in $DTSMs$, we take up in more depth the relative pricing of caps and swaptions.

6.2 Implications for Pricing Swaptions

Focusing first on the “unspanned factors” puzzle, this finding is not, of course, logically inconsistent with arbitrage-free pricing. Indeed, Collin-Dufresne and Goldstein [2001a] provide one explanation within an arbitrage-free, affine $DTSM$ based on the possibility that some elements of $\gamma_Y(T - t)$ in (21) are zero for all $T - t$. This possibility, which requires that $N \geq 3$, arises when there are factors driving the volatility of r that are not “spanned” by the term structure of swap yields. The structure of volatility risk premiums might provide an alternative explanation.

Empirical work addressing the fit of $DTSMs$ to the joint distributions of swap and swaption prices has been limited. Upon fitting an $A_3(3)$ model (with independent factors) to historical swap yields, Jagannathan, Kaplan, and Sun [2001] find that their model is incapable of accurately pricing caps and swaptions.²¹ However, in the light of the preceding discussion, reliable pricing of swaptions would seem to depend on using swaption data in estimation in order to “pick up” the effectively unspanned factors. This is confirmed by Umantsev [2001], who estimated $A_M(3)$ models using data on swap rates and swaption volatilities simultaneously. He finds that $A_M(3)$ models, for $M = 1, 2$, fit the data notably better than an $A_3(3)$ model. Moreover, as anticipated by the descriptive findings of Heidari and Wu [2001], the third factor is related more to volatility in the swaption market than to “curvature” in the swap curve (the more typical third factor in $DTSMs$ fit to yield data alone).

The latter evidence suggests that low-dimensional $DTSMs$ (say N equal to 3 or 4) may well be capable of reliably pricing swaptions. This appears to be at odds with the arguments in Rebonato [1999] and Driessen, Klaassen, and Meleberg [2000].

6.3 Implications for the Relative Pricing of Caps and Swaptions

Where matching the correlation structure does seem to matter is in resolving the “swaptions/caps” pricing puzzle. Elaborating on the nature of this puzzle, many have found that model-implied volatilities extracted from cap prices tend to be larger than those backed out from swaption prices. One interpretation of this puzzle is based on the observation that a

²¹ Since swaptions are effectively options on coupon-paying bonds, closed-form solutions for the relevant exercise probabilities are generally not known, even in the affine case. Chen and Scott [1995] develop pricing relations that are specific to multi-factor $A_N(N)$ (CIR) models, and are used by Jagannathan, Kaplan, and Sun [2001]. Singleton and Umantsev [2002] and Collin-Dufresne and Goldstein [2002] develop accurate, approximate pricing formulas for the entire family of affine diffusion models.

cap can be viewed as a portfolio of options on forward LIBOR rates, whereas a swaption can be viewed as an option on a portfolio of forward LIBOR rates. As such, cap prices are relatively insensitive to the correlation structure of forward LIBOR rates, whereas the swaption prices depend crucially on the correlation structure. Indeed, a one-factor model can be calibrated exactly to all at-the-money cap prices, but it likely mis-price swaptions because forward rates are perfectly correlated in such a model.

Hull and White [1999] and Longstaff, Santa-Clara, and Schwartz [2001b] calibrate multi-factor forward-rate based models to (select portions of) the historical correlation structure of the forward rates²² and explore the relative pricing of caps and swaptions. Hull and White [1999] find that (for a single trading day) at least three factors are needed in order to rationalize observed caps and swaptions prices, whereas Longstaff, Santa-Clara, and Schwartz [2001b] find that, in order to explain caps and swaptions prices over a window of several months, at least four factors are needed. Both studies document periods of significant relative mispricings of caps and swaptions relative to the predictions of their models. Similar findings are reported by Jagannathan, Kaplan, and Sun [2001] and Fan, Gupta, and Ritchken [2001]. The ultimate resolution of this “swaptions/caps puzzle” may require time-varying correlations and possibly factors affecting the volatility of yields that do not affect bond prices. Collin-Dufresne and Goldstein [2001c] posit a “switching correlation” affine model with both of these features and find that it has features consistent with historical swaption and cap volatilities, but they do not formally calibrate their model to historical data.

7 Concluding Remarks

Our overview of the empirical fit of *DTSMs* has underscored several successes, while highlighting several challenges for future research. Specifically, Affine- and Quadratic-Gaussian *DTSMs* are evidently capable of resolving the puzzles associated with the rejections of the expectations hypothesis. At the same time, for both families of models, there appears to be a “tension” between matching properties of (i) the conditional mean, (ii) the conditional volatilities, and (iii) the risk premiums (Duffee [2002], Dai and Singleton [2002b]). Within the family of affine *DTSMs*, clearly we must have $M > 0$ in order for there to be stochastic volatility. However, accommodating stochastic volatility in this manner seems to conflict with matching the first-moment properties of yields. In principle, Duarte [2001]’s extended market price of risk (24) may partially relax this tension, because it allows for richer \mathbb{P} -dynamics of the M volatility factors in $A_M(N)$ models.

Similarly, Ahn, Dittmar, and Gallant found that their most flexible *QG* model exhibited less volatility than was observed historically (as reflected in their semi-non-parametric model), both in terms of level and persistence. Thus, it seems that still undiscovered is a tractable *DTSM* that simultaneously matches *LPY* and *CVY*, at maximum likelihood estimates of the parameters.

We also find it interesting that regime-switching paired with (22) and single-regime,

²²Hull and White [1999] is based on the Libor Market Model, whereas Longstaff, Santa-Clara, and Schwartz [2001b] is based on a discrete approximation of the string model proposed by Santa-Clara and Sornette [2001]. Both studies use a calibration method similar to a principal components analysis.

affine and QG models paired with (23) seem to perform comparably in matching the pattern LPY . It seems that, implicitly, the structure of preferences, and the economy more generally, in these formulations are quite different and a better understanding of these differences would likely improve our understanding of interest rate determination. Some hints about the links between model formulation and the macro economy come from several recent studies. Duffee [2001] explores various macroeconomic interpretations of the finding that the slope of the yield curve forecasts excess returns, the finding that underlies the need for richer risk premiums than (22) in affine models. Bansal and Zhou [2002] find that the regime switching process inferred from bond prices appears to be strongly associated with NBER business cycle indicators for the U.S.. This is consistent with the findings of Ang and Bekaert [2001] from a regime-switching model estimated without the structure of a pricing model. Further exploration of the links between factor risk premiums, regime identities, and the business cycle seems warranted.

The recent efforts at developing discrete-time $DTSMs$, motivated by extant continuous-time formulations, also raise interesting questions for future research. Along some dimensions, the discrete-time models offer more flexibility in specifying the dynamics of the state process. However, specifications of the market prices of risk appear to be richer in the continuous-time formulations. Does more flexibility in modeling Y reduce the need for richer specifications of Λ ?

Turning to the recent empirical work on reduced-form models of defaultable bonds, a natural next step in this literature would be to estimate a model in which $\lambda^{\mathbb{Q}}L^{\mathbb{Q}}$ is a affine function of (possibly multiple) observed credit factors, as well as latent factors. The findings in Collin-Dufresne, Goldstein, and Martin [2001] that most of the variation in credit spreads is independent of a long list of observed macroeconomic and balance sheet variables suggests that supply/demand factors indigenous to the corporate market may be playing a key role in spread variation. Duffie and Singleton [1997] report a similar result for swap spreads. Latent factors would allow modelers to capture the substantial residual variation in the spread s_t that is not captured by the various observable economic factors examined in these studies.

As better fitting models are developed, further analysis of economic forces underlying variation in spreads would also be beneficial. We have, in the end, a quite limited understanding of the determinants of spreads, particularly those in swap and high-grade corporate bond markets. The findings in Eom, Helwege, and Huang [2001] clarify some of the challenges facing structural modeling. Looking beyond the default process, are there unmodeled supply and demand forces affecting spreads and, if so, what are these forces? And what is the role of “liquidity” in spread determination?

Finally, the recent literature on pricing derivatives using $DTSMs$ illustrates the enormous potential for new insights from using derivatives data in model estimations.

References

- Ahn, C. and H. Thompson (1988). Jump Diffusion Processes and Term Structure of Interest Rates. *Journal of Finance* 43, 155–174.
- Ahn, D. and B. Gao (1999). A Parametric Nonlinear Model of Term Structure Dynamics. *Review of Financial Studies* 12, 721–762.
- Ahn, D.-H., R. F. Dittmar, and A. R. Gallant (2002). Quadratic Gaussian Models: Theory and Evidence. forthcoming, *Review of Financial Studies*.
- Ahn, D.-H., R. F. Dittmar, B. Gao, and A. R. Gallant (2000). Purebred or Hybrid? Reproducing the Volatility in Term Structure Dynamics. working paper, University of North Carolina.
- Ait-Sahalia, Y. (1996). Testing Continuous-Time Models of the Spot Interest Rate. *Review of Financial Studies* 9, 385–426.
- Ait-Sahalia, Y. (2002). Closed-Form Likelihood Expansions for Multivariate Diffusions. Working paper, Princeton University.
- Andersen, T. G. and J. Lund (1997). Estimating Continuous-Time Stochastic Volatility Models of the Short-Term Interest Rate. *Journal of Econometrics* 72, 343–377.
- Andersen, T. G. and J. Lund (1998). Stochastic Volatility and Mean Drift In the Short Term Interest Rate Diffusion: Sources of Steepness, Level and Curvature in the Yield Curve. Working paper, Northwestern University.
- Anderson, R. W. and S. M. Sundaresan (1996). The Design and Valuation of Debt Contracts. *Review of Financial Studies* 9, 37–68.
- Ang, A. and G. Bekaert (2001). Regime Switches in Interest Rates. forthcoming, *Journal of Business and Economic Statistics*.
- Ang, A. and M. Piazzesi (2000). A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables. Working Paper, Columbia University.
- Artzner, P. and F. Delbaen (1995). Default Risk Insurance and Incomplete Markets. *Mathematical Finance* 5, 187–195.
- Back, K. (1996). Yield Curve Models: A Mathematical Review. In I. Nelkin (Ed.), *Option Embedded Bonds: Price Analysis, Credit Risk, and Investment Strategies*. Irwin.
- Backus, D., S. Foresi, A. Mozumdar, and L. Wu (2001). Predictable Changes in Yields and Forward Rates. *Journal of Financial Economics* 59, 281–311.
- Backus, D., S. Foresi, and C. Telmer (1998). Discrete-Time Models of Bond Pricing. Working paper, New York University.
- Backus, D., S. Foresi, and S. Zin (1998). Arbitrage Opportunities in Arbitrage-Free Models of Bond Pricing. *Journal of Business and Economic Statistics* 16, 13–26.
- Backus, D. K. and S. E. Zin (1994). Reverse Engineering the Yield Curve. NBER Working Paper #4676.

- Bakshi, G., D. Madan, and F. Zhang (2001a). Investigating the Sources of Default Risk: Lessons from Empirically Evaluating Credit Risk Models. Working Paper, University of Maryland.
- Bakshi, G., D. Madan, and F. Zhang (2001b). Understanding the Role of Recovery in Default Risk Models: Empirical Comparisons and Implied Recovery Rates. Working Paper, Federal Reserve Board.
- Balduzzi, P., S. R. Das, and S. Foresi (1998). The Central Tendency: A Second Factor in Bond Yields. *Review of Economics and Statistics* 80, 62–72.
- Balduzzi, P., S. R. Das, S. Foresi, and R. K. Sundaram (1996). A Simple Approach to Three Factor Affine Term Structure Models. *Journal of Fixed Income* 6, 43–53.
- Balduzzi, P. and Y. Eom (2000). Non-linearities in U.S. Treasury Rates: A Semi-Nonparametric Approach. working paper, Boston College.
- Bansal, R. and H. Zhou (2002). Term Structure of Interest Rates with Regime Shifts. forthcoming, *Journal of Finance*.
- Beaglehole, D. R. and M. S. Tenney (1991). General Solutions of Some Interest Rate-Contingent Claim Pricing Equations. *Journal of Fixed Income*, 69–83.
- Bekaert, G. and R. Hodrick (2001). Expectations Hypotheses Tests. *Journal of Finance* 56, 1357–1394.
- Bekaert, G., R. Hodrick, and D. Marshall (1997). On Biases in Tests of the Expectations Hypothesis of the Term Structure of Interest Rates. *Journal of Financial Economics* 44, 309–348.
- Bekaert, G., R. Hodrick, and D. Marshall (2001). Peso Problem Explanations for Term Structure Anomalies. *Journal of Monetary Economics* 48, 241–270.
- Bielecki, T. and M. Rutkowski (2001). Modeling of the Defaultable Term Structure: Conditionally Markov Approach. Working paper, The Northeastern Illinois University and Warsaw University of Technology.
- Black, F. and J. Cox (1976). Valuing Corporate Securities: Liabilities: Some Effects of Bond Indenture Provisions. *Journal of Finance* 31, 351–367.
- Black, F., E. Derman, and W. Toy (1990). A One-Factor Model of Interest Rates and Its Application to Treasury Bond Options. *Financial Analysts Journal*, 33–39.
- Black, F. and P. Karasinski (1991, July-August). Bond and Option Pricing when Short rates are Lognormal. *Financial Analysts Journal*, 52–59.
- Black, F. and M. Scholes (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* 81, 637–654.
- Bobadilla, G. (1999). Choose the Right Error in Term Structure Models. Working Paper, CEMFI.
- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics* 31, 307–327.

- Boudoukh, J., M. Richardson, R. Stanton, and R. F. Whitelaw (1998). The Stochastic Behavior of Interest Rates: Implications from a Multifactor, Nonlinear Continuous-Time Model. Working Paper, New York University.
- Brace, A., D. Gatarek, and M. Musiela (1997). The Market Model of Interest Rate Dynamics. *Mathematical Finance* 7, 127–154.
- Brandt, M. and P. Santa-Clara (2002). Simulated Likelihood Estimation of Diffusions with an Application to Exchange Rate Dynamics in Incomplete Markets. *Journal of Financial Economics* 63, 161–210.
- Brandt, M. and A. Yaron (2001). Time-Consistent, No-Arbitrage Models of the Term Structure. Working Paper, Wharton School.
- Brenner, R. J., R. H. Harjes, and K. F. Kroner (1996). Another look at Models of the short Term Interest Rate. *Journal of Financial & Quantitative Analysis* 31, 85–107.
- Brito, R. and R. Flores (2001). A Jump-Diffusion Yield-Factor Model of Interest Rates. Working Paper, EPGE/FGV.
- Briys, E. and F. de Varenne (1997). Valuing Risky Fixed Rate Debt: An Extension. *Journal of Financial and Quantitative Analysis* 32, 239–248.
- Brown, R. and S. Schaefer (1999). Why Long Forward Interest Rates (Almost) Always Slope Downwards. Working Paper, London Business School.
- Brown, R. H. and S. M. Schaefer (1994). The Term Structure of Real Interest Rates and the Cox, Ingersoll and Ross Model. *Journal of Financial Economics* 35, 3–42.
- Brown, S. J. and P. H. Dybvig (1986). Empirical Implications of the Cox, Ingersoll, Ross Theory of the Term Structure of Interest Rates. *Journal of Finance* 41, 143–172.
- Buraschi, A. and F. Corielli (2000). Staying Ahead of the Curve: Model Risk and the Term Structure. Working Paper, London Business School.
- Buraschi, A. and A. Jiltsov (2001). How Large is the Inflation Risk Premium? A Monetary Model of the Term Structure. Working Paper, London Business School.
- Campbell, J. Y. and R. J. Shiller (1991). Yield Spreads and Interest Rate Movements: A Bird’s Eye View. *Review of Economic Studies* 58, 495–514.
- Carrasco, M., M. Chernov, J. Florens, and E. Ghysels (2002). Estimating Diffusions with a Continuum of Moment Conditions. Working Paper, University of North Carolina.
- Carverhill, A. (2002). Predictability and the Dynamics of Long Forward Rates. Working Paper, University of Hong Kong.
- Cathcart, L. and L. El-Jahel (1998). Valuation of Defaultable Bonds. *The Journal of Fixed Income*, June, 66–78.
- Chacko, G. and S. Das (2001). Pricing Interest Rate Derivatives: A General Approach. Working Paper, forthcoming, Review of Financial Studies.
- Chapman, D. and N. Pearson (2001). What Can be Learned From Recent Advances in Estimating Models of the Term Structure? *Financial Analysts Journal* July/August, 77–95.

- 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, U.K.
- Chen, R. and L. Scott (1993). Maximum Likelihood Estimation For a Multifactor Equilibrium Model of the Term Structure of Interest Rates. *Journal of Fixed Income* 3, 14–31.
- Chen, R. and L. Scott (1995). Interest Rate Options in Multifactor Cox-Ingersoll-Ross Models of the Term Structure. *Journal of Fixed Income Winter*, 53–72.
- Collin-Dufresne, P. and R. S. Goldstein (1999). On the Term Structure of Interest Rates in the Presence of Reflecting Boundaries. Working Paper, Washington University, St. Louis.
- Collin-Dufresne, P. and R. S. Goldstein (2001a). Do bonds span the fixed income markets? Theory and Evidence for Unspanned Stochastic Volatility. Working paper, GSIA, Carnegie Mellon University, and Ohio-State University.
- Collin-Dufresne, P. and R. S. Goldstein (2001b). Do Credit Spreads Reflect Stationary Leverage Ratios? *Journal of Finance* 56, 1929–1957.
- Collin-Dufresne, P. and R. S. Goldstein (2001c). Stochastic Correlation and the Relative Pricing of Caps and Swaptions in a Generalized Affine Framework. Working Paper, Carnegie Mellon University.
- Collin-Dufresne, P. and R. S. Goldstein (2002). Pricing Swaptions within the Affine Framework. Working Paper, Journal of Derivatives, forthcoming.
- Collin-Dufresne, P., R. S. Goldstein, and J. Martin (2001). The Determinants of Credit Spread Changes. *Journal of Finance* 56, 2177–2208.
- Collin-Dufresne, P. and B. Solnik (2000). On the Term Structure of Default Premia in the Swap and LIBOR Markets. Working Paper, forthcoming, Journal of Finance.
- Constantinides, G. (1992). A Theory of the Nominal Term Structure of Interest Rates. *Review of Financial Studies* 5, 531–552.
- Cox, J., J. Ingersoll, and S. Ross (1980). An Analysis of Variable Loan Contracts. *Journal of Finance* 35, 389–403.
- Cox, J. C., J. E. Ingersoll, and S. A. Ross (1985). A Theory of the Term Structure of Interest Rates. *Econometrica* 53, 385–408.
- Dai, Q. (2002). Asset Pricing in a Neoclassical Model with Limited Participation. working paper, New York University.
- Dai, Q. and K. Singleton (2000). Specification Analysis of Affine Term Structure Models. *Journal of Finance* LV, 1943–1978.
- Dai, Q. and K. Singleton (2002a). Fixed-Income Pricing. Working Paper, New York University.
- Dai, Q. and K. J. Singleton (2002b). Expectations Puzzles, Time-varying Risk Premia, and Dynamic Models of Term Structure. forthcoming, Journal of Financial Economics.

- Das, S. (2002). The Surprise Element: Jumps in Interest Rates. *Journal of Econometrics* 106, 27–65.
- Das, S. and S. Foresi (1996). Exact Solutions for Bond and Option Prices with Systematic Jump Risk. *Review of Derivatives Research* 1, 7–24.
- Driessen, J., P. Klaassen, and B. Meleberg (2000). The Performance of Multi-Factor Term Structure Models for Pricing and Hedging Caps and Swaptions. Working Paper, Tilburg University.
- Driessen, J., B. Melenberg, and T. Nijman (2000). Common Factors in International Bond Returns. Working Paper, Working Paper, Tilburg University.
- Duan, J.-C. and J.-G. Simonato (1999, September). Estimating and Testing Exponential-Affine Term Structure Models by Kalman Filter. *Review of Quantitative Finance and Accounting* 13.
- Duarte, J. (2001). Evaluating Alternative Risk Preferences in Affine Term Structure Models. Working paper, University of Chicago.
- Duffee, G. (1998). The Relation Between Treasury Yields and Corporate Bond Yield Spreads. *Journal of Finance* 53, 2225–2242.
- Duffee, G. (1999). Estimating the Price of Default Risk. *The Review of Financial Studies* 12, 197–226.
- Duffee, G. (2001). Why Does the Slope of the Yield Curve Forecast Excess Returns? Working Paper, University of California, Berkeley.
- Duffee, G. and R. Stanton (2001). Estimation of Dynamic Term Structure Models. Working Paper, University of California, Berkeley.
- Duffee, G. R. (2002). Term Premia and Interest Rate Forecasts in Affine Models. *Journal of Finance* 57, 405–443.
- Duffie, D. (1996). *Dynamic Asset Pricing Theory, 2nd edition*. Princeton University Press, Princeton, N.J.
- Duffie, D. (1998). Defaultable Term Structure Models with Fractional Recovery of Par. Working Paper, Graduate School of Business, Stanford University.
- Duffie, D., D. Filipovic, and W. Schachermayer (2002). Affine Processes and Applications in Finance. Working Paper, Stanford University.
- Duffie, D. and M. Huang (1996). Swap Rates and Credit Quality. *Journal of Finance* 51, 921–949.
- Duffie, D. and R. Kan (1996). A Yield-Factor Model of Interest Rates. *Mathematical Finance* 6, 379–406.
- Duffie, D. and D. Lando (2001). Term Structures of Credit Spreads with Incomplete Accounting Information. Working Paper, Graduate School of Business, Stanford University, forthcoming: *Econometrica*.
- Duffie, D. and J. Liu (2001). Floating-Fixed Credit Spreads. Working Paper, forthcoming: *Financial Analysts Review*.

- Duffie, D., J. Pan, and K. Singleton (2000). Transform Analysis and Asset Pricing for Affine Jump-Diffusions. *Econometrica* 68, 1343–1376.
- Duffie, D., L. Pedersen, and K. Singleton (2002). Modelling Credit Spreads on Sovereign Debt: A Case Study of Russian Bonds. *forthcoming, Journal of Finance*.
- Duffie, D. and K. Singleton (1997). An Econometric Model of the Term Structure of Interest Rate Swap Yields. *Journal of Finance* 52, 1287–1321.
- Duffie, D. and K. Singleton (1999). Modeling Term Structures of Defaultable Bonds. *Review of Financial Studies* 12, 687–720.
- Dullmann, K. and M. Windfuhr (2000). Credit Spreads Between German and Italian Sovereign Bonds— Do Affine Models Work? Working Paper, University of Mannheim.
- Eom, Y., J. Helwege, and J. Huang (2001). Structural Models of Corporate Bond Pricing: An Empirical Analysis. Working Paper, College of Business and Economics, Yonsei University, Seoul Korea.
- Ericsson, J. and J. Reneby (2001). The Valuation of Corporate Liabilities: Theory and Tests. Working Paper, McGill University.
- Evans, M. D. (2000, March). Regime Shifts, Risk and the Term Structure. Working Paper, Georgetown University.
- Fama, E. (1984a). The Information in the Term Structure. *Journal of Financial Economics* 13, 509–528.
- Fama, E. (1984b). Term Premiums in Bond Returns. *Journal of Financial Economics* 13, 529–546.
- Fama, E. F. and R. R. Bliss (1987). The Information in Long-Maturity Forward Rates. *American Economic Review* 77(4), 680–692.
- Fan, R., A. Gupta, and R. Ritchken (2001). On Pricing and Hedging in the Swaption Market: How Many Factors Really? Working Paper, Case Western Reserve University.
- Fisher, M. (1998, January). A Simple Model of the Failure of the Expectations Hypothesis. Working Paper, Federal Reserve Board.
- Fisher, M. and C. Gilles (1996). Estimating Exponential Affine Models of the Term Structure. Working paper, Federal Reserve Board, Washington DC.
- Fleming, M. J. and E. M. Remolona (1999, May). The Term Structure of Announcement Effects. FRB New York Staff Report No. 76.
- Gallant, A. R. and G. Tauchen (1996). Which Moments to Match? *Econometric Theory* 12, 657–681.
- Gallant, A. R. and G. Tauchen (1998). Reprojecting Partially Observed Systems with Application to Interest Rate Diffusions. *Journal of American Statistical Association* 93, 10–24.
- Gerlach, S. and F. Smets (1997). The Term Structure of Euro-Rates: Some Evidence in Support of the Expectations Hypothesis. *Journal of International Money and Finance* 16, 305–321.

- Geske, R. (1977). The Valuation of Corporate Liabilities as Compound Options. *Journal of Financial and Quantitative Analysis* 12, 541–552.
- Gibson, R., F. L'Habitant, and D. Talay (2001). Modeling the Term Structure of Interest Rates: A Review of the Literature. Working Paper, HEC.
- Goldstein, R. S. (2000). The Term Structure of Interest Rates as a Random Field. *Review of Financial Studies* 13, 365–384.
- Gourieroux, C., A. Monfort, and V. Polimenis (2002). Affine Term Structure Models. Working Paper, CREST.
- Gray, S. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics* 42, 27–62.
- Grinblatt, M. (1994). An Analytic Solution for Interest Rate Swap Spreads. Working Paper, UCLA Anderson Graduate School of Management.
- Hamilton, J. (1988). Rational-Expectations Econometric Analysis of Changes in Regime: An Investigation of the Term Structure of Interest Rates. *Journal of Economic Dynamics and Control* 12, 385–423.
- Hardouvelis, G. (1994). The Term Structure Spread and Future Changes in Long and Short Rates in the G7 Countries. *Journal of Monetary Economics* 33, 255–283.
- Heath, D., R. Jarrow, and A. Morton (1992). Bond Pricing and the Term Structure of Interest Rates: A New Methodology. *Econometrica* 60, 77–105.
- Heidari, M. and L. Wu (2001). Are Interest Rate Derivatives Spanned by the Term Structure of Interest Rates? Working Paper, Fordham University.
- Hogan, M. and K. Weintraub (1993). The Lognormal Interest Rate Model and Eurodollar Futures. Working paper, Citibank, New York.
- Honore, P. (1998). *Ph.D. Dissertation: Five Essays on Financial Econometrics in Continuous-Time Models*. Aarhus: Aarhus School of Business.
- Huang, J. and M. Huang (2000). How Much of the Corporate-Treasury Yield Spread is Due to Credit Risk? A New Calibration Approach. Working Paper, Stanford University.
- Hull, J. and A. White (1987). The Pricing of Options on Assets with Stochastic Volatilities. *The Journal of Finance* XLII, 281–300.
- Hull, J. and A. White (1993). One-Factor Interest-Rate Models and the Valuation of Interest-Rate Derivative Securities. *Journal of Financial and Quantitative Analysis* 28, 235–254.
- Hull, J. and A. White (1999, August). Forward Rate Volatilities, Swap Rate Volatilities, and the Implementation of the LIBOR Market Model. Working paper, University of Toronto.
- Jagannathan, R., A. Kaplan, and S. Sun (2001). An Evaluation of Multi-factor CIR Models Using LIBOR, Swap Rates, and Swaption Prices. Working Paper, Northwestern University.

- Jamshidian, F. (1997). Libor and Swap Market Models and Measures. *Finance Stochastics* 1, 293–330.
- Janosi, T., R. Jarrow, and Y. Yildirim (2000). Estimated Expected Losses and Liquidity Discounts Implicit in Debt Prices. Working paper, Cornell University.
- Jarrow, R. and S. Turnbull (1995). Pricing Options on Financial Securities Subject to Default Risk. *Journal of Finance* 50, 53–86.
- Jarrow, R. A., D. Lando, and F. Yu (2000). Default Risk and Diversification: Theory and Applications. Working Paper, Cornell University.
- Jegadeesh, N. and G. G. Pennacchi (1996). The Behavior of Interest Rates Implied by the Term Structure of Eurodollar Futures. *Journal of Money, Credit, and Banking* 28.
- Johannes, M. (2000). A Nonparametric View of the Role of Jumps to Interest Rates. Working Paper, Columbia University.
- Jones, E., S. Mason, and E. Rosenfeld (1984). Contingent Claims Analysis of Corporate Capital Structures: An Empirical Investigation. *The Journal of Finance* 39, 611–625.
- Kennedy (1994). The term structure of interest rates as a Gaussian random field. *Mathematical Finance* 4, 247–258.
- Keswani, A. (2002). Estimating a Risky Term Structure of Brady Bonds. Working Paper, Lancaster University.
- Kim, J., K. Ramaswamy, and S. Sundaresan (1993). Does Default Risk in Coupons Affect the Valuation of Corporate Bonds? A Contingent Claims Model. *Financial Management* 22, 117–131.
- Kugler, P. (1997). Central Bank Policy Reaction and Expectations Hypothesis of the Term Structure. *International Journal of Financial Economics* 2, 217–224.
- Landen, C. (2000). Bond Pricing in a Hidden Markov Model of the Short Rate. *Finance and Stochastics* 4, 371–389.
- Lando, D. (1998). Cox Processes and Credit-Risky Securities. *Review of Derivatives Research* 2, 99–120.
- Langtieg, T. (1980). A Multivariate Model of the Term Structure. *Journal of Finance* 35, 71–97.
- Leippold, M. and L. Wu (2001). Design and Estimation of Quadratic Term Structure Models. Working paper, Fordham University.
- Leippold, M. and L. Wu (2002). Asset Pricing Under the Quadratic Class. *Journal of Financial and Quantitative Analysis* 37, 271–295.
- Leland, H. and K. Toft (1996). Optimal Capital Structure, Endogenous Bankruptcy, and the Term Structure of Credit Spreads. *Journal of Finance* 51, 987–1019.
- Leland, H. E. (1994). Corporate Debt Value, Bond Covenants, and Optimal Capital Structure. *Journal of Finance* XLIX, 1213–1252.
- Litterman, R. and J. Scheinkman (1991). Common Factors Affecting Bond Returns. *Journal of Fixed Income* 1, 54–61.

- Litterman, R., J. Scheinkman, and L. Weiss (1988). Volatility and the Yield Curve. Working Paper, Goldman Sachs.
- Liu, J. (1997). Generalized Method of Moments Estimation of Affine Diffusion Processes. Working Paper, Graduate School of Business, Stanford University.
- Liu, J., F. Longstaff, and R. Mandell (2001). The Market Price of Credit Risk: An Empirical Analysis of Interest Rate Swap Spreads. Working Paper, UCLA.
- Longstaff, F., P. Santa-Clara, and E. Schwartz (2001a). Throwing Away a Billion Dollars. *Journal of Financial Economics* 62, 39–66.
- Longstaff, F. A. (1989). A Nonlinear General Equilibrium Model of the Term Structure of Interest Rates. *Journal of Financial Economics* 2, 195–224.
- Longstaff, F. A., P. Santa-Clara, and E. S. Schwartz (2001b). The Relative Valuation of Caps and Swaptions: Theory and Empirical Evidence. *Journal of Finance* 56, 2067–2109.
- Longstaff, F. A. and E. S. Schwartz (1992). Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model. *Journal of Finance* 47, 1259–1282.
- Longstaff, F. A. and E. S. Schwartz (1995). Valuing Risky Debt: A New Approach. *The Journal of Finance* 50, 789–821.
- Lu, B. (2000). An Empirical Analysis of the Constantinides Model of the Term Structure. Working paper, University of Michigan.
- Lu, B. and G. Wu (2000). Implied Bivariate State Price Density. Working Paper, University of Michigan.
- Lyden, S. and D. Sariniti (2000). An Empirical Examination of the Classical Theory of Corporate Security Valuation. Working Paper.
- Madan, D. and H. Unal (1998). Pricing the Risks of Default. *Review of Derivatives Research* 2, 121–160.
- Martellini, L. and N. E. Karoui (2001). A Theoretical Inspection of the Market Price for Default Risk. Working Paper, University of Southern California.
- McCallum, B. T. (1994). Monetary Policy and the Term Structure of Interest Rates. NBER Working Paper No. 4938.
- Mella-Barral, P. and W. Perraudin (1997). Strategic Debt Service. *Journal of Finance* 52, 531–556.
- Merrick, J. J. (1999). Crisis Dynamics of Implied Default Recovery Ratios: Evidence From Russia and Argentina. Stern School of Business, New York University.
- Merton, R. (1970). A Dynamic General Equilibrium Model of the Asset Market and its Application to the Pricing of the Capital Structure of the Firm. Working Paper, Sloan School of Management, M.I.T.
- Merton, R. (1974). On The Pricing of Corporate Debt: The Risk Structure of Interest Rates. *The Journal of Finance* 29, 449–470.

- Miltersen, K. R., K. Sandmann, and D. Sondermann (1997, March). Closed Form Solutions for Term Structure Derivatives with Log-Normal Interest Rates. *Journal of Finance* 52(1), 409–430.
- Musiela, M. and M. Rutkowski (1997). *Martingale Methods in Financial Modelling*. Springer.
- Naik, V. and M. H. Lee (1997). Yield Curve Dynamics with Discrete Shifts in Economic Regimes: Theory and Estimation. Working paper, Faculty of Commerce, University of British Columbia.
- Nielsen, L. T., J. SaaRequejo, and P. Santa-Clara (1993). Default Risk and Interest Rate Risk: the Term Structure of Default Spreads. Working Paper, INSEAD, Fontainebleau, France.
- Ogden, J. (1987). Determinants of the Relative Interest Rate Sensitivities of Corporate Bonds. *Financial Management* 10, 22–30.
- Pagès, H. (2000). Estimating Brazilian Sovereign Risk from Brady Bond Prices. Working Paper, Bank of France.
- Pearson, N. D. and T. 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.
- Pedersen, A. (1995). A New Approach to Maximum Likelihood Estimation for Stochastic Differential Equations Based on Discrete Observations. *Scand J Statistics* 22, 55–71.
- Peterson, S., R. Stapleton, and M. Subrahmanyam (1998). An Arbitrage-free Two-factor Model of the Term Structure of Interest Rates: A Multivariate Binomial Approach. Working Paper, New York University.
- Piazzesi, M. (2001). An Econometric Model of the Term Structure with Macroeconomic Jump Effects. Working Paper, UCLA.
- Rebonato, R. (1999). On the Simultaneous Calibration of Multifactor Lognormal Interest Rate Models to Black Volatilities and to the Correlation Matrix. *Journal of Computational Finance* 2(4), 5–27.
- Rebonato, R. and I. Cooper (1997). The Limitations of Simple Two-Factor Interest Rate Models. *Journal of Financial Engineering* 5, 1 – 16.
- Roberds, W. and C. Whiteman (1999). Endogenous Term Premia and Anamolies in the Term Structure of Interest Rates: Explaining the Predictability Smile. *Journal of Monetary Economics* 44, 555–580.
- Sandmann, K. and D. Sondermann (1997). A note on the stability of lognormal interest rate models and the pricing of Eurodollar futures. *Mathematical Finance* 7, 119–125.
- Santa-Clara, P. and D. Sornette (2001). The Dynamics of the Forward Interest Rate Curve with Stochastic String Shocks. *Review of Financial Studies* 14, 149–185.
- Singleton, K. (1995). Yield Curve Risk Management for Government Bond Portfolios: An International Comparison. In W. Beaver and G. Parker (Eds.), *Risk Management Problems and Solutions*, pp. 295–322. McGraw Hill.

- Singleton, K. (2001). Estimation of Affine Asset Pricing Models Using the Empirical Characteristic Function. *Journal of Econometrics* 102, 111–141.
- Singleton, K. and L. Umantsev (2002). Pricing Coupon-Bond Options and Swaptions in Affine Term Structure Models. *Mathematical Finance*, forthcoming.
- Stanton (1997). A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk. *Journal of Finance* 52(5), 1973–2002.
- Steenkiste, R. J. V. and S. Foresi (1999). Arrow-Debreu Prices for Affine Models. Working paper, Soloman Smith Barney and Goldman Sachs.
- Sun, T. (1992). Real and Nominal Interest Rates: A Discrete-Time Model and Its Continuous-Time Limit. *Review of Financial Studies* 5, 581–611.
- Sun, T., S. Sundaresan, and C. Wang (1993). Interest rate swaps – An empirical investigation. *Journal of Financial Economics* 34, 77–99.
- Sundaresan, S. (2000). Continuous-Time Methods in Finance: A Review and an Assessment. *Journal of Finance* 55, 1569–1622.
- Tauren, M. (1999). A Model of Corporate Bond Prices with Dynamic Capital Structure. Working Paper, Indiana University.
- Umantsev, L. (2001). Econometric Analysis of European LIBOR-based Options with Affine Term Structure Models. Ph.D. Dissertation, Stanford University.
- Vasicek, O. (1977). An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics* 5, 177–188.
- Wachter, J. (2002). Habit Formation and Returns on Bonds and Stocks. Working Paper, New York University.
- Wu, S. (2002). How Do Changes In Monetary Policy Affect the Term Structure of Interest Rates? Working Paper, University of Kansas.
- Wu, T. (2000). Macro Factors and the Affine Term Structure of Interest Rates. Working Paper, Yale University.
- Yan, H. (2001). Dynamic Models of the Term Structure. *Financial Analysts Journal July/August*, 60–75.
- Zhou, C. (2000). The Term Structure of Credit Spreads with Jump Risk. Working Paper, forthcoming, *Journal of Banking and Finance*.
- Zhou, H. (2001). Jump-Diffusion Term Structure and Ito Conditional Moment Generator. Working Paper, Federal Reserve Board.